Salience and Consumer Choice *

Pedro Bordalo
Royal Holloway, University of London
pedro.bordalo@rhul.ac.uk

Nicola Gennaioli
Università Bocconi, IGIER, and CREI
nicola.gennaioli@unibocconi.it

Andrei Shleifer
Harvard University
shleifer@fas.harvard.edu

First draft October 2011, Revised July 2013

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 attributes, relative to that attribute's average level in the choice set (or more broadly, the choice context). Consumers attach disproportionately high weight to salient attributes and their choices are tilted toward goods with higher quality/price ratios. The model accounts for a variety of disparate evidence, including decoy effects and context-dependent willingness to pay. It also suggests a novel theory of misleading sales.

*We are grateful to David Bell, Tom Cunningham, Matt Gentzkow, Bengt Holmstrom, Daniel Kahneman, David Laibson, Drazen Prelec, Jan Rivkin, Josh Schwartzstein, Jesse Shapiro, Itamar Simonson, Dmitry Taubinski, Richard Thaler and four anonymous referees 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.

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. You again think the French wine is 50% better, but now it is 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 reflects a fundamental feature of decision making namely that the consumer's attention is drawn to – and his choice is shaped by – the most salient aspects in the choice context he faces. In this paper we present a parsimonious model of salience in decision making for riskless choice and show how it helps 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, these subjects 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, to an extent that is hard to account for through income effects (Hastings and Shapiro 2013).

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).

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 (2012a, hereafter BGS 2012a) apply this idea to 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 (2012a) 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.

In our model, a good's salient attributes are those that stand out in the sense of being furthest from their "reference good." In the basic version of the model, the reference good is defined as having the average level of each attribute of goods in the choice set. In Section 4, however, we extend the model to allow the reference good to also depend on the rationally expected prices of the goods in the choice set. In Thaler's beer example, the sunbather expects the beer price at the store to be the usual store price of beer, and the beer price at the resort to be the usual resort beer price. In the Hastings and Shapiro gasoline example, buyers approach the gas station having in mind rational expectations of gasoline prices.

For each good in the choice set, its salient attributes are those whose levels are unusual or surprising relative to the reference good. The consumer's attention is drawn to these attributes, which are then overweighted in the consumer's choice. In many situations, salience induces 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 force generates two broad types of context effects. The first occurs when the reference good is the average good in the choice set, which happens when actual and expected prices coincide. In this case, the quality to price ratio logic implies that consumers display higher price sensitivity, meaning a steeper tradeoff between quality and price, at low price levels. The model also provides new insights into the effects of adding to the choice set "decoy" goods with low quality price ratios. We find that such decoys asymmetrically boost the demand for high quality options, but not low for quality options.

The second class of context effects occurs when realized and expected prices differ: expected prices shape the perception of options under consideration through their influence on the reference price. The model predicts strong reactions to unexpected prices increases, which make prices more likely to become salient. It also generates a context-dependent willingness to pay through an anchoring-like mechanism.[1]

[1]Our approach is related to situations in which decision makers evaluate their options using mental

The context dependent effects created by salience lead to the central empirical implications of the model. In broad terms, consumers in our model are relatively insensitive to price differences between goods of different qualities after an *expected* uniform increase in prices. In contrast, consumers are very price sensitive when faced with *unexpected* uniform price increase. These central implications are summarized in Proposition 4 in Section 4.

Economists have tried several approaches to account for some of the experimental evidence we discuss here: the standard analysis of context effects is information-theoretic (Wernerfelt 1995, Kamenica 2008), while many behavioral models, which we review in Section 2.2, emphasise the dependence of choice on external reference points (Kahneman and Tverksy 1979, Simonson and Tversky 1992). The present model offers several advantages. It provides a tractable framework for a fundamental psychological mechanism, based on ex post attention allocation to well-defined salient features of the environment. It accounts for a broad range of context-dependent choices, both in riskless environments and in choice under risk (BGS 2012a). It provides new insight into how reference points shapes valuation, and as a result can account for evidence that is dumbfounding from the standard perspective, such as Thaler's beer example. Finally, it generates several distinct implications that give new insight into puzzling evidence in several applications.

The paper is organized as follows. Section 2 presents the model when context is given by the choice set, and establishes the central role of the quality price ratio in shaping salience and consumer decisions. Section 3 explores the context effects that arise from manipulations of the choice set, such as changes in price levels and violations of IIA such as decoy and compromise effects. Section 4 broadens the notion of context to include expected prices, and explores its implications for the demand for quality, including the effects of surprising price changes. Here we show how our model accounts for context dependent willingness to pay (the Thaler beer example). In Section 5, we describe several falsifiable predictions of the model when quality levels and/or price expectations are not observable. Section 6 presents a new theory of sales showing how discounts can mislead consumers. Section 7 concludes.

accounts (Thaler 1980). Recent research on the interplay of attention and choice includes Mullainathan (2002), Gennaioli and Shleifer (2010), Schwartzstein (2012), Gabaix (2012) and Woodford (2012). The marketing literature also stresses the effect on choice of the set of alternatives that come to the consumer's mind (see Roberts and Lattin 1997 for a review).

2 The Model

2.1 Setup

A consumer evaluates all $N > 1$ goods in a choice set $\mathbf{C} \equiv \{(q_k, p_k)\}_{k=1,\ldots,N}$. Each good k is characterized by its non-negative quality q_k and price p_k. We occasionally refer to good k using the vector (q_k, p_k) of its attributes. Quality and price are measured in dollars and known to the consumer. We discuss issues concerning the empirical measurement of quality in Section 5. In Appendix B, we extend the model to the case of goods having multiple quality attributes.

Absent salience distortions, a consumer values good k with a linear utility function:

$$u_k = q_k - p_k, \tag{1}$$

which attaches equal weights to quality and price. In Equation (1), utility is also measured in dollars. We later discuss alternative approaches to measuring attributes and utility.[2]

A salient thinker departs from (1) by inflating the relative weights attached to the attributes that he perceives to be more salient. As in BGS (2012a), we say that an attribute (quality or price) is salient for good k in the choice set $\mathbf{C}$ if this attribute "stands out" relative to the good's other attributes. Formally, denote by $(\overline{q}, \overline{p})$ the reference good consisting of average attributes $\overline{q} \equiv \sum_k q_k/N$ and $\overline{p} \equiv \sum_k p_k/N$ in $\mathbf{C}$. The salience of quality for a generic good k is then given by $\sigma(q_k, \overline{q})$, while the salience of price for good k is given by $\sigma(p_k, \overline{p})$. To define the salience function $\sigma(\cdot, \cdot)$, denote by a_k the level of an attribute (quality or price) for good k, and by $\overline{a}$ the attribute's average level in $\mathbf{C}$. We then have:

Definition 1 *The salience function $\sigma(\cdot, \cdot)$ is symmetric, continuous and satisfies:*

[2]Three remarks on assuming the linear utility function (1). First, adopting an additive representation of preferences allows us to apply the formalism we developed in BGS (2012a). Additive preferences are appropriate when quality and price are independent, see Keeney and Raiffa (1976). Appendix B extends the model to arbitrary weights on quality and price. We have not included the consumer's income w in the numeraire good, from which the consumer obtains utility $w - p_k$, because w is not an attribute of the good, so its valuation is not distorted by salience. Second, given separability of the utility function, it is natural to assume linearity in the observed price (measured in dollars) since in most consumer choice settings income effects are not too large. Third, the analysis could be extended to the case of a gain-loss utility, namely when consumers evaluate the utility of a good's quality and price relative to a reference level.

1) *Ordering. Let* $\mu = sgn(a_k - \overline{a})$*. Then for any* $\epsilon, \epsilon' \geq 0$ *with* $\epsilon + \epsilon' > 0$ *we have*

$$\sigma(a_k + \mu\epsilon, \overline{a} - \mu\epsilon') > \sigma(a_k, \overline{a}). \tag{2}$$

2) *Diminishing sensitivity. For any* $a_k, \overline{a} \geq 0$ *and all* $\epsilon > 0$ *we have:*

$$\sigma\left(a_k + \epsilon, \overline{a} + \epsilon\right) < \sigma\left(a_k, \overline{a}\right). \tag{3}$$

We say that, in the choice set $\mathbf{C}$, quality is salient for good k when $\sigma(q_k, \overline{q}) > \sigma(p_k, \overline{p})$, price is salient for good k when $\sigma(q_k, \overline{q}) < \sigma(p_k, \overline{p})$, and price and quality are equally salient when $\sigma(q_k, \overline{q}) = \sigma(p_k, \overline{p})$.

As we discuss in BGS (2012a), the properties of Definition 1 capture two key features of sensory perception. First, our perceptive apparatus is attuned to detect changes in stimuli. This is captured by ordering, whereby salience increases in contrast: the value a_k of an attribute is salient when it is very different from the average value $\overline{a}$ of the same attribute in the choice set. For instance, if a good is much more expensive than average, then its price is very salient. Second, changes in stimuli are perceived with diminishing sensitivity (Weber's law): formally, salience decreases as the value of an attribute uniformly increases for all goods. For instance, the salience of price falls when all prices become uniformly higher. At higher price levels, given price differences are less noticeable.

Ordering and diminishing sensitivity interact in determining salience. Suppose that the price p_k of the most expensive good goes up. By ordering, p_k becomes more salient. At the same time, the increase in p_k increases the average price level $\overline{p}$. By diminishing sensitivity, this reduces the salience of p_k. When, as in this case, ordering and diminishing sensitivity point in different directions, the tradeoff between them is pinned down by the specific salience function adopted. Although most of our results hold under the general Definition 1, in what follows we pin down this tradeoff by assuming that the salience function is homogeneous of degree zero:

A.0: The salience function satisfies ordering and homogeneity of degree zero, which is defined as $\sigma(\alpha \cdot a_k, \alpha \cdot \overline{a}) = \sigma(a_k, \overline{a})$ for all $\alpha > 0$.

Together with ordering, homogeneity of degree zero of the salience function implies diminishing sensitivity for positive attribute levels,[3] so that A.0 implies Definition 1 in this range.

Homogeneity of degree zero characterizes salience in an intuitive way: upon a variation in a good's attribute a_k, ordering dominates diminishing sensitivity if and only if the change in a_k is proportionally larger than the induced change in the average $\overline{a}$. 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.[4] Homogeneity of degree zero is also formally convenient, as it ensures that the salience ranking is invariant under linear transformations of the units (dollars) in which the attributes are measured. Thus, even though we defined salience to be a property of dollar attributes, under A.0. it can be defined over utils provided the latter are a scalar transformation of dollars.[5]

An example of a salience function satisfying homogeneity of degree zero, which we previously used in BGS (2012a), is:

$$\sigma\left(a_k, \overline{a}\right) = \frac{|a_k - \overline{a}|}{a_k + \overline{a}}, \tag{4}$$

for $a_k, \overline{a} \neq 0$, and $\sigma\left(0, 0\right) = 0$.[6]

[3]To see this, is it sufficient to note that ordering and homogeneity of degree zero imply that salience is an increasing function of the ratio $a_k/\overline{a}$ when $a_k > \overline{a}$ and of the ratio $\overline{a}/a_k$ when $a_k < \overline{a}$. Uniform increases in a_k and $\overline{a}$ then reduce salience in the two cases. Homogeneity of degree zero is stronger than diminishing sensitivity, and in particular excludes certain weak forms of the latter. For instance, the salience function $\sigma(x, y) = \frac{|x-y|}{x+y+\zeta}$, with $\zeta > 0$ satisfies Definition 1 but not homogeneity of degree zero (in fact, $\sigma(\alpha x, \alpha y) > \sigma(x, y)$ for $\alpha > 1$).

Homogeneity of degree zero of the salience function raises the issue of how to operationalize the ordering property when the level of an attribute is zero. To do so, we interpret $\sigma(a_k, 0)$ as $\lim_{\underline{z}\to 0} \sigma(a_k, \underline{z})$. Moreover, when comparing $\sigma(q_k, 0)$ and $\sigma(p_k, 0)$, we take the limit with the ratio of the $\underline{z}$ terms constant at 1. As a consequence, $\sigma(q_k, 0) > \sigma(p_k, 0)$ if and only if $q_k > p_k$, so ordering is preserved at zero attribute levels.

[4]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."

[5]Defining salience in terms of utils would not modify our analysis substantially if utils were an affine transformation of dollars. Specifically, the ordering and diminishing properties of Definition 1 would carry through from dollars to utils. This is no longer the case if utils are a nonlinear transformation of dollars.

[6]Because the model is defined for non-negative attribute levels, the specification of the salience function is slightly different from that in BGS (2012a): Definition 1 does not include a reflection property (for negative

Homogeneity of degree zero highlights the role of a good's quality/price ratio in determining the salience ranking of its attributes. Take a choice set $\mathbf{C}$ consisting of $N > 1$ goods. We then prove:

Proposition 1 *Let (q_k, p_k) be a good that neither dominates nor is dominated by the reference good $(\overline{q}, \overline{p})$, that is, $(q_k - \overline{q})(p_k - \overline{p}) > 0$. The following two statements are equivalent:*
1) *The higher quality or lower price of k relative to $(\overline{q}, -\overline{p})$ is salient iff $q_k/p_k > \overline{q}/\overline{p}$.*
2) *Salience is homogeneous of degree zero.*

Under the assumption A.0, salience favors goods with a high quality to price ratio. Proposition 1 captures a central intuition in our model: a good deal is attractive because it draws a consumer's attention to its advantage (high quality or low price) relative to its competitors.

To close the model, consider how salience distorts the valuation of a good. Given a salience function σ, a consumer ranks a good's attributes and distorts their utility weights as follows:

Definition 2 *The salient thinker's valuation of good k enhances the relative utility weight attached to the salient attribute (keeping constant the sum of weights attached to quality and price). Formally:*

$$u_k^S = \begin{cases} \frac{2}{1+\delta} \cdot q_k - \frac{2\delta}{1+\delta} \cdot p_k & \textit{if} \quad \sigma(q_k, \overline{q}) > \sigma(p_k, \overline{p}) \\ \frac{2\delta}{\delta+1} \cdot q_k - \frac{2}{\delta+1} \cdot p_k & \textit{if} \quad \sigma(q_k, \overline{q}) < \sigma(p_k, \overline{p}) \\ \qquad q_k - p_k & \textit{if} \quad \sigma(q_k, \overline{q}) = \sigma(p_k, \overline{p}) \end{cases} . \tag{5}$$

where $\delta \in (0, 1]$ decreases in the severity of salient thinking.

If quality is salient, the relative weight of quality increases, $\frac{2}{1+\delta} > 1$, and the relative weight of price decreases, $\frac{2\delta}{1+\delta} < 1$, as compared to the rational consumer's valuation. In this case, the salient thinker's price sensitivity is lower than the rational consumer's: a rise

attribute levels) and the denominator in Equation (4) does not feature the absolute value of attributes. The model can be extended in a straightforward way to negative attribute levels.

Δp in price disutility is offset by an increase $\delta \cdot \Delta p$ in quality utility, so that the marginal rate of substitution between a quality increase and a price reduction is δ. If instead price is salient, an increase Δp in price is offset by an increase $(1/\delta) \cdot \Delta p$ in quality, so that the salient thinker's marginal rate of substitution is $1/\delta$. As $\delta \to 1$, the salient thinker converges to the rational thinker. As $\delta \to 0$, the salient thinker considers only the most salient attribute and neglects all others. Normalization of the utility weights ensures that valuation of the good lies between q_k and $-p_k$. In Definition 2 utility is assumed to be distorted according to the salience *ranking* of quality and price. Section 2.2 discusses the role of this assumption.

