MEMORY, ATTENTION, AND CHOICE[*]

Pedro Bordalo

Nicola Gennaioli

Andrei Shleifer

Revised January 22, 2020

Abstract. Building on a textbook description of associative memory (Kahana 2012), we present a model of choice in which a choice option cues recall of similar past experiences. Memory shapes valuation and decisions in two ways. First, recalled experiences form a norm, which serves as an initial anchor for valuation. Second, salient quality and price surprises relative to the norm lead to large adjustments in valuation. The model unifies many well documented choice puzzles including the attribution and projection biases, inattention to hidden attributes, background contrast effects, and context-dependent willingness to pay. Unifying these puzzles on the basis of selective memory and attention to surprise yields multiple new predictions.

[*] We are grateful to Dan Benjamin, Paulo Costa, Stefano DellaVigna, Ben Enke, Paul Fontanier, Matt Gentzkow, Sam Gershman, Thomas Graeber, Michael Kahana, Spencer Kwon, George Loewenstein, Sendhil Mullainathan, Josh Schwartzstein, Jesse Shapiro, Jann Spiess, Linh To, Pierre-Luc Vautrey, and five anonymous referees for valuable comments. Gennaioli thanks the European Research Council for Financial Support under the ERC Consolidator Grant (GA 647782). Shleifer thanks the Sloan Foundation and the Pershing Square Venture Fund for Research on the Foundations of Human Behavior for financial support. Corresponding Author: Andrei Shleifer, email: ashleifer@harvard.edu, tel: 6174955046, Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138. Word Count: 17902. JEL: D01, D90.

1. Introduction

Memory plays a central role in even the simplest choices. Consider a thirsty traveler thinking of whether to look for a shop to buy a bottle of water at the airport. He automatically retrieves from memory similar past experiences, including the pleasure of quenching his thirst and the prices he paid before, and decides based on these recollections. Memory serves as an *anchor* for valuation, even before the consumer enters the shop and sees the product.

But memory does more than that. After the traveler enters the shop, seeing the water and its price triggers retrieval of similar past experiences with bottled water, including prices. In the neoclassical model of choice, these memories do not affect behavior, only intrinsic attributes do. Reality is arguably different. If the traveler is at an airport for the first time, he retrieves the prices he is used to paying downtown and is likely shocked by the much higher airport price. He might then refuse to buy, even if very thirsty. On the other hand, a seasoned traveler recalls the high prices seen at other airports in the past: the same high price looks normal to him, and hence acceptable. Here, memories serve as a reference guiding the *adjustment* of valuation to the observed price.

This dual role of memory is reminiscent of Kahneman and Miller's Norm Theory (1986). In Norm Theory, an event spontaneously triggers recall of past similar experiences, which are consolidated into a norm. An event similar to the norm is not surprising, and its evaluation is anchored to the norm. In contrast, an event that is very different from the norm generates a surprise, as in the case of the first-time traveler at the airport shocked by the abnormal water price. This idea is also related to a neuroscientific model of sensory perception, the "predictive brain model" (Hawkins and Blakeslee 2004; Clark 2013). The brain automatically forms mental representations of reality. If these representations are accurate, they are used as a guide to action. If they are inaccurate, our attention is engaged and the initial representation is adjusted.[1]

We show that this memory-based anchoring and adjustment mechanism naturally unifies many well documented choice puzzles such as the attribution bias, the projection bias, inattention to product attributes,

[1] For instance, as we open a door our brain unconsciously predicts the door knob's temperature. If the temperature is close to the usual, the experience is normal and behavior proceeds on the basis of the initial mental prediction. If instead the doorknob is unusually hot, we are surprised: our attention is engaged and our representation is adjusted.

and different reference point effects, including the background contrast effect. Furthermore, the structure of memory offers new predictions for how the strength of these effects is modulated by the consumer's past experiences and by the cues that trigger his recall process.

Our model works as follows. Consider a consumer with true preferences given by $q - p$, where q is a good's quality and p its price. When the consumer thinks about this good, he spontaneously retrieves past experiences of it from memory, and aggregates them into a norm for this good's quality and price (q^n, p^n). In some cases, as when the traveler decides whether or not to go look for water, his decision is entirely anchored to the value $q^n - p^n$ entailed by the norm. If instead the consumer sees the actual price and quality (q, p) of the good, as when the traveler sees water at the airport shop, his norm responds to these stimuli as well. His attention is drawn to discrepancies between the norm and reality, and valuation adjusts to:

$$q^n - p^n + \sigma(q, q^n) \cdot (q - q^n) - \sigma(p, p^n) \cdot (p - p^n). \qquad (1)$$

Adjustment depends on the discrepancy between the normal and the actual attributes through a salience weight $\sigma(x, x^n)$ that captures key features of sensory perception (Bordalo, Gennaioli and Shleifer 2012, 2013). When the seasoned traveler sees a price p that is close to his retrieved norm p^n, salience is low and adjustment is minimal. But for the inexperienced traveler, the same airport price p is much higher than his retrieved downtown norm p^n. The discrepancy is salient and the adjustment is large, causing valuation to collapse.

A key contribution of our model is to formalize retrieval from memory, and hence the norm (q^n, p^n), as a function of the consumer's past experiences and environmental cues. In economic models, a consumer retrieves all his past experiences, but also a lot of publicly available statistical information. In reality, memory is limited and selective. Early models recognizing these basic limitations are Gilboa and Schmeidler (1995) and Mullainathan (2002). In this paper, we go a step further, and adapt to an economic setting the textbook psychology model of associative memory (Kahana 2012).[2] Sensory stimuli such as the price and quality of a good, but critically also contextual stimuli such as location and time, act as cues that trigger recall of similar past experiences. Such recall is *associative*, meaning that a cue triggers recall of items from memory which are similar to that cue. Recall is also subject to *interference*, meaning that recall of a given item is weakened

[2] Some models feature recency effects but abstract from similarity and interference (Taubinsky 2014, Ericson 2017, Bushong and Gagnon-Bartsch 2019, Nagel and Xu 2018, Azeredo da Silveira and Woodford 2019).

or blocked entirely by memories that are more similar to the cue.[3] When thinking about white things in a kitchen, the cue "milk" makes it more likely that "yogurt" is retrieved, but less likely that "flour" is retrieved. The implication is that that the recalled experiences of a good can be very different from a statistically average version, leading to substantial malleability in what is considered "normal".

Similarity based recall tends to retrieve norms that are well adapted to current conditions. The norm is usually accurate, surprise is minimal, and choice is stable (or even rational). Departures from rationality arise when a database of selected experiences or a misleading contextual cue retrieve distorted norms, which anchor valuation or cause artificial surprise.

In Section 2 we show that this framework unifies elements of backward looking (Kahneman and Tversky 1979, DellaVigna et al. 2017) and rational expectations (Koszegi and Rabin 2006) reference points. The price norm of a traveler going for the second time to the airport is still influenced by the disproportionate number of downtown experiences in his memory database. After enough airport visits, however, the airport location selectively triggers the recall of high airport prices, and interferes with the recall of low downtown prices. Effectively, the consumer eventually has two well adapted "rational expectations" norms for water: an expensive one at the airport, and a cheap one downtown.

In Section 3 we show how the interplay between the memory database and contextual cues accounts for choice puzzles involving the valuation of future benefits or costs. The attribution bias reflects the fact that a biased set of experiences in the memory database tends to bias valuation. For example, a person who went to an amusement park in bad weather would have a lower valuation of a subsequent visit than someone who went in good weather (Haggag et al. 2018). The projection bias reflects the role of contextual cues, which can be misleading. For example, people buy more convertible cars on sunny days (Busse et al. 2015), perhaps because the sunny weather cues the retrieval of similarly sunny days and fun drives in the past, interfering with

[3] These principles are illustrated by two experimental paradigms. In *item recognition* tests, subjects assess whether given words are part of a previously shown list. These words cue recall from the list stored in memory. These tests show the role of similarity because i) the probability of recall is higher for items that belong on the list (so the cue is more similar to the item), and ii) subjects are more likely to mistakenly recognize words that are similar to a list member (they recognize yogurt when milk is on the list). In *cued recall* tests, subjects retrieve words that are pairwise associated with a cue, having previously been shown lists of word pairs. These tests show the role of interference: if the cued word appears in many word pairs, recall of each association is less likely (Anderson and Reder 1999). Here interference is stronger for items that are more similar, so similarity shapes cued recall as well.

the recall of driving in the snow. Finally, the combination of a selected database and misleading cues sheds light on persistently neglected attributes. The tendency of consumers to neglect sales taxes (Chetty, Looney and Kroft 2009) can be explained by the fact that previous experiences with the good are dominated by prices observed in the aisle context (sales taxes are included on the whole basket of purchases at the checkout). These prices, and not the prices inclusive of taxes at the checkout, spontaneously anchor valuation and choice.

In Section 4 we show that adjustment to surprise relative to the norm helps explain several puzzles involving the valuation of actual attributes. When norms are well adapted, i.e. close to actual attributes, valuation is strongly anchored to the norm. This yields insensitivity of choice to actual attributes in the sense that small differences from the norm are neglected. However, a selected database or a misleading cue can retrieve norms that are far from the actual choice. Over-reaction to such surprises creates what is known as contrast effects. Insensitivity to stimuli is a key difference between our model of valuation and Prospect Theory, which only features contrast around the reference point. We summarize a good deal of evidence from marketing, economics, and choice experiments pointing to the existence of both insensitivity and contrast depending on the cue. We also focus specifically on contrast effects and show how the model sheds light on several extremely puzzling findings such as Thaler's (1985) beer on the beach experiment and Simonsohn and Loewenstein's (2006) evidence on movers, but also describe the limits of these effects.

Table I summarizes how memory mechanisms – databases and cues – lead to several departures from rationality through their effects on norms and valuation.

Table I: A Summary of Selected Model Implications

Phenomenon	**Cause**		**Effect**	
	Cue	**Memory Database**	**Norm**	**Valuation**
Attribution bias		biased database (good weather only)	recall high utility experiences in good weather	forecast high utility
Projection bias	current environment (good weather)	balanced database (all weather)		
Persistent neglect of attributes	tax-free price, aisle location	mostly tax-free prices in aisle	recall tax-free price	tax-free price
Background contrast effect	high airport price	biased database (downtown prices)	recall downtown prices, as if backward-looking	surprise: "this is so different"
		balanced database (some high prices)	recall airport prices, as if forward-looking	assimilation: "looks like something I've seen before"

Table I also highlights the key source of new predictions: variation in the cues or in the memory database generates different norms and shapes behavior. For instance, the attribution bias is weakened if bad weather is cued, and the first-time flier would be less surprised at the airport if reminded of the price of water at restaurants because that cue would bring to mind similarly high prices. These predictions do not arise with rational expectations reference points, which already integrate all relevant information the agent has.

Additional predictions are due to the fact that the effect of the same cue depends on a person's set of experiences. For instance, in the projection bias the "sunny day" cue should exert a stronger influence on a consumer from Chicago than on one from Miami: because the Chicago consumer has ample experience of both good and bad weather, good weather acts as a potent cue that interferes with the recall of bad weather, drastically shifting the consumer's norm from the baseline. The dependence of the projection bias on personal experiences does not arise in existing theories (Loewenstein, O'Donoghue and Rabin 2003), in which this distortion is due to a mechanical excess persistence of current utility. Similarly, contrast effects should be weaker for consumers who experience more price variability: seeing a high price would disproportionately bring to mind high prices seen in the past, reducing the negative surprise from the current high price.

We study in detail the implications and new predictions of our approach. But the broad point is that memory offers a powerful unifying force. Disparate effects emerge naturally from our anchoring and adjustment model because memory is not just an additional ingredient that we add to existing models, with the aim of explaining some new facts. Rather, memory is an inevitable part of how we think and form valuations. Recall has a well-understood structure, which in turn predicts empirical regularities. Using these insights to understand economic choice delivers unification, clarity, and new testable predictions.

2. A Model of Memory

We first describe the memory database, the process of cued recall, and the formation of memory-based norms. We also discuss how norms relate to reference points.

Episodic memory is a database of past choice experiences. Experiencing a quality-price option (q, p) in context c creates a trace $e = (q, p, c)$ in the consumer's database. Components q and p identify the hedonic attributes of an option, and we assume throughout that true utility is given by $q - p$. In addition, c captures non-hedonic attributes present during encoding, such as location or time. The experience is broader than purchasing. Considering the good in a shop, seeing its price in an advertising campaign, or being told by a friend about it, all leave traces that are potentially available for recall.[4] For simplicity, we abstract from rehearsal of past options (Mullainathan 2002). We assume context c to be cardinal. In reality, c is multi-dimensional and categorical. The model can be extended to capture this richness.

The memory database at time t is summarized by a good-specific distribution $F_t(q, p, c)$ measuring the frequency with which past experiences of this good entailed a quality below q, a price below p, and a context below c. As new experiences come in, the distribution is updated. When context c captures calendar time, the set of times stored in memory expands. To simplify, we focus on a stable dimension c such as location, and on frequently repeated situations in which $F_t(q, p, c)$ has converged to an invariant distribution $F(q, p, c)$. This restriction can also be viewed as focusing on a single static choice, abstracting from the evolution of the database. We use the shorthand $F(e)$ for $F(q, p, c)$.

Because the database is good-specific, we rule out some associations. For instance, a bottle of water may prime a substitute category such as soft drinks. The price may bring to mind other goods one can purchase with the money. To capture these phenomena, recall may be defined over the universe of goods and experiences, but we do not deal with this here. We also abstract from the possibility that surprising events may be more easily retrieved, as in the "peak-end" rule of recall (Kahneman et al. 1993).

2.1 Cues, Similarity, and Norms

The current choice (e.g. "evaluating a bottle of water"), which we denote with subscript t, comes with a cue κ_t and with a database $F(e)$ of relevant experiences. Often, the cue is the full current experience of observing a bottle of water q_t at a price p_t in a location c_t, corresponding to $\kappa_t = e_t = (q_t, p_t, c_t)$. This case captures the traveler who sees quality and price at the airport shop. Other times, only context is observed, so

[4] Whether decision makers process and remember prices is an important topic in the marketing literature. Dickson and Sawyer (1990) survey shoppers in a supermarket about their knowledge of the prices of the goods in their baskets. In their survey, 21% of shoppers do not recall the price but 56% state a price that is within 5% of the correct price.

that $\kappa_t = c_t$. This is the traveler thinking whether to look for the shop, where only the "airport" cue is available. Context is always observed. Distinguishing different cues helps the model make more precise predictions across situations. We implicitly assume that an initial cue (such as "thirst") primes the good-specific database ("water"). In a more complete model, this step can also be formalized.

The cue κ_t stimulates recall of similar past experiences, where similarity is the distance measured along the set of attributes defining the cue κ_t. We follow the standard assumption in psychology (Kahana 2012), and define similarity in terms of the Euclidean distance.[5]

Definition 1 *The similarity of* $e \equiv (q, p, c)$ *to the cue* κ_t *is given by the multiplicatively separable distance:*

$$S(e, \kappa_t) \equiv S_1\big(\lambda_{q,t}|q_t - q|\big)S_2\big(\lambda_{p,t}|p_t - p|\big)S_3(|c_t - c|) \tag{2}$$

where the indicator $\lambda_{x,t}$ *equals 1 if attribute* x *is part of the cue* κ_t *and zero otherwise, and* $S_k\colon \mathbb{R}_+ \to \mathbb{R}_+$ *is decreasing for* $k = 1,2,3$. *The exponential specification takes the form:*

$$S(e, \kappa_t) = \exp\big\{-\delta\big[\lambda_{q,t}(q_t - q)^2 + \lambda_{p,t}(p_t - p)^2 + (c_t - c)^2\big]\big\}, \tag{3}$$

where $\delta \geq 0$ *captures the importance of similarity in recall.*

Multiplicative separability implies that the relative similarity of two experiences to the same cue is only shaped by the dimensions along which they differ. This property sharply characterizes norms. We often use the specification in Equation (3), which follows Kahana (2012). A context cue $\kappa_t = c_t$ retrieves past experiences based only on contextual similarity. A full cue $\kappa_t = e_t$ recruits past experiences based on similarity along all attributes. Past experiences are activated to different degrees, depending on similarity with κ_t. We model this process as a cue-driven change of measure in the historical distribution $F(e)$.

Definition 2. T*he memory weight of experience* e *after the cue* κ_t *is given by:*

$$w(e, \kappa_t) = \frac{S(e, \kappa_t)}{\int S(\tilde{e}, \kappa_t) dF(\tilde{e})}. \tag{4}$$

[5] Equation (2) follows multidimensional scaling (Torgerson 1958) in which the weights capture the unequal salience of different attributes. Tversky (1977) highlights cases in which similarity does not follow geometric properties.