To see how the model works, suppose that a consumer is evaluating two bottles of wine, a high end wine (q_h, p_h) and a low end wine (q_l, p_l), where qualities and prices are known and satisfy $q_h > q_l$ and $p_h > p_l$. The reference wine has quality $\overline{q} = (q_h + q_l)/2$ and price $\overline{p} = (p_h + p_l)/2$. According to Proposition 1, wine h's high quality is salient if and only if $q_h/p_h > \overline{q}/\overline{p}$, which can be written as:

$$\frac{q_h}{p_h} > \frac{\overline{q}}{\overline{p}} > \frac{q_l}{p_l}. \tag{6}$$

Thus, q_h is salient for wine h when the high end wine has a higher quality/price ratio than the low end wine. Proposition 1 similarly implies that, when condition (6) holds, the lower quality q_l is salient for wine l. In sum, when the quality price ratio is higher for the high quality wine, quality is salient for both wines. When the quality price ratio is higher for the lower quality, cheaper wine, price is salient for both wines.

Consider how salience affects choice. When prices are salient, namely when $q_h/p_h < q_l/p_l$, Expression (5) implies that the low end wine l is chosen over the high end wine h provided:

$$\delta \cdot (q_l - q_h) - (p_l - p_h) > 0, \tag{7}$$

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

Analogously, when $q_h/p_h > q_l/p_l$ quality is salient and Expression (5) implies that the low end wine l is chosen over the high end wine h provided:

$$(q_l - q_h) - \delta \cdot (p_l - p_h) > 0, \tag{8}$$

which is harder to meet than its rational counterpart, with $\delta = 1$. Intuitively, when quality is salient, the salient thinker overvalues both wines, but overvalues the high quality wine more because quality is the dimension along which the high end wine does better.

Salience tilts preferences toward the wine offering the highest quality/price ratio. This is a general property of our model. Suppose that the salient 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 gradient in quality and price, formally $q_k - q_{k'} = p_k - p_{k'}$ for all $k, k' = 1, ..., N$. Assume, without loss of generality, that quality and price increase in k (i.e. $q_1 < ... < q_N$ and $p_1 < ... < p_N$), and that the reference good is not an element of the choice set. In Appendix A we prove:

Proposition 2 *Along a rational linear indifference curve, the salient thinker chooses the good with the highest quality/price ratio. Specifically, he chooses good k^* where:*

$$k^* = \arg \max_{k=1,..,N} \frac{q_k}{p_k},$$

and $k^* = 1$ if $q_1/p_1 > 1$, while $k^* = N$ if $q_1/p_1 < 1$.

Along a rational indifference curve, the consumer selects the good with the highest ratio of quality to price.[7] This is the cheapest good ($k^* = 1$) when prices are low, and the most

[7]Consider adding the outside option $(0, 0)$ to the goods on the indifference curve. Because this addition does not affect the quality price ratio of the reference good, it does not change the salience ranking of goods that have above average quality and price. The outside option can only change the salience ranking of goods whose quality and price are such that their location relative to the reference good changes when $(0, 0)$ is introduced. These are typically "intermediate" goods, and they are generally valued less than either the highest quality good or the lowest quality good in the original choice set.

As a consequence, adding an outside option does not affect Proposition 2 as long as the highest quality good is preferred to $(0, 0)$ when quality is salient and the lowest quality good is preferred to $(0, 0)$ when price is salient (i.e. $q_N/p_N > \delta$, $q_1/p_1 > 1/\delta$).

More generally, when the choice context includes goods that do not lie on an indifference curve, salience tilts preferences towards options with *sufficiently high* quality/price ratio, particularly when these options expensive good ($k^* = N$) when prices are high. 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 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 getting a good deal. Instead, a good deal is attractive because it draws the consumer's attention to the dimension in which it does better than its competitor.

In Appendix B we show how the model works when goods are characterized by several quality dimensions. In this setting, diminishing sensitivity implies that improving one dimension of quality at the expense of another can increase the salience of the weaker dimension. As a consequence, the consumer tends to be attracted toward goods that have salient strengths and yet are balanced in their quality attributes. 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.

2.2 Discussion

The central idea of our model is that consumers focus on – and thus overweight – goods' attributes that stand out in the choice context. This generalises to riskless settings the logic of salience, initially developed for choice under risk. This approach is also consistent with recent results in neuroeconomics. Hare, Camerer, Rangel (2009) and Fehr and Rangel (2011) show that subjects evaluate goods by aggregating information about different attributes, with decision weights modulated by attention.[8]

Our model can be easily applied to standard economic problems by introducing salience-based valuation into a "rational" economic model. Doing so requires two key inputs describing the economic problem: i) the choice set, and ii) the attributes of each good that carry utility. We then add ingredients specific to the salience model, namely iii) the reference

are associated with high quality. In this case, it is possible that quality is salient and the high quality good is chosen for $q_N > p_N$.

[8]Exogenously varying the attention received by different attributes (e.g., by instructing subjects to attend to the "healthiness" of a snack) results in both higher brain activity associated with the attribute's decision value, and a higher likelihood that subjects choose the good superior along that attribute.

good, iv) a salience function, and v) a specification of salience weighting of attributes. We now discuss these ingredients in turn.

Rational models typically specify the product attributes in the utility function, as well as the choice set faced by the consumer. To apply the salience distortions of Definition 2 in our model, the utility function also needs to be separable. The most straightforward applications of our model feature only two attributes, quality and price, with utility linear in price.[9] When direct measurement of quality is unavailable, in our approach – as in rational models – a good's quality can be obtained as a latent variable that may depend on the measurable characteristics of the good. We restrict the admissible salience functions by assuming homogeneity of degree zero, which implies that salience is invariant to scalar transformations of attributes.[10] Under this assumption, we can elicit utility of quality directly within our model from willingness-to-pay experiments, see Section 5.

We assume that salience weighting is determined by salience ranking, with the magnitude of the salience distortion characterized by the parameter δ. Together with homogeneity of degree zero, this assumption makes our model significantly more tractable and applicable, since the effects of salience are characterized in terms of rankings of quality price ratios. Psychologically, rank based discounting captures the idea that valuation can be drastically affected by introducing small differences in an attribute such as price (Tverksy 1972, Kim, Novemsky and Dhar 2012). One feature of rank based discounting is that valuation can be non-monotonic, which may be undesirable in some applications. In Appendix C, we show that with a continuous salience weighting, non-monotonicities disappear under general conditions and all of our results qualitatively carry through.

Finally, an important step in applying our model is to appropriately specify the choice context, or the reference good, with respect to which salience is defined. In deterministic settings, the choice context can be assumed to coincide with the choice set, in line with our

[9] If, however, firms advertise or specialize along specific dimensions of quality or price, it may be useful to define salience directly on these dimensions, along the lines of the model of Appendix B. For instance, models of shrouded attributes (Gabaix and Laibson 2006) assume that consumers neglect certain price or quality components. In a multi attribute model, the logic of salience may help to explain which prices of quality components endogenously become shrouded (e.g. why consumers neglect them and why firms choose not to compete on them).

[10] This restriction is convenient because, while attributes are measured in price units (typically dollars), different units of price measurement differ by a scalar transformation (e.g. exchange rates or decimal units).

formal analysis so far. This is also the case in lab experiments where subjects are induced to think only about the choice set.[11]

In stochastic settings, a broader notion of context is needed. In those cases, an attribute's salience also depends on how much this particular realization differs from prior expectations. In Section 4 we extend the model to settings that depend on subjects' expectations (including Thaler's beer example) by assuming that the choice context also includes the agents' rational price expectations. This extension captures the fact that the choice situation brings to the consumer's mind "normal" prices, which then shape the consumer's reference price (Kahneman and Miller 1986). Sometimes, the prices rationally expected in a specific situation might be perceived as being far from normal, for instance the price of a tuna sandwich at an airport, or of a hotdog at a baseball game. However, the assumption of rational expectations is an intuitive and model-consistent way to strike a balance between psychological precision and testability of the model.

Because of the difficulties in measuring quality and context more generally, one might be concerned that our model can predict any behavior by a judicious construction of choice context. To allay these concerns, Section 5 identifies falsifiable predictions of our model that obtain even if the researcher has limited information about qualities or expectations.

We conclude this Section by discussing some related literature. Several models of consumer choice seek to rationalize context effects by incorporating loss aversion relative to a reference good (see Tversky and Kahneman 1991, Tversky and Simonson 1992 and Bodner and Prelec 1994). An implication of these models is a bias toward 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 and Shapiro data.[12]

[11]In this paper, we take the choice set as given, but evidence suggests that consumers typically consider only a subset of the options available in the market. The typical number of options in such consideration sets (or evoked sets) ranges from 2 to 5 goods, see Hauser and Wernerfelt (1990). This observation justifies our occasional focus on small choice sets. Endogenizing the consideration or evoked set is an important direction of future work, see Hauser and Wernerfelt (1990) and Eliaz and Spiegler (2010).

[12]In our model diminishing sensitivity implies a "loss aversion" type of effect: deviations occurring below the reference attribute level are more salient than those occurring above it. For attributes yielding positive utility, this is reminiscent of the idea that "losses loom larger than gains." The implications for valuation,

Gabaix (2012) develops a model of rational inattention, in which attention to different product attributes is efficiently allocated ex-ante, leading the consumer to neglect some of the attributes. In our model, consumer attention to different product attributes is drawn ex- post, depending on which attribute stands out.

Models of relative thinking assume 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, which is reminiscent of the diminishing sensitivity property of salience. Cunningham (2011) reproduces some related patterns of choice, such as the Savage car radio puzzle. This approach however does not account for patterns of choice in which ordering plays a role, such as the Hastings and Shapiro evidence on gasoline (section 4.1).

Koszegi and Szeidl (2013) build a model that centrally features the idea of ordering: their consumers are essentially salient thinkers who focus on and overweigh those attributes in which options differ the most in terms of utility. Koszegi and Szeidl use their model to shed light on biases in intertemporal choice. By neglecting diminishing sensitivity, the Koszegi and 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 high quality manufacturers to avoid shortcomings in any aspect of their merchandise.

By combining diminishing sensitivity with ordering within a choice context, our model generates the central prediction linking price sensitivity to context: consumers are relatively insensitive to price differences among goods of different qualities at expected high price levels, while they are price sensitive when faced with unexpected parallel price increases. This mechanism both provides a unified account of several well-known choice patterns and puzzles and generates new implications.

however, are very different from loss aversion.

3 Salience and Demand for Quality

We now examine the implications of our model for the reaction of consumers to two distinct manipulations of the choice set. We first explore diminishing sensitivity by considering uniform price shifts of all the goods in the choice set. We then explore ordering by considering changes in the reference good due to the addition of an irrelevant alternative.

3.1 Price Differences across Contexts and Diminishing Sensitivity

The wine example from the Introduction suggests that a consumer's price sensitivity depends on the price level, namely that the consumer is more price sensitive when choosing among cheaper goods. This idea is a direct implication of Proposition 2. In particular, suppose that the choice set $\mathbf{C}$ consists of N goods lying along a rational indifference curve. A uniform price shift, $p_k \to p_k + \Delta p$ for $k = 1, \ldots, N$, keeps price gradients $p_k - p'_k$ unchanged and does not affect the preferences of a rational consumer with linear utility (1). However, this may not be the case for the salient thinker. In fact, Proposition 2 has stark implications for the comparative statics:

Corollary 1 *Suppose that the consumer initially chooses the cheapest good in* $\mathbf{C}$*. Then, his demand for quality (weakly) increases in the price shift* Δp*. Specifically, there is a threshold* $\Delta p^* > 0$ *such that the salient thinker continues to choose to lowest quality good for* $\Delta p < \Delta p^*$*, but switches to the highest quality good for* $\Delta p > \Delta p^*$*.*

In the wine example from the Introduction, the consumer chooses between two wines of qualities $q_h = 30$ and $q_l = 20$. At the store, prices are $p_h = \$20$ and $p_l = \$10$. At the restaurant, prices are uniformly increased by $\Delta p = \$40$. Though the quality gradient $q_h - q_l$ and the price gradient $p_h - p_l$ are the same in the two situations, the consumer chooses the cheap wine at the store and the expensive wine at the restaurant. Due to diminishing sensitivity of the salience function, the \$10 price difference among the wines is more noticeable to the consumer at the low price level of the store than at the high price level of the restaurant. To see how this example illustrates the logic of Corollary 1, note that at the store, the cheaper wine has a higher quality-price ratio ($20/10 > 30/20$), while at the restaurant the ranking is reversed ($20/50 < 30/60$). Price is salient at the store; quality is salient at the restaurant.

This mechanism induces consumers to display higher price sensitivity for choice among cheaper goods.[13] As we show in Section 4, in non-deterministic settings the strength of this effect depends on price expectations.

Diminishing sensitivity can lead to outright preference reversals such as Savage's (1954) car radio problem. 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, separately from the car purchase. In fact, diminishing sensitivity implies that the salience of the price p_r of the radio is higher when evaluated in isolation, $\sigma(p_r, p_r/2)$, than against the backdrop of the much higher price p of the car, $\sigma(p + p_r, p + p_r/2)$. Similarly, subjects are willing to travel 10 minutes to save \$5 on a \$15 calculator, but not on a \$125 jacket (Kahneman and Tversky 1984). Diminishing sensitivity implies the discount is more likely to be salient if applied to the calculator than to the jacket, $\sigma(10, 12.5) > \sigma(120, 122.5)$. Intuitively, in the Savage example the cost of the add-on is less salient when it is "hidden" behind the high price of the core good, while in the jacket-calculator problem the given dollar discount is more salient when it reduces the price of the cheaper good. We thus formalize the intuitive argument based on Weber's law offered in Thaler (1980).

3.2 Decoy and Compromise Effects

A well documented anomaly in both marketing and psychology is the so called decoy effect (Huber, Payne and Puto 1983, Tversky and Simonson 1993): adding to a pairwise choice an option dominated by one of the goods boosts the demand for the dominating good. Another well known anomaly is the compromise effect (Simonson 1989): adding an extreme option

[13]This mechanism differs from models based on loss aversion. In Bodner and Prelec's (1994) model, consumers evaluate each good's gains and losses relative to the same reference good, namely the "centroid" (or average) good in the choice set. As prices increase uniformly, the gains/losses relative to the reference price stay constant, leaving choice unchanged. In our model, in contrast, as prices increase a given price difference becomes less salient. This is because salience is evaluated relative to not experiencing an attribute and not with respect to experiencing its reference level. Price levels can also affect the rational consumer's choice through income effects, but in the opposite direction of our prediction: under concave utility, consumers are more price sensitive at higher price levels.

to a pairwise choice induces subjects to change their preferences toward the middle of the road, or compromise, option. Assuming, along the lines of Section 2, that the consumer is perfectly informed about the available goods, both anomalies constitute violations of the axiom of independence of irrelevant alternatives (IIA).[14]

Our model can provide an intuitive account for these phenomena as a consequence of the impact of the added option on salience. We describe conditions under which such preference reversals arise, providing a novel and testable prediction of our model. In Section 6, we use the logic of the decoy effect to model misleading sales from the perspective of consumer psychology.

To see how decoy effects arise in our model, consider again the wine example from the Introduction, with a variation in which a third, more expensive and high quality wine d is added to the wine selection at the store

$$\begin{aligned} \mathbf{C}^{store} &= \{(30,\$20),(20,\$10)\} \\ \mathbf{C}^{store(d)} &= \{(30,\$30),(30,\$20),(20,\$10)\} \end{aligned} \tag{9}$$

Wine d, with quality 30 and price \$30, is dominated by wine h, and yields lower utility than the original options, namely $u(30,-30) = 0 < u(30,-20) = u(20,-10) = 10$. A rational decision maker is indifferent between h and l but prefers both of them to d. The inclusion of d in the choice set does not affect his choice.

Consider the choice of a salient thinker. As shown in Section 3.1, in the context $\mathbf{C}^{store}$ the salient thinker picks the low end wine l because it has the highest quality/price ratio, so prices are salient and this disproportionately hurts the expensive wine h. What happens when 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 fact, the quality/price ratio of d, 30/30, is lower than the quality/price ratio of the high end wine h, 30/20. This is an important change: now, by comparison with d, the high end wine h seems a better deal than in the original choice set!

To see the implications for choice, note that in the set $\mathbf{C}^{store(d)}$, the reference wine is

[14] Wernerfelt (1995) and Kamenica (2008) explain the decoy effects by suggesting that decoys indirectly provide consumers with information about the quality of the products.

(26.7, \$20). The high end wine h delivers above reference quality $30 > 26.7$ at the reference price \$20. As a consequence, the quality of h becomes salient. It is easy to check that the low end wine l remains price salient. Under this new salience configuration, the salient thinker prefers h to l. The model therefore yields a decoy effect: in pairwise choice the salient thinker prefers l to h but he switches to h when an expensive inferior good d is added, thus violating IIA.[15] Intuitively, when the bad deal d is added, h becomes a good deal as its quality becomes salient, and its high price becomes less disturbing.

This logic accounts for the decoy effect but also provides new predictions about its manifestations. First, it implies that this violation of IIA does not rely on introducing a decoy d that is dominated by the originally neglected option h, but rather is driven by shifts in the perception of that option's salient attributes. In particular, a decoy is a good with a significantly lower quality to price ratio than existing goods. Second, our model predicts that decoy effects are asymmetric: they work for high quality goods, but they are much less effective in boosting demand for low quality goods.

To see both effects, consider two goods (q_l, p_l) and (q_h, p_h), and denote by $\Delta u = [q_h - q_l] - [p_h - p_l]$ the rational utility difference between them. For simplicity, we focus on cases in which h is chosen over l if and only if its quality is salient, namely:

$$-(1-\delta)(p_h - p_l) \leq \Delta u \leq (1-\delta)(q_h - q_l), \tag{10}$$

We also assume that decoy options d are such that the high quality good h has above average quality and price even after the decoy is introduced, namely $(q_d + q_h + q_l)/3 \leq q_h$ and $(p_d + p_h + p_l)/3 \leq p_h$.

When Equation (10) holds, we have:

Proposition 3

i) If $\frac{q_l}{p_l} > \frac{q_h}{p_h}$*, so that price is salient and* l *is chosen from* $\{l, h\}$*, then for any* 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* h *is quality salient in* $\{l, h, d\}$*. Moreover, there exist options* d *satisfying the previous condition and* $q_d > q_h, p_d > p_h$ *such that* h *is chosen from* $\{l, h, d\}$*.*

ii) If $\frac{q_l}{p_l} < \frac{q_h}{p_h}$*, so quality is salient and* h *is chosen from* $\{l, h\}$*, then there exist no decoy*

[15] As d lies on a lower indifference curve, and h is quality salient, d is never chosen.

options d such that $\frac{q_d}{p_d} \leq \frac{q_l}{p_l}$ and h is price salient in $\{l, h, d\}$. In particular, for no d satisfying $\frac{q_d}{p_d} \leq \frac{q_l}{p_l}$ is l chosen from $\{l, h, d\}$.

Result *i*) says that the decoy must be a bad deal. When the quality price ratio q_d/p_d of the decoy is low, it lowers the reference quality-price ratio to the point that $q_h/p_h > \overline{q}/\overline{p}$. As a consequence, the quality of h becomes salient, so that h is now overvalued and is chosen from $\mathbf{C}^{store(d)}$, provided the decoy is not so good that the consumer prefers it to h. Critically, the decoy need not be dominated by h: preference reversals can also occur when $q_d > q_h$ and $p_d > p_h$. In this case, in the augmented choice set, h provides intermediate levels of quality and price, but offers a good quality to price ratio when compared to the decoy. This creates a compromise effect in our model, with the same logic as the decoy effect.

Result *ii*) says that the decoy effect is asymmetric, in the sense that it does not reverse an initial preference for high quality goods. When quality is salient in pairwise choice (namely $q_h/p_h > q_l/p_l$), adding a decoy to the lower quality good l may cause its low price to become salient. However, since the decoy reduces the quality-price ratio of the reference good, it cannot at the same time make the high price of h salient. Since h remains quality salient, it is still chosen in the enlarged choice set. There are instances, not contemplated in Proposition 3, in which a decoy may increase the relative valuation of a lower quality good.[16] 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 decoys than their low quality, low price competitors.

The asymmetry of decoy effects is consistent with Heath and Chatterjee's (1995) survey of experimental and field results on decoys. In agreement with our predictions, the authors document a robust asymmetry in the workings of the decoy effect: adding appropriate decoys typically boosts experimental subjects' demand for high quality goods, at the expense of demand for low quality goods. In contrast, adding decoys for low quality goods does not boost the demand for the latter. In this respect, our model differs substantially from formalizations of context dependence based on loss aversion (Tversky and Simonson 1993, Bodner and Prelec 1994), where consumers minimize losses across all attributes and mechanically prefer middle-of-the-road options, so that asymmetries do not arise.

[16] These include decoys with extremely high quality to price ratios, but very low levels of quality.

4 Salience and Expectations

So far we took a narrow view of the choice context by identifying it with the choice set $\mathbf{C}$. A long tradition in psychology, however, stresses that the choice context is not limited to the choice set, but includes also the alternatives that the decision maker expects to find in the current choice setting (Kahneman and Miller 1986). As we stressed previously, this aspect seems relevant to understanding several phenomena. For example, in the Thaler beer example, framing subjects with a specific context (resort or store) makes them think about the price they could expect in that context. In the Hastings and Shapiro example, the consumer approaches the gas station having in mind a price expectation for gas. The logic of our model implies that, once evoked, expected prices shape choice by affecting the salience ranking of different product attributes.

To see how these effects may work, we incorporate expectations in our definition of context. We do so in a straightforward way: we assume that the choice context consists of the goods in the choice set together with those same goods at their *rationally* expected prices.

Definition 3 *The choice context is the set* $\mathbf{C}_{cont} = \mathbf{C} \cup \mathbf{C}_e$, *where* $\mathbf{C}$ *is the externally given choice set while* $\mathbf{C}_e = \{(q_k^e, p_k^e)\}_{k=1,\ldots,N}$ *is the set of goods the consumer expects to find in the choice setting. We assume that:*
i) For each $(q_k, p_k) \in \mathbf{C}$, *there is a* (q_k^e, p_k^e) *whose expected quality satisfies* $q_k^e = q_k$, *and whose expected price* p_k^e *is the rational expectation of* p_k, *namely* $p_k^e \equiv \mathbb{E}[p_k]$.
ii) The choice context is summarized by a reference good $(\overline{q}, -\overline{p})$, *where the reference (or normal) levels of quality and price are their average values in* $\mathbf{C}_{cont}$, *namely* $\overline{q} = \frac{1}{N}\sum_k q_k$, *and* $\overline{p} = \frac{1}{2N}\sum_k (p_k + p_k^e)$.

This definition captures the idea that the choice situation brings to the consumer's mind normal prices, which then shape the consumer's reference price. Assuming rational expectations is an intuitive and model-consistent way to capture normal prices, although not universally applicable (see Section 2.2). The model can be extended to allow for differences between expected and actual qualities, but here we restrict our analysis to price surprises.

Definition 3 has two important implications. First, when actual prices coincide with price expectations (i.e., when $p_k = \mathbb{E}[p_k]$ for all k) the reference good in the choice context coindicdes with the average good in the choice set. As a result, the consumer behaves as if the choice context and the choice set coincide, $\mathbf{C}_{cont} = \mathbf{C}$. In this sense, our previous analysis is directly applicable to deterministic settings or to lab experiments where, unless explicitly primed, consumers have no experience to base their forecasts on.

Second, when expected and actual prices differ, the model uncovers an important new effect: the salience of price is determined not only by comparing prices across goods, but also by comparing actual prices with expected prices. Even if the price of a good is similar to that of other goods, it may be salient if it is unexpectedly high or low. The salient thinker's attention is drawn not just to differences between available options, but also to the surprising features of the environment.

Before studying the implications of this assumption, it is useful to discuss how this model could be applied. Our use of rational expectations is reminiscent of the Koszegi-Rabin (2006) model, in which reference points coincide with expectations, but has two important differences. First, in our model expectations are fully exogenous: they are simply another input into the consumer's decision making and do not depend on his planned (equilibrium) choices. This simplifies the model and facilitates its application because it only requires the modeller to know the exogenous empirical distribution of prices. Second, our mechanism relies on salience relative to the reference good, and not on loss aversion. These two forces may act in the same direction – when a good's disadvantage relative to the reference is salient – but often act in directly opposite directions, when a good's advantage is salient.

4.1 Price Surprises: the Role of Ordering

The ordering property of salience has key implications for the effect of price surprises. Consider a consumer planning a meal at his favorite restaurant. While opening the wine list, the consumer forms rational expectations of the prices he may find at that restaurant. Suppose that these expectations are given by $p_h^e \equiv \mathbb{E}[p_h | rest] = \$60$ and $p_l^e \equiv \mathbb{E}[p_l | rest] = \$50$. Suppose however that actual prices turn out to be higher than expected, and equal to $p_h^{rest} = \$(60+s)$ and $p_l^{rest} = \$(50+s)$, where $s > 0$ is the price surprise. The consumer's choice context is then given by:

$$\mathbf{C}^{rest} = \begin{cases} h^{rest} = & (30, \$(60+s)) \\ l^{rest} = & (20, \$(50+s)) \\ h^{e,rest} = & (30, \$60) \\ l^{e,rest} = & (20, \$50) \end{cases} . \tag{11}$$

In this case, salience is not just affected by the prices in the wine list, but also by the discrepancy of the actual wine prices from the expected prices. Formally, the reference wine is now described by:

$$(\overline{q}, \overline{p}) = \left(25, \$\left(55 + \frac{s}{2}\right)\right).$$

The high end wine h^{rest} still yields above average quality, but it may now feature a lower than average quality/price ratio. This is indeed the case provided:

$$\frac{30}{60+s} < \frac{25}{55+\frac{s}{2}} \Leftrightarrow s > 15. \tag{12}$$

If the price surprise is sufficiently large, the high end wine becomes price salient. This greatly reduces the value of h^{rest} as perceived by the salient thinker. This price surprise might also render the low end wine price salient. However, by drawing the consumer's attention to prices, this price surprise reduces the relative valuation of the high end wine, inducing the consumer to choose the low end wine, regardless of its salience ranking. When the consumer finds wines at the restaurant to be unexpectedly pricey, he switches to lower quality wine, excessively reducing his demand for quality relative to the rational case.

Compare this result with the case where restaurant prices are high but fully expected, namely $s = 0$. In that case, the consumer is adapted to higher wine prices and, by diminishing sensitivity, he focuses on quality. When instead prices at the restaurant are surprisingly high, they become very salient, and disproportionately draw the consumer's attention. Because the consumer views the high quality wine as having a low quality to price ratio (relative to normal prices), he choses the low quality wine.

This logic also explains what happens when restaurant prices are surprisingly low, namely $s < 0$. In this case, Equation (12) is harder to satisfy: relative to historical prices, and thus to the reference good, the high end wine now has a high quality to price ratio. In other words, unexpectedly low prices tend to draw the consumer's attention to quality, reducing his price sensitivity. As a consequence, the high quality wine is chosen over its low quality competitor. Interestingly, this effect only holds for moderate price drops. If the price drop is sufficiently drastic, the consumer's attention is necessarily drawn to prices, which again favors the low quality wine. Our model thus exhibits an asymmetry whereby unexpected price hikes always induce consumers to substitute toward cheaper goods while unexpected price declines tend, but do not always, induce consumers to substitute toward more expensive goods.

Before exploring these effects in more detail, we note that this intuition helps account for the evidence in Hastings and Shapiro (2013). They show that consumers react to parallel increases in gas prices by switching to cheaper (and lower quality) gasoline. One explanation for this behavior is mental accounting (Thaler 1999): when purchasing gas, the consumer thinks about the "gas consumption" account, to which he allocates a fixed monetary budget. The budget is targeted to past prices. As gas prices increase, if the budget is violated, the consumer (who mostly cares about the quantity of gas) substitutes expensive, high grade gas with cheaper, lower grade gas.

In our model, expectations matter not through a fixed monetary budget for gas consumption as in mental accounting, but by shaping the relative salience of gas attributes. To show this, we briefly work through this example following the steps described in Section 2.2: the choice set consists of high octane gas h and low octane gas l, which differ in their quality and price. Consumer utility is separable. To describe the choice context, we assume, as do Hastings and Shapiro (2013), that gas prices follow a random walk. The rationally expected gas price then coincides with the previous price, namely $\mathbb{E}\left[p_{i,t}\left|p_{i,t-1}\right.\right] = p_{i,t-1}$ for both types of gas $i = h, l$. The consumer approaches the gas station with this expectation for price. When the consumer observes surprisingly high prices $p_{i,t} > p_{i,t-1}$ for all gas grades i, the price of the high quality grade – which is above the average price in the choice set to begin with – becomes even more salient. Thus, because the salience of quality is constant, a price increase makes the consumer relatively more sensitive to price differences, and more likely to switch to lower octane, cheaper gas.[17] Suppose instead that the consumer observes

[17] It is useful to compare this result to the Kozsegi and Rabin's (2006) model with loss aversion relative to

surprisingly low prices for all gas grades. Then, as long as the price drop is not too steep, the price $p_{h,t}$ of the high quality grade is closer to the reference price (even if it falls below $p_{l,t-1}$), and therefore is less likely to be salient. In this case, a price drop makes the consumer less sensitive to price differences, and more likely to switch to higher octane, more expensive gas.

In Appendix D, we show how the broad patterns in the demand for gasoline documented by Hastings and Shapiro emerge from our model, namely that the share of regular gasoline rises with gas prices, as consumers' price sensitivity is asymmetrically boosted by price hikes. In their paper, Hastings and Shapiro calibrate a salience model and evaluate its predictions. An important aspect of the calibration is the measurement of quality, which shapes the salience of quality to the consumer, affecting his reaction to given price changes.[18] This particular setting suggests two measures of quality, one numerical (87 versus 91 octane rating) and one descriptive (regular versus premium gasoline). The descriptive measure likely conveys a more salient quality difference than a 4% difference in octane rating. Lacking guidance on which measure to use, in many applications it may be useful to treat quality as a latent variable (even when certain objective proxies for quality are available) to be estimated with the remaining parameters of the model.

4.2 Price Shifts: Diminishing Sensitivity vs. Ordering

Our model generates two seemingly contradictory findings: lower price sensitivity at higher price levels, but higher price sensitivity after unexpected price hikes. We now provide a general characterization of the effects of price changes in our model and show that these results are due to the tension between the basic forces of ordering and diminishing sensitivity.

To do so, take the choice context $\mathbf{C}_{cont}$ as given and consider the effect on salience of a

expectations. Whether that model can rationalize the Hastings and Shapiro data depends on how one models the "losses" account. If the consumer experiences strong losses from not buying high grade gas as originally planned, he will be reluctant to cut high grade consumption. Because the price increase causes the consumer to lose money anyway, he strongly wants to avoid getting worse quality gas than planned. This mechanism dampens the consumer's reaction to price hikes, in the spirit of Koszegi and Rabin's approach to sales. If instead the consumer does not experience strong losses from not buying high grade gas as originally planned, he may try to avoid losses on other goods by switching his gas consumption to regular (lower quality) gas.

[18] This aspect matters for the behavior predicted by the model. As we show in Appendix D, the larger the utility difference among gas qualities, the more likely it is that a parallel price decline increases the relative salience of quality, inducing the consumer to choose the more expensive gas.

marginal price increase in a proper subset of $\mathbf{C}_{cont}$. The choice context includes both available goods and the same goods at expected prices. Partition $\mathbf{C}_{cont}$ into two subsets $\mathbf{C}_{\mathbf{F}}$ and $\mathbf{C}_{\mathbf{C}}$, such that $\mathbf{C}_{\mathbf{F}}$ is the set of goods for which price is held fixed, while $\mathbf{C}_{\mathbf{C}}$ is the set of goods for which we consider a marginal increase in price. Depending on the "experiment", $\mathbf{C}_{\mathbf{C}}$ can include actual prices, or expected prices, or both. The following cases are of particular interest:

a) Expected price hikes, as in the store vs. restaurant example. In this case, both actual and expected prices change, so that $\mathbf{C}_{\mathbf{C}} = \mathbf{C}$ and $\mathbf{C}_{\mathbf{F}} = \emptyset$.

b) Unexpected price hikes, as in the overpriced restaurant example and the Hastings Shapiro gas example. Here only actual prices change while expected prices stay constant in the choice context, so that $\mathbf{C}_{\mathbf{C}}= \mathbf{C}_{choice}$ and $\mathbf{C}_{\mathbf{F}} = \mathbf{C}_{e}$.

c) Relative price changes. Only a subset of actual (and perhaps expected) prices changes. Here $\mathbf{C}_{\mathbf{C}}$ is the subset of actual (and expected) goods whose relative price changes, while $\mathbf{C}_{\mathbf{F}}$ contains all remaining (actual and expected) goods.