The quality and price norms for cue κ_t are the similarity weighted average quality and price:

$$q^n(\kappa_t) \equiv \int q\, w(e, \kappa_t) dF, \qquad p^n(\kappa_t) \equiv \int p\, w(e, \kappa_t) dF. \tag{5}$$

As in Kahneman and Miller (1986), the norms $q^n(\kappa_t)$ and $p^n(\kappa_t)$ aggregate past experiences filtered according to similarity with the cue. The norm satisfies two properties. First, it weighs more similar experiences more heavily. Second, the weight attached to an experience decreases in the similarity of other experiences with the cue κ_t, because $w(e, \kappa_t)$ denotes relative similarity. This captures interference, whereby more similar memories block less similar ones (Kahana 2012).[6] With the similarity function of Equation (3), a higher δ captures stronger interference of similar traces with dissimilar ones.

According to Equation (5), the norm is tilted toward experiences most similar to the cue. To characterize the implications of this idea, we focus on the simplest case in which only one hedonic attribute varies across experiences. In the following Proposition we assume, without loss of generality, that this is the price attribute, so that both the actual quality and its norm are fixed at q. We can then show:

Proposition 1. *Denote by $p^n(c_t)$ and $p^n(p_t, c_t)$ the price norm when the cue is context or context and price, respectively. Assume that prices and context are independent in $F(e)$. In this case, the observed context is irrelevant for norms. In addition, denoting by $\overline{p}$ the average price in the database, we have:*

i) When the only cue is context, the price norm is the unconditional average experienced price, $p^n(c_t) = \overline{p}$.

ii) When the cue is context and price, the norm is $p^n(p_t, c_t) = p^n(p_t)$, where $p^n(p_t)$ is the price norm prevailing in the hypothetical case in which the cue is only price. If the marginal distribution of prices entailed by $F(e)$ is symmetric and unimodal, the norm $p^n(p_t)$ lies in between p_t and $\overline{p}$.

When context is uncorrelated with price, it has no effect on either the price norms or choice. This usefully implies that researchers should focus on measuring only contextual variables that are correlated with price (and/or quality), and can neglect the rest. The experiences of a consumer purchasing water downtown

[6] Equation (5) yields the well documented laws of recency and repetition. The "contextual drift" hypothesis states that c_t moves slowly over time (e.g., our state of mind changes slowly) so context cues recent experiences. In turn, the law of repetition follows from the fact that the distorted measure $w(e, \kappa_t) \cdot dF(e)$ attaches a larger weight to experiences with higher frequency $dF(e)$, which thus influence the norm more than less frequent ones do.

accumulate in contexts that differ in a myriad of attributes such as location, layout, etc. This bewildering variation, however, is unrelated to the price of water. When cued with a specific store the consumer recalls past experiences at similar stores but those entail the same average price as would be retrieved in another location. In this case, the norm is context-insensitive, and depends only on the average past price $\overline{p}$. This is similar to adaptive, backward looking reference points, but also coincides with rational expectations reference points because here context is uninformative of price.[7]

In contrast, as highlighted by property ii), the price norm is sensitive to a price cue, because the latter triggers recall of similar prices. The norm still depends on the average past price $\overline{p}$, but due to similarity it also adjusts toward the current price p_t. Similarity is thus a mechanism of "on the fly" adaptation to the current data. Comparing again our model to reference points, memory based norms are more flexible than both backward looking and rational expectations reference points. Indeed, the latter do not depend on available cues, such as the current price. Ex post adaptation is evident in the following special case.

Corollary 1 *If the marginal distribution $dF(p)$ is Gaussian with mean $\overline{p}$ and variance π^2, and the similarity function satisfies (3), then the measure $w(p, \kappa_t)dF(p)$ is Gaussian with variance $\pi^2/(1+2\delta\pi^2)$ and mean:*

$$p^n(p_t) = \frac{\overline{p} + 2\delta\pi^2 p_t}{1 + 2\delta\pi^2}. \tag{6}$$

Equation (6) is illustrated in Figure I, where the green line depicts the norm conditional on price p.

[7] More generally, to test our model one need not measure all cues in the environment, but rather to identify the cues that most strongly correlate with quality and price and hold the rest constant. This requirement is no different from any other model, both rational (e.g., what information do people see?) or behavioural (e.g., in Haggag et al (2018) the weather during the visit to the park is not the only aspect influencing utility). As we highlight throughout the paper, the role of cues is not a shortcoming but rather adds new predictions relative to other models.

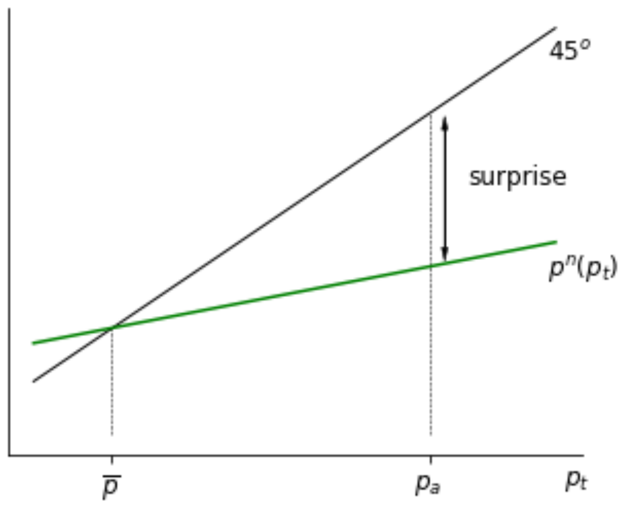

Figure I
Price norm and price surprise.
The figure plots the norm p^n (in green) as a function of observed price p_t when the database is Gaussian and context is uncorrelated with price (Equation 6).

In equation (6) the norm is a weighted average of the past and current price. This norm increases with the observed price, but less than one for one due to the disproportionate recall of the average (and modal) past price $\overline{p}$. When seeing the high price of water at the airport p_a for the first time, recall of the downtown price $\overline{p}$ brings the norm down. This entails a big surprise, as shown in Figure I.

The norm adjusts more to the observed price, including the high airport price, when δ is higher, that is, when recall of similar past prices interferes more with retrieval of different, even if frequent, prices. In Figure I, this means that when δ is higher the green line is closer to the 45^o line that goes through $\overline{p}$. Evidently, a steeper norm in Figure I reduces the surprise associated with seeing a high price p_t. Critically, the norm also adjusts more to price if past price variability π^2 is higher. In this case, a consumer seeing a high price recalls many past instances of similarly high prices, which interfere with the retrieval of low prices. The norm is steeper, the surprise is smaller. Similarity based recall thus implies that reaction to a cue is individual specific and stronger for those individuals who have more numerous past experiences similar to the cue. In general, we say that when δ and π^2 are higher the norm adapts better to the current price p_t.

Consider next the empirically more interesting case where price and context are correlated. In this case, norms adjust and hence adapt to the current context.

Proposition 2 *Suppose that $F(e)$ is Gaussian with mean $(\overline{p}, \overline{c})$, variances of price and context π^2 and γ^2, and correlation ρ between them. Using the similarity function in Equation (3) the price norms are given by:*

$$p^n(c_t) = \frac{\overline{p} + 2\delta\gamma^2 \mathbb{E}_F(p|c_t)}{1 + 2\delta\gamma^2}, \tag{7}$$

$$p^n(c_t, p_t) = \frac{\overline{p} + 2\delta\gamma^2 \mathbb{E}_F(p|c_t) + [2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1-\rho^2)]p_t}{1 + 2\delta\gamma^2 + [2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1-\rho^2)]}. \tag{8}$$

The norms in Proposition 2 adjust to context through the term $\mathbb{E}_F(p|c_t)$, which denotes the average price observed in c_t in the database $F(e)$. A traveler entering the airport recalls past airport visits and the high price associated with them, because in the airport context $\mathbb{E}_F(p|c_t)$ is higher (Equation 7). Contextual similarity creates a form of statistical conditioning. This conditioning is reminiscent of rational expectations reference points (Koszegi and Rabin 2006): upon seeing context c_t, the consumer retrieves the average price seen in this context $\mathbb{E}_F(p|c_t)$, which coincides with the rationally expected price when either the consumer's database has converged to the true distribution, or when he is learning. As Figure II illustrates, relative to the downtown context, the price norm at the airport is shifted upwards to reflect the higher expectation $\mathbb{E}_F(p|c_t)$. This implies that a price $p_a \gg \overline{p}$ is no longer surprising at the airport.