Denote by η the fraction of goods in the choice context that belong to $\mathbf{C}_{\mathbf{C}}$. Denote by $\overline{p}$ the reference price in $\mathbf{C}$, and by $\overline{p}_{\mathbf{X}}$ the average price in subset $\mathbf{C}_{\mathbf{X}}$, $\mathbf{X} = \mathbf{F}, \mathbf{C}$ (formally, all these prices are computed prior to the price increase in $\mathbf{C}_{\mathbf{C}}$). In example a), $\eta = 1$ and $\overline{p}_{\mathbf{c}} = \overline{p}$. In example b), $\eta = 1/2$, and $\overline{p}_{\mathbf{F}} = \overline{p}_{\mathbf{C}} = \overline{p}$. In example c), average prices depend on the specific goods considered and η may be small, close to 0. We then show:

Proposition 4 *If $\overline{p}_{\mathbf{C}} \geq \overline{p}$, then a uniform marginal increase in the prices of goods in $\mathbf{C}_{\mathbf{C}}$ boosts the salience of price for the most expensive good in $\mathbf{C}_{\mathbf{C}}$ only if:*

$$\frac{p_{\mathbf{C}}^{\max} - \overline{p}_{\mathbf{C}}}{\overline{p}_{\mathbf{F}}} < \frac{1-\eta}{\eta}, \tag{13}$$

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

Take a category of goods that are originally at least as expensive as average (i.e. $\overline{p}_{\mathbf{C}} \geq \overline{p}$). A uniform increase in the prices in this category boosts the salience of price for its most expensive members – and thus increases the consumer's price sensitivity for these goods – provided the category affected by the price hike is sufficiently small, namely η is small.

The size η of the category affected by the price hike is important precisely because it modulates the strength of ordering versus diminishing sensitivity. When the category is large, the price hike induces a commensurate increase in the reference price $\overline{p}$, so the price differential between category prices and the reference price does not grow significantly. As a consequence, diminishing sensitivity dominates, Equation (13) does not hold, and a price increase reduces the salience of price. This is what happens in the case of fully expected price hikes in setting a) above: actual and expected prices uniformly increase ($\eta = 1$) and price salience falls for all goods, including the very expensive ones, reducing the consumer's price sensitivity for all goods.

When instead the category affected by the price hikes is small (in the limit, when $\eta \approx 0$), the reference price $\overline{p}$ is only modestly affected by the price hike, so that the price differential between category prices and $\overline{p}$ grows disproportionately. In this case, ordering dominates, Equation (13) holds, and a price increase boosts the salience of price. The surprising price hike in setting b) falls into this case: only actual prices increase while expected prices stay constant (so $\eta = 1/2$). Because the prices of the available goods are very high relative to the reference price, the most expensive goods become price salient, increasing the consumer's price sensitivity to these goods.

For a given category size, ordering is likely to dominate when the proportional increase in the price $p_{\mathbf{C}}^{\max}$ of the most expensive category item is larger than the proportional increase in the average price $\overline{p}$ in the entire choice context. As shown by Equation (13), this occurs when the price range $p_{\mathbf{C}}^{\max} - \overline{p}_{\mathbf{C}}$ in the category is sufficiently small.

As illustrated in setting c) above, this phenomenon describes consumers' reaction not only to surprises in overall price levels, but also to relative price changes. Imagine for instance a consumer choosing among different qualities of Bordeaux wines. Equation (13) says that as the price of Bordeaux wines uniformly increases, the consumer is more likely to substitute towards cheaper Bordeaux (or potentially to leave the category altogether) if Bordeaux wines are on average expensive and display relatively low price dispersion.

Our model thus yields testable predictions on whether price hikes boost or dampen the demand for quality, depending on the magnitude of the price hike and on the market structure (measured by η and price dispersion). Proposition 4 provides the central comparative statics on how price sensitivity depends on context: consumers are relatively insensitive to price differences among goods at expected high prices, while they are very price sensitive when faced with unexpected price increases.

4.3 Manipulation of Expectations

The Willingness To Pay (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 (q_0, p_0), where typically $q_0 = p_0 = 0$. In standard theory, knowledge of q and of q_0, p_0 are sufficient to determine WTP for q (assuming quasi-linear utility, as we do here).

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. 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).

Our model indicates that WTP may vary across contexts because the contexts differ in the associated price expectations. As the experimenter mentions the nearby location, he prompts the decision maker to form an expectation for the price of beer, which is included in the choice context. When thinking of the high expected price at the resort, the salient thinker is willing to pay a high price for the beer and still perceive quality as salient. When thinking of the low expected price at the store, however, the salient thinker is not willing to pay a high price for the beer, because a high price would be salient at the store. Through its effect on the reference good and salience, the expected price acts as an anchor for the consumer.

Formally, suppose the consumer must state his WTP for quality q while expecting a good (q, p^e_γ), where p^e_γ is the expected price at which quality q is sold in context γ. In the

Thaler experiment, $\gamma = resort, store$. The consumer evaluates the good (q,p) at a price p, so his choice context is $\mathbf{C} \equiv \{(0,0),(q,p^e_\gamma),(q,p)\}$, where $(0,0)$ is the outside option of not consuming q. We define the consumer's willingness to pay for q in the context γ as:

$$\begin{aligned} &\text{WTP}(q|(q,p^e_\gamma)) = \sup p \\ &\text{s.t.} \quad u^S((q,p)|\mathbf{C}) \geq u^S((0,0)|\mathbf{C}). \end{aligned} \tag{14}$$

WTP is still defined as the maximum price p that the consumer is willing to pay for q as opposed to getting the outside option $(0,0)$, but the superscript S indicates that the consumer's preferences are distorted by salience. Crucially, different values of the good's price p can alter the salience of its attributes, changing the consumer's valuation. Thus, maximization in (14) tends to select a price p such that good (q,p)'s quality is salient.

In the choice context $\mathbf{C}$, the reference good has quality $\overline{q} = q \cdot \frac{2}{3}$ and price $\overline{p} = \frac{p+p^e_\gamma}{3}$. We can then show:

Proposition 5 *The consumer's WTP for q depends on the expected price p^e_γ as follows:*

$$WTP(q|\mathbf{C}) = \begin{cases} \delta q & if \quad p^e_\gamma \leq \delta q \\ p^e_\gamma & if \quad \delta q < p^e_\gamma \leq \frac{1}{\delta} \cdot q \\ q/\delta & if \quad \frac{1}{\delta} \cdot q < p^e_\gamma \leq \frac{7}{2\delta} \cdot q \\ \delta q & if \quad p^e_\gamma > \frac{7}{2\delta} \cdot q \end{cases} \tag{15}$$

As $\delta \to 1$, the willingness to pay tends to q and becomes independent of context p^e_γ.

The price expectation p^e_γ only affects WTP if the consumer is a salient thinker, i.e. if $\delta < 1$. Importantly, for $p^e_\gamma \leq \frac{7}{2\delta}q$ the consumer's WTP weakly increases in the expected price p^e_γ. In contexts where quality is more expensive, the consumer is willing to pay a higher price p and still view quality as salient.[19] Through salience, a higher price p^e_γ acts like an anchor, increasing WTP. This mechanism can explain Thaler's beer experiment because the expected price in the resort is larger than that in the store (i.e., $p^e_{resort} > p^e_{store}$).

[19]Put differently, as p^e_γ increases the consumer perceives (q,p) as a good deal even at higher prices p.

Interestingly, Proposition 5 suggests that when the reference price is extremely high, this effect vanishes. In this respect our model yields a different prediction than the one based on transaction utility. If the expected price is too high, formally if $p^e_\gamma >> q/\delta$, price becomes salient and the consumer's WTP drops. Intuitively, if the consumer expects beer at resorts to be outrageously expensive, then he focuses on the high prices at resort, which causes his valuation for beer to drop in this context. As a result, the consumer refuses to buy a beer even if the price is close to his true valuation. In fact, this causes the consumer's WTP to drop below his true valuation. The WTP in (14) is graphically represented in Figure 1.

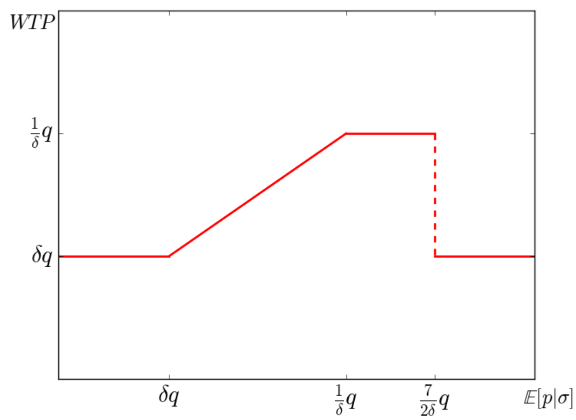

Figure 1: Willingness to Pay for q as a function of reference price $\mathbb{E}[p|\gamma]$.

5 Falsifiable Predictions of the Model

In the previous sections we characterized the predictions of the model that hold when the choice context $\mathbf{C} \cup \mathbf{C}_e$ is known. In realistic settings, though, only some features of the choice context (e.g., the prices of goods in the choice set) are readily available to an observer. Other inputs to the model such as the precise quality levels or the expected prices are difficult to observe. This can create the impression that our model may explain any possible behavior by a judicious choice of these unobservable variables.

To allay this concern, we highlight some falsifiable predictions of our model when the researcher does not have direct knowledge of quality levels and/or expected prices. We study predictions that hold under any salience function satisfying A.0, and under rank based weighting. We continue to assume that the utility function is linear.[20] We consider two scenarios:

1) The choice context coincides with the choice set. Quality rankings are observed, but quality levels are not.

2) The choice context includes expected prices. Quality rankings are observed, but neither quality levels nor expected prices are observed.

Scenario 1 holds in lab experiments, in which subjects do not have strong pre-existing expectations about prices (unless these are induced by the experimenter), but it also holds in the field when consumers are familiar with the available goods and prices can (roughly) be assumed to be stable. One interesting property of this scenario is that even if the quality q of a good is not directly observable, it can be recovered by the researcher by running a pricing experiment eliciting a subject's willingness to pay. In such an experiment, the choice set is given by $\mathbf{C} \equiv \{(q,p),(0,0)\}$. Because the choice set coincides with the choice context, the reference good is $(\overline{q},\overline{p}) \equiv (q/2, p/2)$. Under a homogeneous of degree zero salience function, we then have that:

$$\sigma(q, q/2) = \sigma(p, p/2) = \sigma(1, 1/2).$$

That is, quality and price are equally salient for any willingness to pay p (more generally, for any asking price). In this pricing experiment valuation is non distorted, so that, for any salience weighting (rank based or continuous), WTP$= q$. Without price expectations, the willingness to pay for a given good is equal to its quality level, just as in the rational model.

The ability to recover quality through WTP is important, because it can in principle allow researchers to directly test the predictions of the model, such as those highlighted in Proposition 4, in deterministic or lab settings where price expectations do not play a role. But even if WTP cannot be readily elicited through price experiments (e.g. in certain field applications), information on quality ordering still yields falsifiable predictions based on price changes (which are observable) in situations where the choice context coincides with

[20]Though in this section goods are characterized by their quality and price, the results carry through if quality is specified differently, e.g. as several separate quality dimensions. This is because the results in this section relate only to price experiments (i.e. consumer reactions to price shifts).

the choice set.

In particular, suppose that we are in scenario 1) above, and the researcher only observes prices and the ranking of qualities. The simplest setting consists of a binary choice between a high quality and a low quality good. In this case, one can directly test the prediction of Corollary 1, whereby the demand for the high quality good should (weakly) rise after an expected uniform price increase. In fact, this prediction depends only on knowledge of: i) prices, and ii) the quality ranking of the two goods.

With more than 2 goods the analysis is more complicated, but our model yields testable predictions in this case as well. Suppose that the choice set has $N > 2$ goods, and that the researcher observes quality rankings $q_1 < q_2 < ... < q_N$, and prices $p_1 < p_2 < ... < p_N$. Because we are in Scenario 1, the researcher observes the reference price $\overline{p} = (1/N) \sum_k p_k$, and can manipulate all prices that determine $\overline{p}$. We can then prove:

Prediction 1 Let $p_j > \overline{p} > p_i$ and suppose that a consumer chooses good i at prices $\{p_k\}_{k=1,...,N}$ and good $j > i$ when all prices experience a given uniform increase. Then, if starting from the price level $\overline{p}$ the prices of all goods $k \neq i, j$ experience a uniform increase so that the new average price is equal to $p_j > \overline{p}$, then good j is preferred to good i, so the latter cannot be chosen.

This result highlights the mechanics underlying violations of the independence of irrelevant alternatives in our model.[21] It exploits the fact that the consumers' price sensitivity towards the high quality good j is reduced under two distinct patterns of price increases. Consider the assumption of Prediction 1: if a uniform price hike reduces the consumer's price sensitivity, and induces him to switch to higher quality $q_j > q_i$, then it must be that the consumer choses q_j as its price becomes less salient. In this case, Prediction 1 identifies a way to induce a violation of IIA: increase the prices of all other goods so that the reference price rises to p_j. This renders good j quality salient so that, by assumption, j is preferred to i. A change in the prices of the "irrelevant" goods $k \neq i, j$ reverses preferences among i

[21] To avoid a situation in which price manipulations influence the formation of price expectation in the lab, this prediction is best evaluated in a cross-subjects experiment. The prediction can also be tested in a within-subject paradigm, but only when it is possible to ensure that subjects fully expect the implemented price shifts (as in the store vs. restaurant example).

and j, in contrast to the rational model.[22]

Both properties of salience play a role here: the assumption of Prediction 1 uses diminishing sensitivity to infer that the higher quality good j is chosen over good i if its quality is salient; the prediction uses ordering to render good j's quality salient by changing only the prices of irrelevant goods. A similar logic can be used to obtain falsifiable predictions when the consumer has different expectations about prices but the researcher knows these expectations or has a good proxy for them, as when expectations are rational. Also in this case, violations of IIA can be engineered by comparing the effect of uniform price shifts that render quality more salient (which now may be price cuts) to the effect of asymmetric changes in the prices of irrelevant goods, or to the addition of irrelevant alternatives as in the decoy effect.

We now turn to case 2), in which price expectations are present but the researcher cannot observe them. In this case, the location of the reference price is not known, which makes it difficult to control the salience ranking of specific goods. This makes testing our model more difficult. However, Proposition 5 provides a clue on how the demand for quality q depends on its price p, given an arbitrary expected price p^e: the demand for quality is inverse-U-shaped in the sense that a good's valuation is maximized when current prices are close to expected prices (and quality is salient), and is lower otherwise (when price is salient).[23] In the simple case of pairwise choice, this behavior leads to the testable Prediction 2.

Prediction 2 Suppose the choice set consists of two goods ($N = 2$), with $q_1 > q_2$ and $p_1 > p_2$, and that the expected prices are fixed but not observed. Under uniform price shifts, $p_k \to p_k + \Delta p$ for $k = 1, 2$ and $\Delta p > 0$, there can be at most two changes in

[22] In Prediction 1, the asymmetric price changes in the second part ensure that good j is preferred to good i but not necessarily that good j itself is chosen. Two mechanisms are involved: first, in the second step of the Lemma other high quality goods k' may become quality salient and chosen over good j, even if j is quality salient itself. Second, with discrete salience ranking valuation may be non monotonic in price, so that making a cheap good k'' more expensive may decrease the salience of its price and thus increase its valuation.

[23] In particular, an increase in the observed price levels increases the salience of price provided observed prices are at or above the expected prices. This is the case in the Hastings and Shapiro example of demand for gas, where expected prices are known (they are specified to be the past period prices) and price shifts are defined relative to expectations. However, if observed prices are low relative to expectations, then a price hike boosts the relative salience of quality, by bringing observed prices closer to expected prices. Thus, if expectations are known (as in scenario 1) but do not necessarily coincide with observed prices (as in scenario 2), the model has further testable predictions.

pairwise choice. If there exist price shifts $\Delta p_a < \Delta p_b < \Delta p_c$ such that good 2 is chosen at prices $p_k + \Delta p_a$ and $p_k + \Delta p_c$ but good 1 is chosen at prices $p_k + \Delta p_b$, then good 2 is chosen at any price level $p_k + \Delta p$ satisfying $\Delta p \in [0, \Delta p_a]$ or $\Delta p \geq \Delta p_c$.

As a function of the price level, the quality of the chosen good is either flat (if the same good is chosen under all salience rankings), monotonic, or inverse-U-shaped (the case considered in the lemma). Prediction 2 can be interpreted as describing preference reversals between two goods within arbitrary choice sets. In this sense, data on consumers' full preference rankings, rather than simply choice, provide a better test of the model when expectations are not observed.

6 Misleading Sales

To illustrate how our model can clarify field evidence on context effects, we develop an application to sales. Retailers frequently resort to sales events as a means to sell their products. In 1988 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 (Varian 1980, Lazear 1986, Sobel 1984). 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 some 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 comparison 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. In particular, we illustrate the steps involved in developing a parsimonious specification of salience and sales, we discuss how salience naturally generates the misleading sales phenomenon, and we highlight two new predictions of our salience-based model of sales:

- In a store selling different qualities, misleading sales only boost demand for high quality goods, and this occurs at the expense of demand for lower quality goods.

- Misleading sales boost demand only for non-standard goods.

Recall from Section 2.2 the ingredients required to specify the model: typically described in a rational model are the set of attributes in the utility function (which must be separable), and the choice set faced by the consumer. Here we take the utilities derived from the goods' qualities to be known (see discussion in Section 5 about how quality can be elicited). As before, we assume that the salience function is homogeneous of degree zero, and that salience weighting is determined by salience ranking. Finally, we define the choice context, by which we mean price expectations about goods in the choice set.

Specifically, 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 has no substitutes and is only available at this store; we later consider the case of standard goods, which can be easily found at different stores. As in Section 5, 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. In this case, 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, with some probability π, the consumer is offered a given quality q at the sale price p_s rather than at the full regular price $p_f > p_s$. Since the consumer has rational expectations regarding the possibility of sales, his expected price is $\mathbb{E}[p] = \pi p_s + (1-\pi)p_f$. When deciding whether or not to buy the good, this expected price becomes part of the consumer's choice context, $\mathbf{C}_{sale} \equiv \{(0,0),(q,p_s),(q,\mathbb{E}[p])\}$.

Consider the standing of the option (q,p_s) in the new choice context $\mathbf{C}_{sale}$. The salience of quality is $\sigma(q, 2q/3)$, while the salience of price is $\sigma(p_s, \frac{\mathbb{E}[p]+p_s}{3})$. The central implication is that the retailer can manipulate the salience of price by manipulating the price discount p_s/p_f. In particular, we can establish

Proposition 6 *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* $p_f \in \left(q/\delta, q/\delta \cdot \frac{7-2\pi}{2-2\pi}\right)$.

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 maximum consumer valuation q/δ. This is because the consumer views the discount as a good deal, boosting his valuation of quality. The model limits the maximal regular price and thus the maximal discount to $p_s/p_f \geq (2-2\pi)/(7-2\pi)$. The reason is that, as in Proposition 5, an excessively high regular price makes prices salient, reducing the consumer's valuation.

By using regular prices to inflate consumers' valuations of goods, the misleading sales logic suggests that firms generate most returns during the sales events themselves, consistent with Ortmeyer, Quelch and Salmon (1991). This is in contrast with psychological model of sales based on loss aversion (e.g. Heidhues and Koszegi 2008) where firms make money on sales at regular prices, and sales events are used as a (potentially unprofitable) bait to generate loss aversion in consumers.

Up to this point, the central modelling decision was that of the choice context $\mathbf{C}_{sale}$. We now examine the robustness of this sales mechanism to the specification of the choice set. In particular, we consider the case where the outside option is a different good, with positive quality and price. This leads us to our first prediction, namely that a "misleading sale" is effective only for a high quality good.

Suppose that the store has a high quality good $h = (q_h, p_h)$ and a lower quality good $l = (q_l, p_l)$, 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).[24] 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 occasionally available (say, with probability $1-\pi$) also at a full price $p_{fh} > p_h$. Similarly, for the low quality good, the store can set a full price $p_{fl} \in (p_l, p_h)$ with the same probability. Suppose that when both goods are offered at "sale" prices p_l, p_h, the good h is sold if and only if it is quality salient, implying that condition (10) holds and $q_h - \delta p_h > 0$. We then find:

[24] A general analysis of sales policies, including the case where a store is able to choose the goods' prices, is left for future work.

Proposition 7 *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 h where the full price p_{hf} is suitably chosen. In contrast, a sale is innefectual for the low quality good: if the consumer chooses h in the absence of a sale, there exists no full price $p_{fl} \in (p_l, p_h)$ for l that makes h price salient, and l be chosen, in the context of the sale.*

It is always possible to engineer sales inducing the salient thinker to overvalue the high quality good h relative to 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 and the salience of price for the low quality good (since price is its relative advantage). As a result, the sale boosts the overvaluation of the high quality good and may cause an undervaluation of the cheaper good. Both of these effects imply that sales on the low quality good are unlikely to work.

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 h boosts the salience of its quality, increasing this good's valuation relative to l (regardless of the latter's salience ranking). Thus, sales should be effective specifically for high quality goods that, in the absence of sales, would be price salient. The same mechanism for the asymmetry is at work here as for decoy effects, since the high regular price effectively acts as a decoy.

Proposition 7 describes in a sales setting the asymmetry of the decoy effect established in Proposition 4. That prediction is supported by experimental data (Heath and Chatterjee 1995). The model's novel prediction on the asymmetry in the effectiveness of sales – as well as the implication that demand for the high quality good grows at the expense of that for the low quality competitor – has also been documented in the field. In their review on the literature on promotions, Blattberg, Briesch and Fox (1995) present this asymmetry as one of the stylised facts in the field of marketing.

Consider now our second prediction, namely that sales are unlikely to work with standard goods, for which market prices are well known. A 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 that each of these sellers implements the same misleading sales policy consisting of a regular price p_f and a sales price p_s, each occurring with probability $\pi = 1/2$ and where $p_f/p_s = k \in (1, 6)$ (see Proposition 6 above).

In this case, the consumer's choice context consists of $2N$ goods (two goods for each of the N sales), and the outside option of not buying $(0, 0)$. Formally, $\mathbf{C}_{sale}$ is given by $\{(0, 0), (q, p_s), \ldots, (q, \mathbb{E}[p])\}$ where (q, p_s) and $(q, -\mathbb{E}[p])$ are repeated N times, and $\mathbb{E}[p] = p_s \cdot (1 + k)/2$. For the items on sale, then, the salience of quality is $\sigma(q, q\frac{2N}{2N+1})$, and that of price is $\sigma(p_s, p_s\frac{N\cdot(3+k)/2}{2N+1})$. Due to homogeneity of degree zero, these expressions imply that when the number of sellers is sufficiently large, namely when

$$N > \frac{2 + \sqrt{3 + k}}{k - 1},$$

the items on sale have salient price [i.e., $\sigma(q, q\frac{2N}{2N+1}) < \sigma(p_s, p_s\frac{N\cdot(3+k)/2}{2N+1})$], rather than salient quality as in the non-standard good case of Proposition 6.

This result is intuitive, and holds for any magnitude k and frequency π of the sale. As the number of sellers N increases, the average quality $\overline{q} = q\frac{2N}{2N+1}$ in the choice set gets arbitrarily close to the quality q of the standard good. As a result, 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 standard good the salient thinker focuses on price because price is the attribute that varies most across sellers (almost by definition of standard goods)! This implies that a generalized policy of misleading sales does not work in the case of standard goods: because the inherent price variation induces consumers to focus on prices, it reduces their willingness to pay, so that the overall demand for the standard good falls.

This intuition 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 Storchmann (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. These are predictions that arise naturally in our model of salience.

7 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, expectations about prices also influence what decision makers are thinking about, and the choice context shaping the reference good is larger than the choice set. To discipline the model, we assume that price expectations are rational, but this assumption may need to be revised in some applications.

The second idea, which extends our earlier work on choice under risk (BGS 2012a), 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 new insight into several puzzles of consumer choice. The model makes stark predictions for choice in experimental settings, in which the choice context is fully controlled by the experimenter. By showing how irrelevant alternatives change the reference good – and thus the salient attributes of existing alternatives – the model accounts for two well-known violations of independence of irrelevant alternatives, namely decoy and compromise effects. But our mechanism also makes a novel prediction that decoy and compromise effects are asymmetric in that they differentially benefit more extreme goods (e.g. expensive, high-quality goods), a prediction that has strong experimental support. In the design of desirable goods, the model predicts a preference for some specialization as long as a minimum balance across attributes is provided.

By allowing expected prices to shape the reference good, the model has stark implications for responses to price changes. Consistent with previous research, salient thinkers may exhibit lower price sensitivity when price levels are high, but only when the high price level is fully expected. In contrast, the model makes the novel prediction that surprising price increases increase price sensitivity. In particular, the model accounts for context-dependent willingness to pay, exemplified by Thaler's celebrated beer example. Taken together, these predictions suggest that the salience mechanism can be seen as a simpler alternative to loss aversion in generating context effects.

We show how these results provide a unified way of thinking about field evidence previously described as mental accounting, in particular describe how consumers react to changes in the prices of individual goods or whole categories of goods. We provide a novel 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 also 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. We have noted throughout the paper a number of possible extensions and empirical tests, which we leave to future work.

References:

Azar, Ofer. 2007. "Relative Thinking Theory." *Journal of Socio-Economics*, 36 (1): 1 – 14.

Blattberg, Robert, Richard Briesch and Edward J. Fox. 1995. "How Promotions Work", *Marketing Science* 14(3), G122 – G132.

Bodner, Ronit and Drazen Prelec. 1994. "The Centroid Model of Context Dependent Choice", Unpublished Manuscript, MIT.

Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2012a. "Salience Theory of Choice under Risk", *Quarterly Journal of Economics*, 127 (3): 1243 – 1285.

Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2012b. "Salience and Experimental Tests of the Endowment Effect", *American Economic Review: Papers & Proceedings*, 127 (3): 1243 – 1285.

Chetty, Raj, Adam Looney and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence", *American Economic Review*, 99(4): 1145-1177.

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.

Fehr, Ernst and Antonio Rangel. 2011. "Neuroeconomic Foundations of Economic Choice – Recent Advances", *Journal of Economic Perspectives*, 25 (4): 3–30.

Feigenson, Lisa, Stanislas Dehaene and Elizabeth Spelke. 2004. "Core Systems of Number", *Trends in Cognitive Sciences*, 8 (7): 307 – 314.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, 121(2): 505-540.

Gabaix, Xavier. 2012. "A Sparsity-Based Model of Bounded Rationality, Applied to Basic Consumer and Equilibrium Theory", NYU Working Paper.

Gennaioli, Nicola and Andrei Shleifer. 2010. "What Comes to Mind", *Quarterly Journal of Economics*, 125 (4): 1399–1433.

Hare, Todd, Colin Camerer and Antonio Rangel. 2009. "Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System", *Science*, 324 (5927): 646–648.

Hastings, Justine and Jesse Shapiro. 2013. "Fungibility and Consumer Choice: Evidence from Commodity Price Shocks". *Quarterly Journal of Economics*, forthcoming.

Hauser, John and Birger Wernerfelt. 1990. "An Evaluation Cost Model of Consideration Sets". *Journal of Consumer Research*, 16 (4): 393 – 408.

Heath, Timothy and Subimal Chatterjee. 1995. "Asymmetric Decoy Effects on Lower-Quality versus Higher-Quality Brands: Meta-Analytic and Experimental Evidence". *Journal of Consumer Research*, 22 (3): 268 – 284.

Heidhues, Paul and Botond Koszegi. 2008. "Competition and Price Variation when Consumers are Loss Averse". *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 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 Miller. 1986. "Norm Theory: Comparing Reality to its Alternatives", *Psychological Review*, 93 (2): 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.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux.

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.

Kim, Jongmin, Nathan Novemsky and Ravi Dhar. 2012. "Adding Small Differences Can Increase Similarity and Choice." *Psychological Science*, 24 (2): 225 – 229.

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. 2013. "A Model of Focusing in Economic Choice", *Quarterly Journal of Economics*, 128 (1): 53 – 104.

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.

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.

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.

Stanovich, Keith and Richard West. 2000. “Individual Differences in Reasoning: Implications for the Rationality Debate,” *Behavioral Brain Sciences* 23 (5): 645 – 726.

Taylor, Shelley and Suzanne Thompson. 1982. “Stalking the Ellusive 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. 1972. “Choice by Elimination ”, *Journal of Mathematical Psychology*, 9 (4): 341 – 367.

Tversky, Amos and Danier 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.

A Proofs (For Online Publication)

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

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

Proof of Proposition 2. Consider an indifference curve characterized by $u(q, p) = q - p = u$. As in the text, order the elements of the choice set by increasing quality and price, so that (q_1, p_1) is the cheapest good. The goods' quality-price ratios satisfy $\frac{q_i}{p_i} = 1 + \frac{u}{p_i}$, and in particular the reference good $(\overline{q}, \overline{p})$ satisfies $\frac{\overline{q}}{\overline{p}} = 1 + \frac{u}{\overline{p}}$. As in the text, we assume that $\overline{\mathbf{q}}$ is not in the choice set.

1) $\frac{q_1}{p_1} > 1$ 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 $\overline{p}$ and have high quality/price ratios), while it is the relative disadvantage of expensive goods (whose prices are under $\overline{p}$ and have low quality/price ratios). Since the cheapest good is the best option along the salient price dimension, it is chosen and $k^* = 1$. Formally, all goods are undervalued, $u^S(q_i, p_i) = \frac{\delta q_i - p_i}{\delta + 1}$, but the cheapest good is the least undervalued.

2) $\frac{q_1}{p_1} < 1$ 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 and $k^* = N$. Formally, all goods are overvalued, $u^S(q_i, p_i) = \frac{q_i - \delta p_i}{1 + \delta}$, but the highest quality good is the most overvalued.

3) $\frac{q_1}{p_1} = 1$ 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 salient thinker evaluates each good correctly (as the rational agent) and is thus indifferent

between them. $\blacksquare$

Proof of Proposition 3. A sufficient condition for reversal between l and h is that good 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 $\Delta u = [q_h - q_l] - [p_h - p_l]$. Together, these conditions are equivalent to (10).

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

We are mainly interested in regime a), in which the decoy is located close to the other goods, i.e. $\overline{q} < q_h$ and $\overline{p} < p_h$, and it is a "bad deal", i.e. it has a low quality-price ratio. In fact, in this regime the condition that 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 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), $\overline{q}_d < q_h$, $\overline{p} > p_h$ and the condition on $q_h \cdot p_h$ yields $q_d < q_h\left[3p_h/\overline{p} - 1\right] - q_l$. Regimes c) and d) instead impose lower bounds on q_d.)