Even though selective retrieval goes in the right direction, it does not reflect optimal use of the consumer's information. First, because the consumer is cued not only by context, but also by the good itself, his norm in Equation (7) is still anchored on his experiences $\overline{p}$, reflecting the backward looking nature of norms. Second, if price p_t is also a cue, it also shapes retrieval, reminding the consumer of prices similar to itself, as in Equation (8). This helps the consumer adapt to the realized price, as shown by the bold green line in Figure II, but causes a further departure from the rationally expected price $\mathbb{E}_F(p|c_t)$. Finally, adjustment to current context $\mathbb{E}_F(p|c_t)$ is greater when experienced contextual variability γ^2 is high, and adjustment to current price p_t is greater when experienced price variability π^2 is high. This is because greater variability of context or price cues facilitate interference with the retrieval of more frequent experiences in memory, captured by $\overline{p}$.

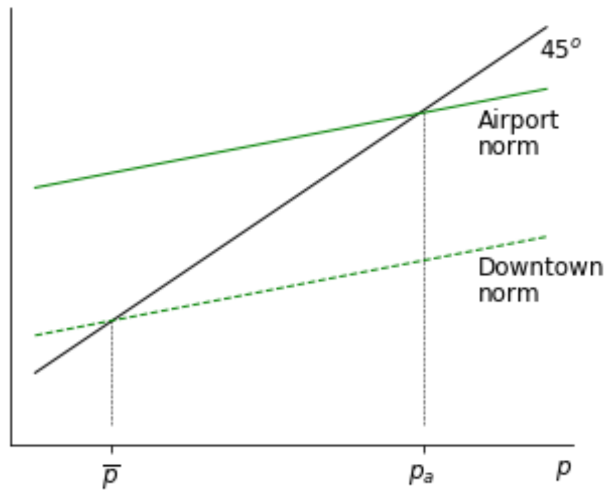

Figure II
Price norms and context.
The figure plots the norm p^n as a function of observed price p and where $c_{downtown}, c_{airport}$ (in bold and dotted green, respectively) identify different values of $c \in \mathbb{R}$ when context is correlated with price (Equation 8).

As we show in Section 4, ours is a theory of reference-dependent choice in which memory based norms shape valuation similarly to reference points. We have already noted that memory based norms capture elements of backward looking reference points such as Kahneman and Tversky's (1979) "status quo" and DellaVigna et al.'s (2017) mechanically adaptive reference points, via the average past price $\overline{p}$ in Equation (8). But these norms also capture forward looking elements similar to Koszegi and Rabin's (2006) rational expectations references through the conditional objective expectations $\mathbb{E}_F(p|c_t)$. Unification through memory yields new predictions for when backward or forward-looking elements should dominate. In situations where prices are stable, the database is populated by the few very frequent experiences. Memory-based norms are then strongly anchored to these experiences, and behave like backward looking reference points. Even if the traveler who is at the airport for the first time is told that prices at the airport are high – so that under rational expectations his reference price would immediately and fully adjust – he still retrieves his modal experiences with low prices downtown. In Equation (8), this occurs when π^2 and γ^2 are low, so that the price norm is strongly shaped by past experience $\overline{p}$.

On the other hand, as the consumer's experience with different prices in different contexts grows, norms become flexible in a forward looking way, getting closer to rationally expected reference points. After many airport experiences, the traveler's memory database becomes populated with high prices there. When cued by high prices at the airport, recall of these experiences interferes with recall of low downtown prices, causing the norm to adjust upward. In Equation (8), this is captured by high π^2 and γ^2, which cause the norm to be barely influenced by past experience $\overline{p}$. Unlike with mechanically adaptive expectations, the norm is not updated globally: there is a norm for the airport and a norm for downtown, and both are driven by contextual similarity and selective recall.

A key implication of associative memory is that norms respond to irrelevant cues. The current price cues similar past prices, causing the norm to adjust (Equation 8). This "ex post" adaptation of norms does not occur in other models, including rational expectations reference points. Moreover, the same associative nature of memory implies that an irrelevant contextual cue c_t can also change the norm, and affect choice. In Thaler's (1985) experiment, reminding a beachgoer that the beer comes from a resort triggers associative recall of high prices, raising the willingness to pay. Recency effects are another example of irrelevant contextual similarity. Recent prices or wages easily come to mind since recent experiences are close to the current one on the time dimension, and influence judgment even if they are normatively irrelevant (DellaVigna et al. 2017).[8]

We next investigate the link between memory and choice in two settings. In Section 3 we consider pure anchoring, which occurs when the consumer does not observe hedonic attributes, so norms fully shape valuation. In Section 4 the consumer observes hedonic attributes, so valuation is anchored to norms, but also adjusts to surprise relative to norms.

3. Memory and Choice I: Anchoring

A large body of work documents systematic biases in the assessment of the future quality or price of a good. Work on the attribution bias (e.g., Haggag et al. 2018) or on experience effects (e.g., Malmendier and Nagel 2011) shows that the evaluation of the future benefits of, say, a stock investment is unduly influenced by the past personal experiences with it. Work on the projection bias (e.g., Conlin, O'Donoghue and Vogelsang 2007) shows that the assessment of the value of warm sweaters increases under normatively

[8] This effect can be easily captured in our model by including calendar time as a dimension of context, or by formalizing contextual drift (Kahana 2012). That is, assuming that context at t is the combination $\alpha c_t + (1-\alpha)c_{t-1}$ with $\alpha < 1$. Mullainathan (2002) offers an early discussion of recency effects in economics.

irrelevant current conditions, such as cold days. Finally, work on shrouded attributes (e.g., Gabaix and Laibson 2004) or inattention (e.g., Chetty, Looney, and Kroft 2009) shows that, when valuing goods, consumers neglect important attributes not immediately available to them. In all these cases, the good's attributes are not observed, so valuation is distorted due to misleading mental representations.

Our model unifies these phenomena by viewing mental representations as the product of selective memory, which creates two sources of bias. First, because norms are based on past experiences, a biased database creates a biased valuation. Second, because recall is associative, norms are tilted towards past experiences most similar to the cue (even a spurious cue), and neglect experiences that are less similar to the cue due to interference. This approach implies that valuation biases are individual-specific, because they are shaped by personal experiences.

To see these effects, suppose to begin that neither the quality nor the price of a good is observed. The only cue available to the consumer is the choice context itself, namely $\kappa_t = c_t$. In this case, valuation is pinned down by the norm and decision value is $q^n(c_t) - p^n(c_t)$. Consider the simplest case in which only quality is uncertain, such as when a consumer assesses the utility of buying a warm sweater, or of the returns on a stock investment. The assessed quality is then the cued norm which, from Equation (7), is given by:

$$q^n(c_t) = \frac{\overline{q} + 2\delta\gamma^2 \mathbb{E}_F(q|c_t)}{1 + 2\delta\gamma^2}. \tag{9}$$

There are three determinants of quality evaluation:

1. *Past experience*: higher average experienced quality $\overline{q}$ increases the good's estimated quality.
2. *Cued context*: exposure to a context c_t associated to higher average quality $\mathbb{E}_F(q|c_t)$ increases the estimated future quality of the good.
3. *Variability of experiences*: higher variability of experiences γ^2 (or stronger interference δ) increases the malleability of quality valuation to a contextual cue c_t.

We next describe how these determinants account for and unify the effects documented in previous work.

3.1 Biased Database and the Attribution Bias

According to the "attribution bias", the valuation of a good to be consumed in the future is unduly influenced by the context of past experiences with it. For instance, consumers expect higher utility from going to an amusement park if during a past visit to the park the weather was good (Haggag et al. 2018). In our model, the attribution bias reflects a biased database of past experiences, the first term in the numerator of equation (9). When assessing the value of going to the amusement park, a consumer retrieves his own past experience from a database $F(q)$. If consumer i experienced better weather at this amusement park than consumer j, then his memory-based valuation is higher, $\overline{q}_i > \overline{q}_j$ because i's experiences are disproportionally formed in good weather contexts when the valuation is higher. Even if consumers know in principle that weather affects the enjoyment of the amusement park, memories are retrieved at "face value" because of the spontaneous association between the park and the pleasure of a past visit.

Other findings are consistent with an impact of selected databases on judgments. Several papers document that individuals' assessments of expected inflation are based on the goods they buy frequently (Georganas, Healy and Li 2014, Cavallo, Cruces, Perez-Truglia 2017), and their expectations about aggregate outcomes such as home prices or unemployment are based on their own recent experiences (Kuchler and Zafar 2019). When cued to recall price changes, subjects retrieve the average experienced price change which is tilted towards the most frequently bought goods. As illustrated in Equation (9), frequent experiences dominate judgments unless the cue has a strong similarity with other experiences.

Another piece of evidence comes from the experience effects documented by Malmendier and Nagel (2011): individuals who have experienced low stock market returns are less willing to take financial risk, and report worse returns expectations, than individuals who have experienced higher stock market returns (see also Malmendier and Nagel 2016). This finding runs counter to the idea that individuals form rational expectations of future returns using all publicly available data. Equation (9) goes some way toward accounting for this effect: if an investor i has had better stock market experiences than j, namely $\overline{q}_i > \overline{q}_j$, he will also expect a higher stock valuation today. However, our current model does not fully explain this phenomenon. In particular, it cannot explain why a few disastrous experiences exert such a strong effect despite the many more numerous experiences of good stock returns. Wachter and Kahana (2019) build a more complete theory of experience effects by allowing for stronger encoding of experiences that are more extreme or that occur earlier in life. For simplicity we abstract from these effects here.

3.2 Cued Context and the Projection Bias

Conlin, O'Donoghue and Vogelsang (2007) show that catalog orders of conventional cold-weather items spike in very cold days, and items ordered in such days are more likely to be returned. Busse et al. (2015) show that consumers are more likely to buy a convertible car if the weather is sunnier on the day they test-drive it, even if they have already owned a convertible in the past. Chang, Huang, and Wang (2016) find that on days in which air pollution is high, there is a spike in the purchases of health insurance in China, but consumers subsequently cancel their contracts if air quality improves. In these findings, consumers appear to unduly weigh current conditions when estimating future benefits, and remarkably do so even in cases where they have sufficient experience to know the value of the good under different conditions.

Selective retrieval yields the second term in the numerator of equation (9). Upon seeing an extreme context c_t, say cold weather, memory retrieves an extreme prediction $\mathbb{E}_F(q|c_t)$ of the utility of a warm sweater, which raises valuation.[9] In principle, the consumer has enough experience to form a stable and accurate valuation of clothes by retrieving warm days also. However, he does not access it because recall of the experience of cold weather interferes with recall of the experience of warm weather. Of course, when the weather improves, the consumer retrieves a lower predicted utility $\mathbb{E}_F(q|c_t)$, potentially causing him to regret the prior choice and return the sweater.

In this account, the projection bias reflects a valuation focused on specific contingencies selected via similarity-based recall. Factors influencing the availability of certain thoughts and the strength of interference should then modulate the projection bias. Defining the projection bias from current context c_t as the difference between the valuation in Equation (9) and the average memory-based norm $\overline{q}$, we have:

[9] A related phenomenon is the anchoring heuristic (Tversky and Kahneman 1973), whereby judgments of unfamiliar quantities can be tilted towards irrelevant anchors. As an element of context, an anchor may act as a cue that retrieves target instances of similar magnitude. This may also help explain why an anchor whose magnitude is unreasonable for the target question does not increase the effect on judgment (Mussweiler and Strack 2001).

Corollary 2. *The average change in valuation caused by observing c_t is given by:*

$$q^n(c_t) - \overline{q} = \frac{2\delta\gamma^2}{1 + 2\delta\gamma^2}[\mathbb{E}_F(q|c_t) - \overline{q}] = \frac{2\delta\gamma^2}{1 + 2\delta\gamma^2}\frac{\gamma}{\pi}\rho(c_t - \overline{c}).$$

This implies, ceteris paribus, that: i) the magnitude of the projection bias from observing a high cue c_t is lower for consumers who have experienced high average context $\overline{c}$, and ii) projection is more sensitive to the cue for consumers who have experienced higher contextual variability γ^2.

Prediction i) addresses availability: when test driving a convertible on a sunny day, consumers coming from cold regions (and hence having a low $\overline{c}$) should ceteris paribus increase their valuation of the car more than consumers coming from warm regions. This is because for the former consumers more cold weather experiences come to mind, so cueing them with sunny weather changes their assessments more. As Busse et al (2015) find, Miami residents do not rush to buy convertibles on a sunny day in the same way as Chicago residents do. Prediction ii) instead relies on interference. The projection bias is higher for consumers who have experienced stronger variability γ^2 in climatic conditions. Again, warm weather cues a resident in Chicago to recall many such experiences, and interferes with recall of cold winter days. This may explain why sales of convertibles in Chicago go up significantly on surprisingly warm days, even in November (Busse et al 2015).[10] The typical November cue is bad weather, so an occasional good weather cue sparks the retrieval of good weather experiences, boosting the valuation of a convertible car.

Loewenstein, O'Donoghue and Rabin (2003) model the projection bias not as one in the representation of future contingencies but rather as mis-predicted preferences in the form of a perceived excess persistence of tastes. Formally, they postulate that in current state c the projected utility q of the good in a future state c' is given by $q(c'|c) = (1 - \alpha)q(c') + \alpha q(c)$, where $\alpha > 0$ captures excess persistence in current preferences. Given that the future weather c' is not known, the consumer computes the expected projected utility as:

$$\mathbb{E}[q(c'|c)] = (1 - \alpha)\mathbb{E}[q(c')] + \alpha q(c).$$

[10] These examples recall Schacter's (2007) observation that imagining the future is fundamentally similar to recalling the past: "a crucial function of memory is to make information available for the simulation of future events" (p. 659).

Unlike our model, this approach allows no role for the memory database: the estimated value of the good is mechanically anchored to the current quality $q(c)$ while in our model it is anchored to past personal experiences that are retrieved by the cue, including the average past experience $\overline{q}$ (Corollary 2). In addition, the mapping from current tastes and future utility is not mediated by a fixed parameter α, but depends on personal experiences. Our mechanism can thus explain the differentially strong effect of weather in Chicago relative to Miami. More broadly, mis-predicting preferences is fundamentally different from selectively retrieving contingencies, because selective memory influences valuation over and above the current personal utility state.[11] For instance, reminding the Chicagoan that winter is coming may well cause him to pass on the convertible on a sunny November day. Chang, Huang and Wang's (2018) evidence on demand for insurance responding to current air pollution can be seen through this perspective as well. Relatedly, news about natural disasters – for example, floods in Florida – increase demand for insurance against such hazards (Slovic, Kunreuther, White 1974, Gallagher 2014) even when such news carry little new information. Instead, news may act as reminders of contingencies that were previously not top of mind.[12]

Experimental evidence shows that, in line with our model, judgments about the future are shaped by cues that evoke past experiences. Bornstein and Norman (2017) find that choices among risky alternatives can be significantly altered by showing subjects images that co-occurred with past gains, even if these images are uninformative about the current probability of gain. Bornstein and Norman (2017) also find neural evidence consistent with a memory mechanism: subjects' behavioural responses to the cue are commensurate with the reinstatement in the brain's visual areas of patterns associated with the cue. Enke, Schwerter and Zimmerman (2019) show that selective retrieval shapes valuation by explicitly associating certain images with good news, and other images with bad news, about a hypothetical asset. When assessing the asset, subjects overreact to good news that occur with images associated with past good news, and underreact to the same good news that

[11] In particular, our model connects the attribution and projection biases. Contextual cues can be used to strengthen projection at the expense of attribution. For instance, a consumer who visited an amusement park with good weather will moderate his valuation of the park if cued with "bad weather".

[12] This mechanism seems to also be at the heart of advertising: ads for cars prominently show beautiful mountain roads, when instead most driving is commuting in traffic. These advertising strategies make sense with interference but are harder to explain with persistence of current utility. See Section 3.4 for further discussion.

occur with images associated with past bad news, consistent with the idea that context (an image) cues recall of past news associated with it.

3.3 Biased Database and Cued Context: Inattention and Shrouding

A large literature highlights consumer inattention to shrouded product attributes (Gabaix and Laibson 2004). Chetty, Looney and Kraft (2009) find that displaying full prices, inclusive of sales taxes, in supermarket aisles reduces demand, even though consumers correctly recall sales taxes across a range of products when asked directly about them. Our model sheds light on these phenomena because it endogeneizes which pieces of information fail to come to mind, and when this neglect persists despite experience, as in the case of add-on fees or taxes. Such inattention arises in our model due to a combination of the two forces described in the previous two subsections, selected database and misleading cues, as captured by the two terms in the numerator of equation (9).