In regime a), h is quality salient so (10) guarantees it is preferred to l. To see that the alternative d is never chosen, two cases are distinguished: either d has higher quality and lower quality-price ratio than h, in which case it is price salient; or it has lower quality and lower quality-price ratio than 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 h is overvalued relative to d. Thus, a small enough δ can be found such that h is chosen. A sufficient condition for h to be chosen, for any δ, is that the decoy lies on a lower rational indifference curve than h. This is guaranteed for dominated d, and by continuity for some d with $q_d > q_h$ as well. In fact, given the assumptions that $\theta_1 = \theta_2$ and that 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 h has a relatively high quality price ratio, quality is salient in $\{l, h\}$ and h is chosen. Given the constraints $\overline{q} < q_h$ and $\overline{p} < p_h$, adding a decoy d to the choice set makes h price salient when it increases the quality price ratio of the average good to the level where $q_h/p_h < \overline{q}/\overline{p}$. However, this is excluded by the condition that the decoy is a "bad deal", namely $q_d/p_d < \max\{q_l/p_l, q_h/p_h\}$. ■

Proof of Proposition 4. Suppose the prices of all goods in $\mathbf{C}_C$ are shifted by a small $\gamma > 0$. Then the average price in $\mathbf{C}$ shifts by $\eta \cdot \gamma$, where η is the share of goods in $\mathbf{C}_C$. Consider the salience of price for goods in $\mathbf{C}_C$ which have price p^*, i.e. $\sigma\left(p^* + \gamma, \overline{p} + \eta\gamma\right)$. Diminishing sensitivity implies that salience decreases in γ whenever $\eta = 1$, or when $\eta < 1$ but $p^* < \overline{p}$. This is because in either situation the average payoff level increases but the difference between payoffs weakly decreases.

For salience to increase in γ, 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 $\eta < 1$ and $p^* > \overline{p}$. 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

$$\partial_\gamma \sigma\left(p^* + \gamma, \overline{p} + \eta\gamma\right) > 0 \Leftrightarrow \partial_\gamma \frac{p^* + \gamma}{\overline{p} + \eta\gamma} > 0$$

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

Proof of Proposition 5. The average quality in $\mathbf{C} \cup \{(q, p)\}$ is $\overline{q} = q\frac{2}{3}$. The average price is $\overline{p} = \frac{1}{3}\left[p + \mathbb{E}[p|\sigma]\right]$. Thus, the salience of quality and price of good (q, p) are, respectively

$$\sigma\left(1, \frac{2}{3}\right), \quad \sigma\left(1, \frac{1}{3}\left[1 + \frac{\mathbb{E}[p|\sigma]}{p}\right]\right)$$

It follows that quality is salient when

$$p \in \left(\mathbb{E}[p|\sigma] \cdot \frac{2}{7}, \mathbb{E}[p|\sigma]\right), \quad \text{or} \quad \mathbb{E}[p|\sigma] \in \left(p, \ \frac{7p}{2}\right)$$

Note that as p varies in this range, it can take values larger or smaller than the reference price $\overline{p}$. In turn, price is salient when

$$\mathbb{E}[p|\sigma] < p \ \text{ and } \ \mathbb{E}[p|\sigma] > \frac{7p}{2}$$

Recall the definition of willingness to pay:

$$\text{WTP}(q|\mathbf{C}) = \sup p \text{ s.t. } u^S((q,p)|\mathbf{C} \cup \{(q,p)\}) \geq u^S((q_0, p_0)|\mathbf{C} \cup \{(q,p)\}).$$

Consider first the case where the good is expensive relative to the reference price, $\mathbb{E}[p|\sigma] < p$. 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, WTP$= \delta q$ whenever $\mathbb{E}[p|\sigma] < \delta q$.

Consider now the case where quality is salient, so the good is cheaper than the reference price, $\mathbb{E}[p|\sigma] \geq p$, 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/δ, as long as it does not change the salience ranking: WTP$= \max\left\{\frac{q}{\delta}, \mathbb{E}[p|\sigma]\right\}$. As a consequence, for $\mathbb{E}[p|\sigma] \leq \frac{q}{\delta}$, WTP$= \mathbb{E}[p|\sigma]$. For $\frac{7q}{2\delta} > \mathbb{E}[p|\sigma] > \frac{q}{\delta}$, we find WTP$= \frac{q}{\delta}$.

Finally, consider the case $\mathbb{E}[p|\sigma] > \frac{7q}{2\delta}$. Now the reference price is so high that even at the highest possible price for the good, namely q/δ, its price is salient. As a result, WTP goes back down to δq. ■

Proof of Prediction 1. Under the "if" step of the Prediction, the consumer's choice changes from i to j as prices increase in parallel. This implies that a) there is a shift in the salience ranking of at least one of the goods i and j, which benefits the valuation of j

relative to i, and b) the shift in salience rankings is sufficient to change the consumer's choice. Because expected prices coincide with observed prices, diminishing sensitivity implies that under a price increase any shift in salience ranking is a shift from price- to quality- salience.[1] Since $q_j > q_i$, i.e. demand for quality increases with price level, it must be that good j becomes quality-salient, and thus overvalued relative to good i.

Turning to the "then" step of the Prediction, note that because $p_i < \overline{p}_0 < p_j$, uniformly increasing the prices of all other goods $k \neq i, j$, such that the resulting reference price equals p_j results in good j being strictly quality salient, while strictly increasing the salience of price for good i (thus precluding the case where good i goes from being price salient to quality salient). As a consequence, under the new price distribution good j is preferred over good i, so that good i is no longer chosen. Yet the characteristics of goods i and j have not changed, and the other goods have become more expensive. As a consequence, this is a violation of the axiom of independence of irrelevant alternatives. ■

Proof of Prediction 2. We start the proof with two observations. First, as observed prices uniformly shift by Δp, the reference price shifts by $\Delta p/2$, so that this process typically affects the salience rankings of the two goods. In particular, if the observed price of a good becomes very different from the expected price (either very low or very high), price might become salient. If instead the actual price is close to the expected price, then quality is salient. Thus, as both goods' prices increase by Δp, a given good's salience ranking follows (a subset of) the chain $P \to Q \to P$, where P, Q represent price- and quality-salience respectively. As a consequence, the set of transitions of salience rankings of the two goods as both goods' observed prices uniformly increase are described by (a subset of) one of the following chains:

$$PP \to QP \to \ \{PP \ \ or \ \ QQ\} \ \to PQ \to PP \tag{1}$$

$$PP \to PQ \to \ \{PP \ \ or \ \ QQ\} \ \to QP \to PP \tag{2}$$

Here the first letter denotes the salient attribute of good 1 and the second letter denotes the salient attribute of good 2. Not all possible rankings necessarily obtain. In particular, it

[1] Without loss of generality, we ignore the non-generic case of equal salience of price and quality, as well as the case where $q_j = \overline{q}$. The argument carries through to these cases as well.

may be impossible to have both goods be quality salient if their prices are too far apart.[2]

Our second observation is that the set of possible salience rankings can be ordered in terms of the difference in valuation between the lower quality good and the high quality good. Write Δv_{XY} as the difference in valuation of good 1 versus good 2 when good 1's attribute X is salient and good 2's attribute Y is salient.

$$\Delta v_{XY} = \begin{cases} (q_1 - q_2) - \delta(p_1 - p_2), & XY = QQ \\ (q_1 - \delta q_2) - (\delta p_1 - p_2), & XY = QP \\ (\delta q_1 - q_2) - (p_1 - \delta p_2), & XY = PQ \\ \delta(q_1 - q_2) - (p_1 - p_2), & XY = PP \end{cases}$$

For example, the overvaluation of good 1 is larger when the salience rankings are QP than when they are QQ, namely $\Delta v_{QP} > \Delta v_{QQ}$. It is easy to see that, under the quality and price rankings of goods 1 and 2, the relative valuations satisfy $\Delta v_{QP} > \Delta v_{QQ} > \Delta v_{PP} > \Delta v_{PQ}$.

Putting the two observations together, it is easy to check that only a certain set of shifts in demand for quality can occur under parallel price shifts. For instance, if $\Delta v_{PQ} > 0$ then good 1 is chosen under any salience ranking, so that the consumer chooses good 1 at any price (we say that demand for quality is flat as a function of price levels). Similarly, if $\Delta v_{QP} < 0$ good 2 is always chosen. If $\Delta v_{PQ} < 0 < \Delta v_{QP}$, then demand for quality along the entire chains (1) and (2) is single-peaked, namely of the form $1 \rightarrow 2 \rightarrow 1$ or $2 \rightarrow 1 \rightarrow 2$. However, if only a subset of these chains obtain under parallel price shifts with $\Delta p \geq -p_2$, then demand may be monotonic (or even flat). Because salience rankings are not observable, such restrictions are not predictable ex ante. Therefore, the testable prediction is that, if two changes in preferences have been observed between $\overline{p}_0 + \Delta p_1$ and $\overline{p}_0 + \Delta p_2$, and between $\overline{p}_0 + \Delta p_2$ and $\overline{p}_0 + \Delta p_3$, then there are no other preference changes at other price levels: demand for quality is constant (and equal) for $\Delta p \in [-p_2, \Delta p_1]$ and $\Delta p > \Delta p_3$.

Note that if the observer has some information about the expected prices, then further

[2] Moreover, the endpoints of the chain may not obtain either. For instance, as Δp becomes large, the salience of each good's price approaches $\sigma(\Delta p, \Delta p/2)$. If the salience function is homogeneous of degree zero, this equals a finite number, so that quality might be salient in this limit. Also, in each chain's middle step either PP or QQ obtain, but not both.

predictions can be made. Suppose the expected prices satisfy

$$\overline{p}_e < p_2 - \frac{p_1 - p_2}{2} \tag{3}$$

In this case, the lower quality good costs more than the reference price $(\overline{p}_e + \overline{p}_0)/2$, so any increase in the observed prices boosts the price salience of both goods. Moreover, since $p_1 > p_2$, ordering implies that $\sigma(p_1, \overline{p}) > \sigma(p_2, \overline{p})$, while diminishing sensitivity (together with the assumption that expected qualities coincide with observed qualities) implies that $\sigma(q_1, \overline{q}) < \sigma(q_2, \overline{q})$. As a consequence, if the low quality good is price salient, then so is the high quality good. Conversely, still assuming that (3) holds, if the high quality good is quality salient, so is the low quality good. These considerations restrict the possible salience rankings to the subset $QQ \to PQ \to PP$ of chain (1). In this case, suppose that under a price increase there is an increase in demand for quality. This means both goods are price salient – we are in the rightmost PP step of chain (1) – and there are no further shifts in demand for higher price levels. Alternatively, suppose that under a price drop respecting (3) there is an increase in demand for quality. This means both goods are quality salient – we are in the QQ step of chain (1) – and price cannot go lower without violating (3). ■

Proof of Proposition 6. As in the text, consider the choice context $\mathbf{C}_{sale}$ equal to $\{(0,0), (q, p_s), (q, \mathbb{E}[p])\}$. Consider the valuation of the good on sale, (q, p_s). The salience of its quality is (using homogeneity of degree zero) $\sigma\left(q, \frac{2q}{3}\right) = \sigma\left(1, \frac{2}{3}\right)$. The salience of its price is $\sigma\left(p_s, \frac{p_s + \mathbb{E}[p]}{3}\right) = \sigma\left(1, \frac{1 + \frac{\mathbb{E}[p]}{p_s}}{3}\right)$. Therefore, quality is more salient than price as long as $\frac{p_f}{p_s} \in \left(1, \frac{7-2\pi}{2-2\pi}\right)$. 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. ■

Proof of Proposition 7. The store can always make the high quality good quality salient by holding a sale with a full price $p_{fh} = (4 - \pi)p_h - 2p_l$ (in which case p_h coincides with the expected quality in the choice context, $\mathbb{E}[p]$).

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 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 h is quality salient and chosen by the consumer, holding the sale for l has no effect on the consumer's choice. ■

B Extension: Goods with Multiple Quality Attributes

In this Appendix, we extend the model to the case where goods are characterized by multiple attributes. We then study in detail the case where goods differ along two quality attributes.

As in Section 2, a consumer evaluates all $N > 1$ goods in a choice set $\mathbf{C}_{choice} \equiv \{\mathbf{q}_k\}_{k=1,...,N}$. Each good k is a vector $\mathbf{q}_k = (q_{1k}, \ldots, q_{mk}, p_k) \in \mathbb{R}^m$ of $m > 1$ quality attributes, where q_{ik} $(i = 1, \ldots, m)$ measures the utility that attribute i generates for the consumer. The case studied in the main text has $m = 1$.[3]

The consumer has full information about the attributes of each good and, absent salience distortions, evaluates $\mathbf{q}_k$ with a separable utility function:

$$u(\mathbf{q}_k) = \sum_{i=1}^{m} \theta_i q_{ik} - \theta_p p_k, \tag{4}$$

where θ_i is the weight attached to quality attribute i and θ_p is the weight attached to the numeraire in the valuation of the good. We normalize $\theta_1 + ... + \theta_m + \theta_p = 1$. Parameter θ_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 θ_i/θ_j is the rational rate of substitution among attributes j and i.

The choice context is defined similarly to the case where $m = 1$.

Definition 4 *The choice context is the set $\mathbf{C} = \mathbf{C}_{choice} \cup \mathbf{C}_e$, where $\mathbf{C}_{choice}$ is the externally given choice set while $\mathbf{C}_e = \{\mathbf{q}_k^e\}_{k=1,...,N}$ is the set of goods the consumer expects to find in the choice setting. We assume that:*

i) $\mathbf{q}_k^e$ shares the same non-price attributes of choice option $\mathbf{q}_k$, namely $q_{ik}^e = q_{ik}$ for $i = 1, ..., m$. The expected price p_k^e is the rational expectation of p_k, namely $p_k^e \equiv \mathbb{E}[p_k]$.

[3] The extension to a case where a good has multiple price components is straightforward.

ii) The choice context is summarized by a reference good $\overline{\mathbf{q}} = \{\overline{q}_1, \ldots, \overline{q}_m, \overline{p}\}$*, where the reference (or normal) level of attribute* i *is the average value of that attribute in* $\mathbf{C}$*, namely* $\overline{q}_i = \frac{1}{2N}\sum_k (q_{ik} + q_{ik}^e)$ *and* $\overline{p} = \frac{1}{2N}\sum_k (p_k + p_k^e)$*. The reference good* $\overline{\mathbf{q}}$ *need not be in* $\mathbf{C}$*.*

Given a salience function $\sigma(\cdot,\cdot)$, the salience distortions of decision weights are again defined similarly to the case where $m = 1$.

Definition 5 *Quality attribute* i *is more salient than quality attribute* j *for good* $\mathbf{q}_k$ *if and only if* $\sigma(q_{ik}, \overline{q}_i) > \sigma(q_{jk}, \overline{q}_j)$*. Quality attribute* i *is more salient than price for good* $\mathbf{q}_k$ *if and only if* $\sigma(q_{ik}, \overline{q}_i) > \sigma(p_k, \overline{p})$*. Let* r_{ik} *be the salience ranking of quality attribute* i *and* r_{pk} *the salience ranking of price for good* $\mathbf{q}_k$*, where the most salient attribute has rank 1. Attributes with equal salience receive the same (lowest possible) ranking. The salient thinker evaluates good* $\mathbf{q}_k$ *by transforming the weights* θ_i *attached to quality attribute* $i \in \{1, ..., m\}$ *and the weight* θ_p *attached to the numeraire into:*

$$\widehat{\theta}_i^k = \theta_i \cdot \frac{\delta^{r_{ikt}}}{\sum_j \theta_j \delta^{r_{jk}} + \theta_p \delta^{r_{pk}}} \equiv \theta_i \omega_i^k, \qquad \widehat{\theta}_p^k = \theta_p \cdot \frac{\delta^{r_{pk}}}{\sum_j \theta_j \delta^{r_{jt}} + \theta_p \delta^{r_{pk}}} \equiv \theta_i \omega_i^t \tag{5}$$

where $\delta \in (0, 1]$*. The salient thinker's evaluation of good* $\mathbf{q}_k$ *is given by:*

$$u^S(\mathbf{q}_t) = \sum_{i=1}^{m} \widehat{\theta}_i^t \cdot q_{it} - \widehat{\theta}_p \cdot p_k. \tag{6}$$

Having examined the tradeoff between quality and price in the main text, we now consider the trade-off between two quality dimensions. We show that diminishing sensitivity naturally creates a taste for goods delivering balanced utilities across different attributes: for unbalanced goods, the salient attributes are their shortcomings rather than their strengths. This mechanism is richer than loss aversion accounts and yields novel predictions.