The findings of Chetty, Looney and Kraft (2009) can be accounted for by the two mechanisms of selective memory: i) the price database of consumers is flooded with prices observed many times in the aisles (which do not include the tax) and ii) when thinking of whether to buy a product, the aisle location and the price of the good cue retrieval of similar prices experienced in the same context. According to Equation (8) the normal cost $p^n(c_t, p_t)$ retrieved in the context $c_t = aisle$ when the price p_t is observed is given by:

$$p^n(aisle, p_t) = \frac{\overline{p} + 2\delta\gamma^2\mathbb{E}_F(p|aisle) + [2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1-\rho^2)]p_t}{1 + 2\delta\gamma^2 + [2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1-\rho^2)]}.$$

Here $\overline{p}$ is the average experienced price of the good, which includes experiences of seeing pre-tax prices on the aisles of different shops, and the arguably less numerous experiences of tax inclusive prices at the counter when checking the receipt. As per point i) above, the database $\overline{p}$ insufficiently accounts for the tax because sales taxes are not visibly associated with the specific good. But there are two other sources of tax neglect related to point ii). First, the current price p_t seen on the aisle cues recall of similar prices. Second, the aisle context retrieves an average aisle price $\mathbb{E}_F(p|aisle)$, which also does not include the tax. We do not normally see taxes in the aisle, and so do not think about them (just like we do not think about tax and tip when picking dishes from a restaurant menu). These forces entail neglect of hidden sales taxes and are ultimately due to the contextual dissociation between the price seen in the aisle and the full price at the counter. This dissociation hinders recall of taxes, especially in the aisle context.

In our model, consumers neglect sales taxes because the good itself is not a strong enough cue for them. This implies that attributes like fees are likely to be neglected when buying and paying are stored as separate experiences in memory. Some papers show that payment methods that decouple buying and paying prevent accurate recall of prices at the moment of purchase (Finkelstein 2009, Soman and Gourville, 2001, Chatterjee and Rose 2011), even though consumers may actually know the correct prices. Neglect of attributes may be less persistent when the purchase decision cues recall of hidden fees. To take Gabaix and Laibson's (2004) running example, a first time buyer of a printer may forget about the cost of replacing the cartridge. We expect, however, that a second time buyer would be much less likely to do so. The painful experience of overpaying for replacing the cartridge is now directly associated with the good, facilitating recall. This mechanism is also consistent with the bill shock on hidden fees such as overage charges in telephone usage (Grubb and Osborne 2015).

3.4 Reminders, aka "Cueing Selective Retrieval"

Reminders and information provision improve decisions in many settings, including savings, loan repayments, medication adherence, and gym attendance (Karlan et al. 2016, Cadena and Schoar 2011, Calzolari and Nardotto 2019). In the rational approach, reminders close the gap between a forgetful, biased assessment and the rational representation. However, evidence on the efficiency of reminders is mixed, in that reminders often fail to impact the perceived value of a choice option. Our model offers a different way to think about this problem, sheds light on some puzzling evidence, and offers new predictions.

The fundamental implication of equation (9) is that successful reminders should identify contexts that selectively retrieve desirable qualities of the choice, i.e., the situations in which the choice is most valuable. To illustrate, consider advertising. Many advertising campaigns cue context c_t rather than quality per se. For beer, Budweiser ads cue friendship, Corona ads cue young people partying on a beach. These ads evoke a context c_t in which the valuation of the given beer is maximized.

The impact of such reminders is stronger than that for rational consumers. For the latter, the reminder brings to mind the average value of the action. In our model, in contrast, the selective reminder interferes with recall of less desirable experiences, in line with the memory-jamming view of advertising in Shapiro (2006).[13] In the case of Budweiser, the recall of pleasurable experiences with friends interferes with recall of other conditions, such as getting drunk and throwing up at a party, leading to a higher quality norm. In the case of Corona, the ads evoke fun and sex, rather than the taste, presumably also helping the norm. In a fully rational model of information provision, recall is perfect and so these strategies are not effective.

This logic has the obvious implication that ads should cater to their audiences: the Budweiser drinker presumably associates beer more strongly with lasting friendship than with beach parties. But this logic goes further: because the effect of a reminder depends on what memories it brings to mind, it can generate very different responses in people with different experiences. At the extreme, in our model of selective recall, exposure to information can make people diverge in their norms and assessments. Suppose consumers i and j are identical in all respects except that i has experienced more beach parties than j. Assuming that such experiences are enjoyable, consumer i values beer on average more than consumer j. Cueing these consumers with a Corona ad, then, boosts valuation of beer by bringing beach experiences to mind. But for consumer i, the impact of the cue is disproportionately larger: not only does he have more numerous experiences that are similar but these interfere with other memories. As a result, i's and j's valuations of the beer after the ad are more different than before. In this sense, the ad "works" for consumer i but less so for consumer j.[14] Memory-based norms thus generate interesting implications for when disagreement arises and how it persists, which we leave for future work.

[13] As Shapiro (2006) shows, this view helps explain several stylized facts about advertising, such as the fact that highly familiar brands continue to advertise and that advertising intensity often grows as brands mature.

[14] Applying this logic to the case of experience effects (Malmendier and Nagel 2011), the same low stock return may especially depress the stock valuation of an investor who had worse experiences than another. In this instance, the two consumers respond in the same direction but with different strengths. It is possible, however, that the memory databases of different consumers differ in ways that cause them to react in opposite directions to the same cue. One such possibility arises when cued context exhibits positive correlation ρ for one consumer and negative correlation for another. For instance, priming public spending might induce a rightwing voter to think about misuse of public funds and a leftwing voter to think about poverty alleviation, enhancing their disagreement over taxes.

4. Memory and Choice II: Adjustment

Many choice puzzles involve the valuation of given attributes (q, p). Even when the consumer receives all the normatively relevant information, memory may still affect valuation. In the background contrast experiments (Simonson and Tversky 1992), for example, subjects are more likely to buy a good if they have previously experienced the same good with a higher price. We can address such evidence by analyzing how memory affects valuation of observed price and quality according to Equation (1).

In this case, the decision value of the observed price and quality depends on two forces. First, the retrieval of similar past prices and qualities anchors valuation to a norm. This mechanism tends to create a form of inattention to current attributes and thus rigidity in choice. Second, heightened attention to price and quality that are surprising relative to the memory norms tends to create excess sensitivity of valuation and contrast effects in choice, accounting for the Simonson and Tversky (1992) experiments. By clarifying when either force dominates, memory unifies different findings and yields new predictions.

Consider a consumer observing hedonic attributes, $\kappa_t = (q_t, p_t, c_t)$. This cue triggers retrieval of price and quality norms $p^n(\kappa_t)$ and $q^n(\kappa_t)$. Valuation is anchored to these norms, but also adjusts to the surprises $q_t - q^n(\kappa_t)$ and $p_t - p^n(\kappa_t)$ as in Equation (1). We now define this process more precisely.

Definition 3. *Given a cue $\kappa_t = (q_t, p_t, c_t)$, the decision utility from good (q_t, p_t) is given by:*

$$q^n(\kappa_t) - p^n(\kappa_t) + \sigma\big(q_t, q^n(\kappa_t)\big) \cdot [q_t - q^n(\kappa_t)] - \sigma\big(p_t, p^n(\kappa_t)\big) \cdot [p_t - p^n(\kappa_t)]. \quad (10)$$

The first two terms describe as before the decision utility given by the norm; the last two terms describe the adjustment due to the surprising price and quality. This adjustment is guided by the Weber-Fechner law of perception. We assume, as in Bordalo, Gennaioli and Shleifer (2012, 2013), that salience $\sigma(x, y) \geq 0$ increases in the proportional difference between the attribute and its norm. Specifically, $\sigma(x, y) \geq 0$ is symmetric, homogeneous of degree zero, and increasing in x/y for $x \geq y > 0$. When the surprise is small, little attention is paid to it, and here we assume $\sigma(y,y)=0$. When the surprise is large, attention is directed toward it, but we bound this effect by assuming that $\lim_{x/y\to+\infty}\sigma(x/y,1)=\sigma>1$.[15]

Equation (10) introduces two changes in the specification of salience relative to our prior work. First, salience weights are attached to deviations from the norm, not to attribute levels themselves as in Bordalo, Gennaioli and Shleifer (2013). Despite this change, the key properties of our original model are preserved.[16] Second, in Bordalo, Gennaioli and Shleifer (2013) attention is drawn to attributes as a function of their relative salience, so attention to price directly reduces attention to quality. This externality plays no role here.

To highlight the main features of Equation (10), and to simplify the analysis, we focus on the case in which quality is deterministic. Equation (10) becomes:

$$q-p^n-\sigma(p_t,p^n)\cdot(p_t-p^n),$$

where the key object is (the negative of) the decision value of the current price p_t based on a norm p^n:

$$DV(p_t,p^n)\equiv p^n+\sigma(p_t,p^n)\cdot(p_t-p^n). \tag{11}$$

Two key properties distinguish Equation (11) from Prospect Theory, the leading model of reference-dependent preferences. First, Equation (11) implies that *DV* is sometimes anchored to the norm p^n, and not always contrasted away from it as implied by Prospect Theory. To see this, consider the effect of changing the norm on *DV*:

$$\frac{\partial DV(p_t,p^n)}{\partial p^n}=1+\frac{\partial\sigma(p_t,p^n)}{\partial p^n}\cdot(p_t-p^n)-\sigma(p_t,p^n). \tag{12}$$

Because of anchoring, raising the norm increases *DV* (the first term). Because of adjustment, raising the norm makes p_t look surprisingly low, reducing *DV* (the second and third terms). In reference dependent

[15] Koszegi and Szeidl (2013) and Bushong, Rabin and Schwartzstein (2017) explore how attention is allocated as a response to choice data. They focus on the role of the range of attributes, and in particular on whether larger ranges attract or dampen attention. Endogenizing memory-based norms suggests both intuitions are valid: while extreme price realizations unambiguously attract attention (as in the water at the airport example), experiencing a greater variance of prices facilitates the decision-maker's adaptation to any price, and dampens overreaction.

[16] In the current formulation, DV in equation (11) is monotonic in price, so that a salient low price always increases valuation. In Bordalo, Gennaioli and Shleifer (2013), by contrast, price salience reduces valuation of all goods, but reduces it more for more expensive goods. This distinction is not critical for Salience Theory, but monotonicity is intuitive.

models, only the latter effect is present.[17] In our model, instead, anchoring dominates the standard adjustment effect when the current price is close to the norm $p_t \approx p^n$, while it subsides when the surprise $|p_t - p^n|$ is large. Psychologists call the first case *assimilation*, and the second case *contrast*. Our model, unlike Prospect Theory, captures both.

The second key difference with Prospect Theory is that in our model valuation is altered by irrelevant cues, such as the cued, and potentially irrelevant, context c_t. Because the effects of cues depend on past experiences, assimilation and contrast depend on the memory databases of different consumers. These features give rise to new testable predictions. First, contextual cues can be directly manipulated in laboratory or field experiments. But even in field data, there are proxies for cues that should not normatively affect valuation, such as current weather when buying a sweater. Second, the impact of contextual cues on behavior should depend on past experiences. It is increasingly common to see datasets reporting the conditions faced by consumers in the past (e.g. Simonsohn and Loewenstein 2006). We explore these implications below.

4.1 Memory-Based Decision Value

To study the interaction between memory and attention, consider the decision value of a given price p_t in context c_t. The price-context stimulus (p_t, c_t) cues retrieval of a norm $p^n(p_t, c_t)$ that leads to *DV*:

$$DV\left(p_t, p^n(p_t, c_t)\right) = p^n(p_t, c_t) + \sigma\left(p_t, p^n(p_t, c_t)\right)[p_t - p^n(p_t, c_t)]. \quad (13)$$

This expression generalizes the case in which only the price acts as a cue ($\rho = 0$) discussed in Proposition 1. We can characterize the *DV* function in Equation (13) as follows.

Proposition 3. *Suppose that the database $F(p, c)$ is Gaussian and similarity is measured by Equation (3). Suppose further that the salience function exhibits increasing concavity, namely $2\sigma'(x, 1) + \sigma''(x, 1)(x - 1)$ is monotonically decreasing for $x > 1$ and negative at $x = \lim_{p_t \to +\infty} \frac{p_t}{p^n(p_t, c_t)}$. Then for any context c_t such that*

[17] Formally, in KT, the valuation of p_t is $v(p_t - p^n)$, where $v(.)$ is an increasing gain-loss utility. In KR, the valuation of p_t is $p_t + v(p_t - p^n)$, where again $v(.)$ is increasing. In both cases, price valuation falls in the reference p^n.

$p^n(\overline{p}, c_t) > 0$, *there exist two price thresholds* p_L, p_H *satisfying* $0 < p_L < p^n(\overline{p}, c_t) < p_H$ *such that* $\{p_L, p^n(\overline{p}, c_t), p_H\}$ *are the inflection points of* DV, *and coincide with its crossing points.*

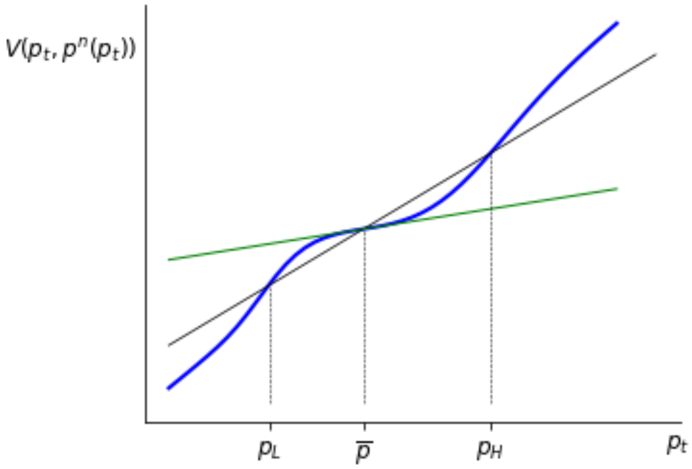

Figure III
Decision value of price.
The figure plots the decision value, $p^n + \sigma(p_t, p^n)(p_t - p^n)$ (in blue) and the norm p^n (in green) as a function of observed price p_t when the database is Gaussian and context is uncorrelated with price (Equation 6). The salience function is of the form $\sigma(x, 1) = \sigma \frac{e^{\theta(x-1)^2}}{1+e^{\theta(x-1)^2}} - \sigma_0$ for $x \geq 1$, where $\sigma > 1 + \sigma_0$, $\theta > 0$ and $\sigma(1,1) = 0$. The rational benchmark p_t is shown in black. In the figure, $\sigma = 3$, $\sigma_0 = 3/2$, and $\theta = 50$.

The blue curve plots the DV of price while the green line plots the price norm conditional on p_t. When the current price p_t is close to $p^n(\bar{p}, c_t)$, the norm is accurate. The small discrepancy between p_t and the norm does not attract attention, so DV is assimilated to the norm. Here DV is shifted toward the norm, the flat curve $p^n(p_t, c_t)$, so the consumer's price sensitivity is dampened relative to the rational case. The seasoned traveler seeing a \$4 bottle of water at the airport recalls similar airport prices and buys as in the past, paying little attention to small price differences from his past airport experiences.

For prices far enough from the average, the surprise relative to the retrieved norm is large and salient. The consumer pays a lot of attention to it and valuation is now contrasted away from the norm. The consumer's price sensitivity is steeper than the 45^o line, higher than in the rational case. The inexperienced traveler seeing the same \$4 bottle of water perceives it as exorbitantly expensive relative to his normal \$1 price. He focuses on the high price and refuses to buy. As prices get extremely high or low, diminishing sensitivity of salience prevails, so valuation flattens again.

The DV in Figure III, and more generally Equation (1), can be interpreted as reflecting the use of mental categories. The norm is the exemplar of the category of normal bottled water. If the current stimulus (q, p) falls within the normal range, it is assimilated to the category norm, as in models where category members are assimilated toward the category exemplar (Mullainathan 2000, Fryer and Jackson 2008). If instead the current stimulus is far from the normal category, it is placed in the "cheap" or "expensive" category, and its valuation is shaped by contrast relative to the norm. We can thus view the price thresholds p_L and p_H as determining the cheap, expensive, and normal price categories, so that price sensitivity is highest at category boundaries, consistent with the evidence on categorization (Goldstone 1994).[18]

The possibility of assimilation implies that DV in Figure III differs markedly from the value function in Prospect Theory, in which price sensitivity is highest around the reference price boosted by loss aversion for prices above the reference level. One way to reconcile the intuitions of Prospect Theory and our model is to note that in Prospect Theory gains and losses belong to sharply different categories, consistent with price sensitivity being the highest at the category boundary.