Consider goods (q_{1k}, q_{2k}, p) that differ in their qualities but not in their prices, so that price is the least salient dimension (we assume price is deterministic, so $p = \overline{p}$). For notational convenience, we omit the price. In this setup, Definition 1 implies that q_{1k} is more salient than q_{2k} for good k if and only if $\sigma(q_{1k}, \overline{q}_1) > \sigma(q_{2k}, \overline{q}_2)$. Once more, the salience ranking of a good in quality-quality space is determined by its location relative to the reference $\overline{\mathbf{q}} = (\overline{q}_1, \overline{q}_2)$. Suppose that $q_{1k} > \overline{q}_1$ and $q_{2k} < \overline{q}_2$. Then, homogeneity of degree zero implies that the upside q_{1k} of good k is salient whenever $\sigma(q_{1k}/\overline{q}_1, 1) > \sigma(1, \overline{q}_2/q_{2k})$, which is equivalent to:

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

The salience ranking is determined by the quality-quality product $q_{1k} \cdot q_{2k}$. In this regard, 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. Unbalanced goods, which increase the level of one attribute at the cost of weakening the other, have low values of $q_1 \cdot q_2$. Balanced goods, whose strengths and weaknesses are comparable, have high values of $q_1 \cdot q_2$. We then show:

Proposition 8 *Let all goods in the choice context be located on a rational indifference curve, with reference good* $\overline{\mathbf{q}} = (\overline{q}_1, \overline{q}_2)$*. The consumer chooses the good* k *which is furthest from* $\overline{\mathbf{q}}$*, i.e. maximizes* $|q_{1k} - \overline{q}_1|$*, conditional on being more balanced than* $\overline{\mathbf{q}}$*, i.e.* $q_{1k} \cdot q_{2k} > \overline{q}_1 \cdot \overline{q}_2$*. If all goods are less balanced than* $\overline{\mathbf{q}}$*, the salient thinker chooses the most balanced good* k*, namely the good that maximizes* $q_{1k} \cdot q_{2k}$*.*

The salient thinker picks the good that is most specialized relative to the reference good, provided that good's weakness is not so bad that it is noticed. This choice trades off two forces. On the one hand, keeping the salience ranking fixed, the salient thinker tries to maximize the salient quality along the rational indifference curve. If the good is more balanced than the reference, its salient quality is its advantage relative to the reference. The salient thinker chooses the good which maximizes this advantage, which is measured by the distance $|q_{1k} - \overline{q}_1| = |q_{2k} - \overline{q}_2|$ from the reference. 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.[4] These effects imply that the consumer tends to be

[4]Thus, in quality-quality tradeoffs the salient 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 (Proposition 2). In contrast, along a quality-quality indifference curve one quality increases at the expense

attracted toward goods that are closer to the reference good $\overline{\mathbf{q}}$.

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 (2013)'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.

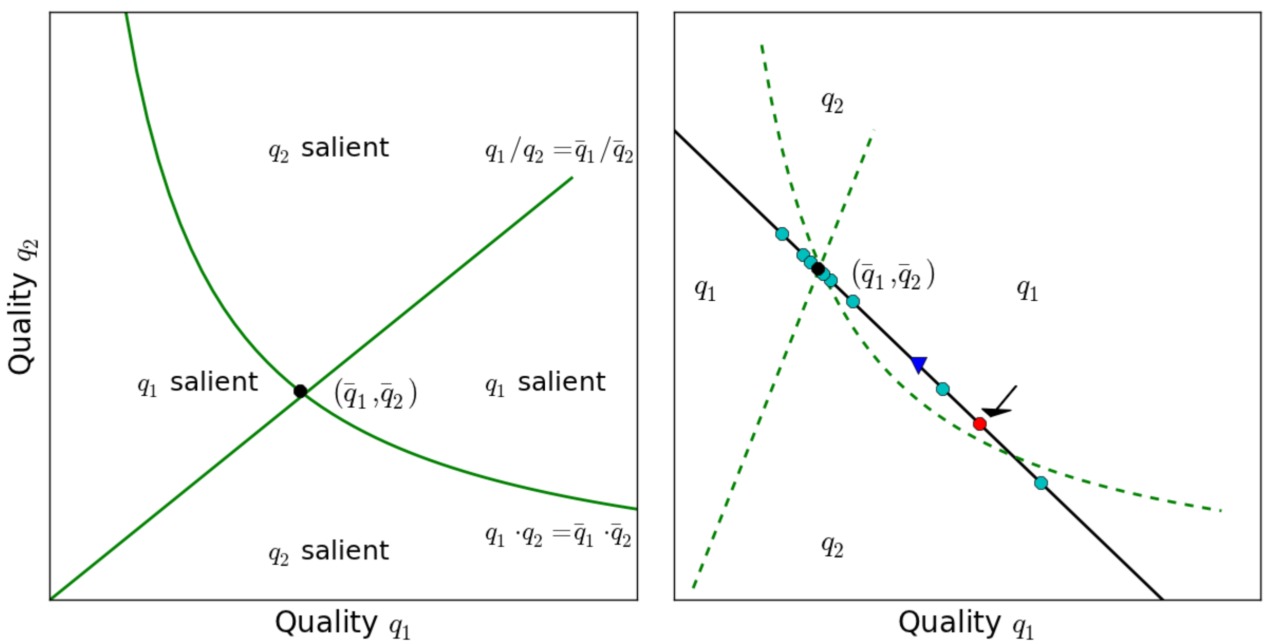

Figure 1: Salience ranking of goods with two quality attributes.

Proposition 8 is illustrated in the right panel of Figure 1. The solid downward sloping curve represents a rational indifference curve with utility level u, along which the choice set is distributed. The average good has a relatively low value $\bar{q}_1$ and a high value $\bar{q}_2$. The convex (dotted) curve represents the (choice-set dependent) iso-salience curve, namely the set of attribute combinations (q_1, q_2) for which $q_1 \cdot q_2 = \bar{q}_1 \cdot \bar{q}_2$. Crucially, the product $q_1 \cdot q_2$ is

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.

concave along the rational indifference curve, with a maximum at the middle good $\left(\frac{u}{2\theta_1}, \frac{u}{2\theta_2}\right)$ (denoted by a triangle). Thus the convex curve separates three regions of the rational indifference curve: a region of low $q_{1,k}$ where $q_1 \cdot q_2 < \bar{q}_1 \cdot \bar{q}_2$ (goods in this region have their weakest quality q_1 salient), a region of low q_2 where $q_1 \cdot q_2 < \bar{q}_1 \cdot \bar{q}_2$ (goods in this region have their weakest quality q_2 salient), and the intermediate range of goods which lie above the iso-salience curve. These goods have a high (balanced) quality product $q_{1,k} \cdot q_{2,k} > \bar{q}_1 \cdot \bar{q}_2$, as well as a higher level of q_1 than the average good. As a consequence, these goods have their q_1 quality salient. It is easy to see that such goods are overvalued relative to all other goods in the choice set. The most overvalued good is the one with the highest q_1 value (arrow) which is chosen.

To jointly characterize the salience ranking of all goods in a general choice set $\mathbf{C}$ we simply need to compute the reference attribute levels, and then place the goods in a diagram such as that of Figure 1 above. The Figure's left panel clearly shows that a good's quality q_{ik} is salient in regions where it is far from the reference quality level $\overline{q}_i$, with $i = 1, 2$, thus allowing us to develop visual intuitions for the role of salience in explaining choices.[5]

Proof of Proposition 8. Consider an indifference curve characterized by $u(q_1, q_2) = q_1 + q_2 = u$, where for simplicity we set $\theta_1 = \theta_2$. The average good $(\overline{q}_1, \overline{q}_2)$ also lies on the indifference curve, and good k's advantage relative to the reference good is salient whenever $q_{1k} \cdot q_{2k} > \overline{q}_1 \cdot \overline{q}_2$. The central point of the indifference curve $(u/2, u/2)$, which maximizes the product of qualities, satisfies $q_{1k} \cdot q_{2k} \leq \frac{u}{2} \cdot \frac{u}{2}$ for all k.

Let $\mathbf{C}_{bal}$ be the set of goods satisfying $q_{1k} \cdot q_{2k} \geq \overline{q}_1 \cdot \overline{q}_2$, where $\overline{\mathbf{q}}$ is the reference good in the choice set. Goods in $\mathbf{C}_{bal}$ have their advantages relative to $\overline{\mathbf{q}}$ salient. Importantly, since all such goods lie closer to the central point of the indifference curve than $\overline{\mathbf{q}}$, they have the same advantage relative to the reference. By diminishing sensitivity, this coincides with $\overline{\mathbf{q}}$'s weak attribute, namely the quality dimension in which $\overline{\mathbf{q}}$ delivers lower utility. Goods in $\mathbf{C}_{bal}$

[5] To identify the upward sloping curve, note that when $\mathbf{q}_k$ dominates the reference (i.e. $q_{1k} > \overline{q}_1$ and $q_{2k} > \overline{q}_2$), then q_{1k} is salient if and only if $\sigma\left(q_{1k}/\overline{q}_1, 1\right) > \sigma(q_{2k}/\overline{q}_2, 1)$, namely if and only if $q_{1k}/q_{2k} > \overline{q}_1/\overline{q}_2$. Instead, when $\mathbf{q}_k$ is dominated by the reference, its quality q_{1k} is salient if and only if $q_{1k}/q_{2k} < \overline{q}_1/\overline{q}_2$.

The figure thus also illustrates the quality price trade off considered in the main text. If we re-interpret the dimension q_2 as a price dimension, it follows that – in the regions where there is a trade-off between $\mathbf{q}_k$ and the reference good $\overline{\mathbf{q}}$, namely $q_{1k} < \overline{q}_1$, $p_k < \overline{p}$ or $q_{1k} > \overline{q}_1$, $p_k > \overline{p}$ – good $\mathbf{q}_k$'s advantage relative to the reference $\overline{\mathbf{q}}$ if salient if and only if $\mathbf{q}_k$'s quality-price ratio is higher than that of the reference.

maybe undervalued (if their weakness coincides with that of the reference) or overvalued. However, since they lie close to the central point, they are less affected by salience than the good lying outside $\mathbf{C}_{bal}$.

Consider now those goods that are less balanced than $\overline{\mathbf{q}}$, namely which lie outside $\mathbf{C}_{bal}$. Diminishing sensitivity implies that these good's disadvantages relative to $\overline{\mathbf{q}}$ are salient. Since any such good lies farther from the central point than $\overline{\mathbf{q}}$, its disadvantage relative to the reference coincides with its weak dimension. As a result all such goods are undervalued. Note that, within this set of goods, the more balanced goods closer to the centre of the indifference curve (namely, with higher $q_{1k} \cdot q_{2k}$) are preferred to the more extreme goods, because their salient disadvantages are less extreme.

To conclude, if $\mathbf{C}_{bal}$ is non-empty, the consumer chooses the good in $\mathbf{C}_{bal}$ which has the highest value along the reference's weak dimension. If $\mathbf{C}_{bal}$ is empty, then the chooses the good which maximizes $q_{1k} \cdot q_{2k}$. ■

C Continuous Salience Distortions

The dependence of valuation distortions on the salience ranking of different attributes (Definition 2) implies that the salient thinker's valuation can jump discontinuously at attribute values where salience ranking changes. Here we provide a continuous formulation where this behavior does not occur. Continuous salience distortions also allows to rule out non-monotonicity in valuation, which may sometimes arise in the salience ranking specification (which may even lead, in finely tuned examples, to a dominated good being preferred over a dominating good).

Take a choice context $\mathbf{C}$ characterized by a given reference good $(\overline{q}, \overline{p})$. We define the salient thinker's valuation of an individual good (q, p) to be:

$$u(q, p) = q \cdot w(q, \overline{q}) - p \cdot w(p, \overline{p}), \tag{7}$$

where w is a continuous weighting function encoding the properties of salience. We later offer a specification that makes this link transparent. Note that this formulation imposes two restrictions: i) salience weights are determined independently for different attributes, and ii) salient weights have the same functional form for all attributes.

The weighting function satisfies the properties of ordering, symmetry and homogeneity of degree zero. Formally, let $k > 0$ be the level of a good's attribute (either quality or price) and let $\overline{k}$ be the reference level of that attribute in a given choice context. Then:

$$\partial_k w(k, \overline{k})\big|_{k \geq \overline{k}} \;>\; 0 > \partial_k w(k, \overline{k})\big|_{k < \overline{k}}\,, \tag{8}$$

$$w(k, \overline{k}) \;=\; w(\overline{k}, k), \tag{9}$$

$$w(k, \overline{k}) \;=\; w(\alpha k, \alpha \overline{k}), \text{ for any } \alpha > 0 \tag{10}$$

That is, the weight attached to any attribute (quality or price) increases as the value of that attribute becomes more distant from its reference value. The property of reflection follows from the specification that w takes (positive) prices as arguments. As we saw in the text, ordering and homogeneity of degree zero together imply diminishing sensitivity of the weighting function (for positive quality and price levels). For convenience, we also assume that w is bounded.

Due to the assumed continuity of w, valuation in Equation (7) is continuous at any (q, p). For differentiable w, monotonicity in quality and price read as:

$$\partial_q u(q, p) \;=\; w(q, \overline{q}) + q \cdot \partial_q w(q, \overline{q}) \geq 0, \tag{11}$$

$$\partial_p u(q, p) \;=\; -w(p, \overline{p}) - p \cdot \partial_p w(p, \overline{p}) \leq 0, \tag{12}$$

We proceed in three steps: first, we derive the conditions under which – keeping the reference good fixed – valuation is monotonic. In other words, a cheaper good is perceived to have a price advantage over a more expensive good. Second, we examine when *valuation* exhibits diminishing sensitivity, namely when the price advantage of the cheaper good becomes less pronounced as prices increase (as in the store vs. restaurant example). Finally, we turn to violations of IIA, and show the workings of the decoy effect and of willingness to pay when salience weighting is continuous.

Using homogeneity of degree zero, write $w(q, \overline{q}) = f(q/\overline{q})$. Then the ordering property simply states that $f(x)$ gets larger as x gets further from 1, namely $f'(x) > 0$ for $x > 1$ and

$f'(x) < 0$ for $x < 1$. Moreover, symmetry implies that $f(x) = f(1/x)$ (in particular, $f(x)$ need not be differentiable at $x = 1$).

We now re-write the monotonicity conditions in terms of f and show under what conditions they are satisfied. Consider monotonicity in price. Then (12) becomes $f(p/\overline{p}) + p \cdot \partial_p f\,(p/\overline{p}) > 0$. Note that for $p > \overline{p}$, this condition is guaranteed by the ordering property, namely the second term is positive. As a consequence, monotonicity need only be checked for attribute values below the reference levels for which the second term is negative. Suppose $p < \overline{p}$ and p increases, while $\overline{p}$ stays fixed. Then we get

$$f\left[\overline{p}/p\right] > \frac{\overline{p}}{p} \cdot f'\left[\overline{p}/p\right] \tag{13}$$

Since p and $\overline{p}$ are arbitrary, the function $f(x)$ must be concave for $x > 1$.

As an example of a salience weighting function, consider

$$w(p, \overline{p}) = \frac{\left[1 + \sigma(p, \overline{p})\right]^{1-\delta}}{2}$$

where $\sigma(\cdot, \cdot)$ is a salience function that satisfies the properties of ordering, symmetry and homogeneity of degree zero (and diminishing sensitivity) which it receives from the weighting function w. Using $f(x) = [1 + \sigma(x, 1)]^{1-\delta}/2$, we can rewrite the monotonicity condition (13) in terms of σ as $x \cdot \partial_x \sigma(x, 1) < \frac{1+\sigma(x,1)}{1-\delta}$ for $x > 1$. Our standard salience function (4) satisfies this condition.

Consider now other properties of the model, starting from the store vs. restaurant comparison. Consider a pairwise choice between goods (q_h, p_h) and (q_l, p_l) where $p_h > p_l$ and where we denote $\overline{p} = (p_h + p_l)/2$. Then a uniform increase Δp in the level of prices induces the consumer to substitute toward the more expensive good provided the difference

$$(p_h + \Delta p) \cdot f\left((p_h + \Delta p)/(\overline{p} + \Delta p)\right) - (p_l + \Delta p) \cdot f\left((\overline{p} + \Delta p)/(p_l + \Delta p)\right)$$

decreases in Δp. Write $R_{\Delta p} = \frac{p_h + \Delta p}{p_l + \Delta p}$, with $\Delta p > 0$. Also, denote $r_{h,\Delta p} = \frac{2R_{\Delta p}}{1+R_{\Delta p}}$ and $r_{l,\Delta p} = \frac{1+R_{\Delta p}}{2}$ the arguments of the salience function for the expensive and cheap good, respectively. Note that $r_{h,\Delta p} < r_{l,\Delta p}$. The above expression can then be rewritten as:

$$(p_h + \Delta p) \cdot f(r_{h,\Delta p}) - (p_l + \Delta p) \cdot f(r_{l,\Delta p}), \tag{14}$$

which should decrease with Δp. Differentiating with respect to Δp, we find

$$\begin{aligned} & f(r_{h,\Delta p}) - f(r_{l,\Delta p}) + \\ & + \partial_{\Delta p} R_{\Delta p} \cdot \frac{p_l + \Delta p}{1 + R_{\Delta p}} \left[f'(r_{h,\Delta p}) \cdot r_{h,\Delta p} - f'(r_{l,\Delta p}) \cdot r_{l,\Delta p} \right] \end{aligned} \tag{15}$$

To analyze this expression, recall that $r_{h,\Delta p} < r_{l,\Delta p}$. The first line is negative by monotonicity of f. This is the direct effect of diminishing sensitivity of the salience function, which ensures that the salience of price is lower for the expensive good than for the cheap good. The second line captures instead that differential effect of price level in the price salience of the two goods. Because f is concave, this effect is larger (more negative) for the more expensive good, $f'(r_{h,\Delta p}) > f'(r_{l,\Delta p})$. In particular, a sufficient condition for (15) to be negative is that

$$\partial_x \left[f'(x) \cdot x \right] \leq 0$$

This holds as long as f grows at most as fast as the logarithmic function. In particular, it holds for our example $f = [1 + \sigma(x, 1)]^{1-\delta}/2$.

We now turn to the analysis of violations of IIA. We begin with the decoy effect. The workings of the decoy effect follow in a straightforward manner from the ordering property. To see that, consider again the context of a pairwise choice. As before, the price advantage of the cheaper good is given by

$$p_h \cdot f(p_h/\overline{p}) - p_l \cdot f(\overline{p}/p_l) \tag{16}$$

Suppose a decoy option (q_d, p_d) is introduced in the choice set, such that $q_d \geq q_h$ and $p_h > p_h$.

The resulting reference price is equal to $\overline{p}' = (p_h + p_l + p_d)/3$ which locates closer to p_h relative to $\overline{p}$. Then the price advantage of the cheaper good strictly decreases because of ordering, since the price p_h becomes less salient (the ratio $p_h/\overline{p}$ goes down) while the price

p_l becomes more salient (the ratio $\overline{p}/p_l$ goes up). Similarly, the quality advantage of the higher quality good decreases: since the reference quality moves closer to q_h, the quality salience of the high quality good decreases relative to that of the low quality good. The net effect on the relative valuation of the goods depends on which effect dominates: if the price advantage decreases more than the quality advantage, then the decoy benefits the high quality good. Intuitively, this holds when the reference price becomes close to p_h, while q_h is still significantly higher than $\overline{q}$.

To proceed, suppose for simplicity that $q_h = p_h$ and $q_l = p_l$. In particular, the consumer is indifferent between the goods in a pairwise choice. Suppose further that the effect of the decoy good is to change the reference good as $\overline{q} \to \lambda_q \overline{q}$ and $\overline{p} \to \lambda_p \overline{p}$, where λ_q, λ_p are small. Then one can show that the price advantage of good l decreases by more than the quality advantage of good h if and only if $\lambda_p > \lambda_q$, in particular if and only if the decoy leads to a drop in the quality price ratio. Note that this setting describes two possible types of decoy: a decoy for the high quality good, where e.g. $q_d \geq q_h$ and $p_d > p_h$, but also a decoy for the low quality good, where e.g. $q_d \leq q_l$ and $p_d > p_l$.

Finally, we turn to the determination of willingness to pay for quality. Let the choice context be $\mathbf{C} = \{(0,0), (q,p), (q,p_\gamma)\}$, where p_γ is the expected price of quality q in context γ (e.g. the store or the resort). Since the reference quality is $\overline{q} = 2q/3$, the salience weight of quality is $f(3/2)$. The salience weight for price is, in turn, $f(p/\overline{p})$, where $\overline{p} = (p + p_\gamma)/3$. According to the definition in the text, the willingness to pay $\text{WTP}(q)$ for quality q in the choice context $\mathbf{C}$ is the maximum price p such that $q \cdot f(3/2) - p \cdot f(p/\overline{p}) \geq 0$. In other words, $\text{WTP}(q)$ satisfies

$$\text{WTP}(q) \cdot f\left(1, \frac{p_\gamma/\text{WTP}(q) + 1}{3}\right) = q \cdot f(3/2) \tag{17}$$

To gain insight into this expression, note that the salience weighting on the LHS reaches its minimum $f(1)$ when $\text{WTP}(q) = p_\gamma/2$. Suppose p_γ is such that $p_\gamma/2 \cdot f(1) = q \cdot f(3/2)$. In this case, the willingness to pay is exactly $p_\gamma/2$, as can be seen by direct substitution into (17). Moreover, if $p_\gamma\ 2 \cdot f(1) < q \cdot f(3/2)$, it must be that $\text{WTP}(q) > p_\gamma/2$, since the left hand side of (17) increases as WTP rises above $p_\gamma/2$.[6] As a consequence, WTP(q) increases with p_γ. To see this, suppose (17) is satisfied and then increase p_γ. Then the willingness to pay increases in order to compensate the reduction in the salience weighting.

Consider now the case where $p_\gamma/2 \cdot f(1) > q \cdot f(3/2)$. A reasoning similar to the above shows that now WTP(q) $< p_\gamma/2$. Note that in this regime a solution to (17) always exists since the LHS goes to zero with WTP(q) (as long as f is bounded, as assumed). Moreover, as p_γ increases, the salience weighting increases as well, causing WTP(q) to fall.

Sumarizing, the condition (17) defines WTP(q) as a function of the expected price p_γ, taking q as given. This function is inverse-U-shaped, increasing with expected price p_γ for p_γ up to $q \cdot K$ (where $K = 2f(3/2)/f(1) > 1$) and decreasing with expected price above that. At its maximum value, willingness to pay satisfies WTP(q) $= q \cdot f(3/2)/f(1) > q$.

D Price Shocks and Consumer Demand

Hastings and Shapiro (2013) show that consumers react to parallel increases in gas prices by switching to cheaper (and lower quality) gasoline, and to parallel decreases in gas prices by switching to more expensive (higher quality) gasoline. Here we show how this pattern emerges in our model when consumers hold rational expectations for gasoline prices at the time of choosing which gasoline to purchase.

There are two grades of gas, with qualities $q_h > q_l$ and prices p_{ht}, p_{lt} at time t. At each t, the consumer must buy one unit of gas and must decide which grade to buy. Here, we assume that gas prices follow a random walk, so that the consumer's expectation for gas prices for the current period t is simply the realisation of prices in the previous period $t-1$. This captures the intuition that when the consumer chooses gas, he recalls gas prices from

[6] There can also be a solution to (17) below $p_\gamma/2$ but by definition WTP is the largest solution satisfying (17).

the last time he bought gas. [7] As a result, his choice context is:

$$\mathbf{C}_t = \{(q_h, p_{ht}), (q_l, p_{lt}), (q_h, p_{h,t-1}), (q_l, p_{l,t-1})\}.$$

Following Hastings and Shapiro (2013), we focus on parallel price shifts $p_{ht} - p_{h\ t-1} = p_{lt} - p_{l\ t-1} = \Delta_t$.

In the choice context $\mathbf{C}_t$, the reference quality and price are given by

$$\overline{q}_t = \frac{q_h + q_l}{2}, \qquad \overline{p}_t = \frac{p_{h,t-1} + p_{l,t-1} + \Delta_t}{2}.$$

Suppose that the two grades yield the same intrinsic utility to the consumer, namely $q_h - p_{ht} = q_l - p_{lt}$. In this case, demand is fully determined by salience: the consumer chooses the high grade gas if and only if its quality is salient. The salience function $\sigma(\cdot, \cdot)$ satisfies the usual properties of diminishing sensitivity, ordering and symmetry, as well as homogeneity of degree zero. The salience of quality and price for the high quality gas are:

$$\sigma(q_h, \overline{q}) = \sigma\left(\frac{2}{1 + q_l/q_h}, 1\right), \qquad \sigma(p_{ht}, \overline{p}_t) = \sigma\left(\frac{2}{1 + \frac{p_{t-1,l}}{p_{th}}}, 1\right). \tag{18}$$

The most intuitive case is one in which, after the parallel price change Δ_t, the high grade gas is still more expensive than the reference price $\overline{p}_t$. This condition is equivalent to $\Delta_t + (p_{ht} - p_{lt}) > 0$. It is satisfied as long as the price shock is not too negative between two visits at the gas station. We later discuss what happens when $\Delta_t + (p_{ht} - p_{lt}) < 0$.

From Equation (18), q_h is salient (and thus the high grade gas is chosen) when:

$$\frac{q_h}{p_{h,t-1} + \Delta_t} > \frac{q_l}{p_{l,t-1}}. \tag{19}$$

which is satisfied provided Δ_t is sufficiently low (it is always fulfilled for $\Delta_t + (p_{ht} - p_{lt}) = 0$).

The demand for low quality gas decreases, namely Equation (19) is more likely to hold,

[7] An alternative specification would be to assume a static price distribution. In this case the expected price would be fixed for all t. If realised prices are above the expected price (e.g. due to a temporary oil shock), then salience of gas price increases with the realised price, from which the Hastings and Shapiro evidence follows. By assuming instead that prices follow a random walk, we show that this prediction is very robust to assumptions about price paths.

when there is a sufficiently large drop in gas prices (i.e., Δ_t is sufficiently negative). The demand for low quality gas increases, namely Equation (19) is less likely to hold, when there is a sufficiently large hike in gas prices (i.e., Δ_t is sufficiently positive). In particular, suppose that in the previous two visits at the gas station the price of gas was stable, namely $\Delta_{t-1} = 0$. Then, the change in the demand for the low grade gas between $t-1$ and t as a function of the price change Δ_t is plotted in Figure 1 (where W_{t-1} is a constant determined below).

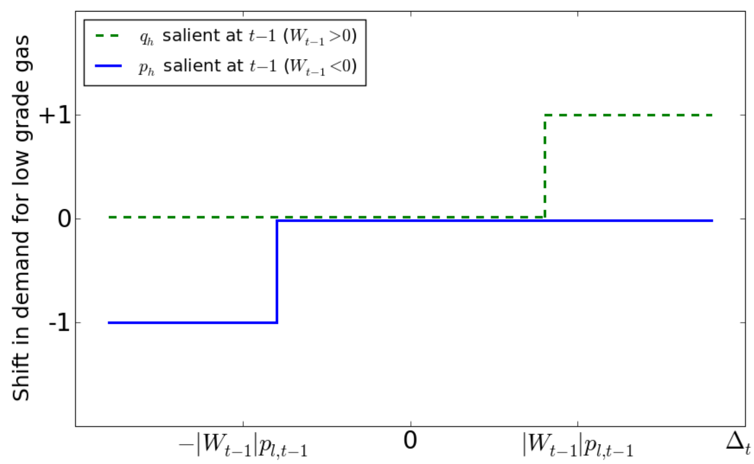

Figure 2: Price shocks and shifts in demand for the low grade gas.

Three features stand out:

- The demand for low grade gas tracks price changes. A sufficiently large price hike ($\Delta_t > 0$) increases the demand for low grade gas, while a sufficiently large price drop ($\Delta_t < 0$) decreases it. The intuition is that when the price of gas increases, the consumer views the current high grade price as a bad deal relative to yesterday. This renders its price salient. When the price of gas drops, the consumer sees the current high grade as a good deal relative to yesterday. This renders its quality salient. Thus, salience predicts history dependence in the demand for gas at given price levels.

- Demand changes only if the price change is sufficiently large. This is because small price changes do not affect salience.

- Demand is more sensitive to a given price change Δ_t when the price level $p_{l,t-1}$ is low.

This is because at lower price levels a given price change is more noticeable, due to diminishing sensitivity. Thus, salience predicts history dependence in the reaction of demand for gas to a given price change, even with linear utility.

Two further comments. First, consider large price drops such that $\Delta_t + (p_{ht} - p_{lt}) < 0$. In this case, it is still true that demand for the low grade gas decreases, but only up to a threshold drop $\widehat{\Delta} < 0$. For $\Delta_t < \widehat{\Delta}$ price becomes salient and thus the consumer again chooses the low grade gas. We can ignore this case, however, as for a reasonable difference of grade qualities q_h, q_l the required price drop $\widehat{\Delta}$ is of the order of the price level $p_{l,t-1}$ itself.[8]

Second, to fully appreciate the implications of history dependence, the model should be studied for all possible past price changes Δ_{t-1} (remember that here we restricted to the case $\Delta_{t-1} = 0$ for simplicity).

Let us go back to the determination of the threshold level W_{t-1}. To study the change in demand between $t-1$ and t we need to determine demand at $t-1$ when $\Delta_{t-1} = 0$. Iterating Equation (19) backward, the consumer picks the high grade gas at $t-1$ if and only if:

$$\frac{q_h}{p_{h,t-1}} > \frac{q_l}{p_{l,t-1}}. \tag{20}$$

According to Equations (19,20), the demand for high grade gas increases fom 0 to 1 when

$$\frac{p_{h,t-1}}{p_{l,t-1}} + \frac{\Delta_t}{p_{l,t-1}} < \frac{q_h}{q_l} < \frac{p_{h,t-1}}{p_{l,t-1}}.$$

This requires a sufficiently large price drop $\Delta_t < 0$. In contrast, the demand for high grade gas decreases from 1 to 0 when

$$\frac{p_{h,t-1}}{p_{l,t-1}} < \frac{q_h}{q_l} < \frac{p_{h,t-1}}{p_{l,t-1}} + \frac{\Delta_t}{p_{l,t-1}},$$

which requires a sufficiently large price hike Δ_t. To construct Figure 1, denote $W_{t-1} = \frac{q_h}{q_l} - \frac{p_{h,t-1}}{p_{l,t-1}}$. Condition (20) becomes $W_{t-1} > 0$, while condition (19) reads $\Delta_t < W_{t-1} \cdot p_{l,t-1}$. Note that the thresholds $|W_{t-1}| \cdot p_{l,t-1}$ increase in absolute value with the price level $p_{l,t-1}$.

[8] The precise threshold is $\widehat{\Delta}_t = \frac{1+\lambda}{3\lambda-1} p_{l,t-1} - p_{h,t-1}$, where $\lambda = q_h/q_l$. In particular, $\widehat{\Delta}_t = -p_{l,t-1}$ when $p_{h,t-1}/p_{l,t-1} = 4\lambda/(3\lambda - 1)$.