Our approach also differs from models of efficient coding (e.g. Wei and Stocker 2015, Woodford 2012, Polania, Woodford and Ruff 2019, Frydman and Jin 2018), in which sensitivity to a stimulus is highest around its modal level. In our model price sensitivity is lowest at the modal price. This is due to the fact that in our model attention is allocated ex-post, not ex-ante. At the modal price, the norm is accurate, price surprise $|p_t - p^n|$ is small, and so attention is disengaged, causing the price sensitivity to be low. When the observed price is sufficiently surprising, $|p_t - p^n|$ is large enough, attention is engaged and price sensitivity is high (and then monotonically declines due to diminishing sensitivity of salience).[19] This role of attention is consistent with evidence from psychology. When experimental subjects are primed or encouraged ex-post to compare stimuli, which is akin to enhancing the salience $\sigma(p_t, p^n)$ of given price differences, contrast is more likely to dominate (Cunha and Shulman 2011).

[18] Increasing concavity of the salience function ensures that the valuation function has exactly three inflection points. But the fact that price valuation increases in p_t as well as the ranges of assimilation and contrast identified by the thresholds p_L and p_H continue to hold if this condition is relaxed.
[19] Valuation is steepest around the modal price when prices are highly stable, i.e., π^2 is very small. In this special case, our model approximates the Prospect Theory value function with a norm equivalent to the "status quo" price.

Evidence from marketing is consistent with the predictions of our model. Individual consumers appear insensitive to small price changes relative to their reference price. Studies of aggregate demand across dozens of retail product categories and different retailers suggest that demand is rigid around normal prices while it is more elastic for larger price changes in either direction (e.g. Casado and Ferrer, 2013, Cheng and Monroe 2013). Our model offers a possible account of these findings. Small price changes within the "normal range" are not attended to, but large changes are, perhaps too much so. The contrast effect of Prospect Theory could not account for such insensitivity. Real rigidities stemming from rational (Sims 2003) or sparse (Gabaix 2014) inattention, or adjustment costs, may explain this evidence. However, these mechanisms do not account for over-reaction to large price changes.

In economics, several studies show that households do not react to changes in the price of their health plans, which are typically small (Chandra, Handel and Schwartzstein 2019). In retail markets, Nakamura and Steinsson (2008) find that firms often implement small price increases (against the logic of menu cost models), but also hold large temporary sales. Ortmeyer, Quelch and Salmon (1991) show that the majority of revenue of large department stores come from occasional deep discounts, which is difficult to account for within a price discrimination framework. These pricing policies may be optimal when, as in Figure III, households are inattentive to small deviations from the reference and over-react to large price drops.[20]

Our model yields several predictions for the intermediate range of assimilation.

Proposition 4. *Consider the intermediate price range* (p_L, p_H) *of Proposition 3 in which valuation is assimilated toward the norm. Our model yields the following comparative statics.*

1) The range (p_L, p_H) *moves up and expands for a consumer who has experienced higher prices, formally* $\frac{\partial p_L}{\partial \overline{p}} > 0, \frac{\partial p_H}{\partial \overline{p}} > 0,$ *and* $\frac{\partial (p_H - p_L)}{\partial \overline{p}} > 0$*, or after he is cued with a context* c_t *associated with higher prices.*

2) Higher volatility of prices π^2 *expands the range* (p_L, p_H)*, formally* $\frac{\partial p_L}{\partial \pi^2} < 0, \frac{\partial p_H}{\partial \pi^2} > 0.$

[20] In our memory based approach, recent prices can play the role of reference prices (so that attention is modulated by price changes) due to two reasons. First, because they are similar to current prices, so they are more easily retrieved. Second, due to recency effects (i.e. contextual similarity).

According to 1), neglect of price changes around the norm occurs to a greater extent at a higher experienced price level $\overline{p}$. This is due to the diminishing sensitivity of salience: an error of \$5 is less salient compared to a high price norm such as \$100 than to a low price norm of \$10. This result is reminiscent of Dehaene's (1997) evidence of Weber's law in number perception, where the price difference required for a given rate of discriminability increases proportionately with the price level. Several papers in marketing offer laboratory and field evidence of a similar phenomenon: the "latitude of acceptance" of a certain price grows with the price level (Koschate-Fischer and Wullner 2017). This idea may provide a psychological foundation for the well-documented finding that the price dispersion for a good increases proportionately with its average price (Pratt, Wise, and Zeckhauser 1979). Price dispersion may arise due to imperfect competition, as a function of search costs (Pratt, Wise, and Zeckhauser 1979). Yet both the level of price dispersion, and its increase with the price level, are hard to explain based on search costs alone, which are plausibly similar across a wide range of price levels. Proposition 4 offers a complementary mechanism: flat valuation around the normal price amplifies the effect of search costs, and the wider normal range at higher price levels accounts for the increase of dispersion with the price level.[21]

The most distinctive predictions of Proposition 4 are due to the associative structure of memory. First, a cue that retrieves contexts associated with low prices shrinks the price inattention range. When buying a handicraft at a market in a developing country, tourists may bargain hard over amounts that they would not even notice when shopping back home. Dissimilarity from a rich country market shuts down recall of the traveler's normal price level at home, in contrast with purely backward looking reference points. Once back home, the consumer again adapts to the usual prices. In contrast to rational expectations reference points, cues that remind the consumer about low prices in the developing country jolt him to see the usual prices at home as very high. Contextual similarity brings past price norms to mind, switching attention to prices on and off.

Second, selective memory implies that higher price volatility also expands the range of inattention in response to price cues. Exposure to a higher π^2 means the consumer has many instances of each price to draw from. Similarity-based recall then implies that, even when faced with a price away from the mode, he recalls

[21] Tversky and Kahneman (1981) also suggested that diminishing sensitivity interact with search costs to explain patterns of price discrimination. In their thought experiments, subjects stated a higher willingness to travel across town to save \$5 off a \$15 calculator than off a \$125 jacket, even if they were buying both items (see also Cunningham 2013).

past instances of it and views it as normal. As a result, a given price change is more likely to be assimilated to the norm when prices are very volatile.[22]

There is supportive evidence for these predictions. Niedrich, Sharma and Wedell (2001) show subjects price series drawn from different distributions, and then ask them to report the attractiveness of the product. They find that when prices are drawn from a bimodal distribution, subjects are less attracted by low prices and less disappointed by high prices relative to subjects trained with less volatile distributions. This is consistent with our model: extreme prices look more normal when drawn from a more volatile distribution. Once again, this role of the price distribution in shaping the flexibility of the reference is inconsistent with Prospect Theory, in which the price observed ex-post does not affect the reference point. Future work may seek a finer test of our model by trying to elicit judgments of price normality.

The same idea can explain why the efficacy of "strategic sales" is limited if these sales occur too often: consumers become adapted, and hence they are not surprised when they see them. Frequent shallow sales lower consumers' "internal reference price" much more dramatically than do infrequent deep sales (Cheng and Monroe 2013). This is consistent with our model: shallow sales are both more frequent and more similar to regular prices, and thus may entail assimilation rather than contrast.

In sum, our anchoring and adjustment mechanism unifies assimilation and contrast effects in a way that cannot be obtained under existing reference dependent theories such as Prospect Theory. The memory structure of norms yields testable predictions on when assimilation and contrast should prevail. In particular, irrelevant cues about hedonic attributes or context can generate artificial contrast, as we now show in detail.

4.2 Background Contrast Effects

A conventional analysis of the effect of past experience on choice is related to the phenomenon of background contrast. In background contrast experiments (e.g. Simonson and Tversky 1992) subjects are initially presented with goods at either high or low prices. In a second stage, subjects choose whether to buy

[22] Under rational inattention, greater ex-ante price variability should increase attention to prices. In our model there are two conflicting effects. On the one hand, a higher π^2 reduces prices sensitivity by fostering more flexible norms. On the other hand, a higher π^2 increases price sensitivity by making it more likely that a surprising price is realized.

similar goods at an intermediate price. Exposure to high prices increases the likelihood of purchase in the second stage, while exposure to low prices decreases it. Such contrast effects have been documented in a number of other settings, such as housing choices by movers (Simonsohn and Loewenstein 2006), reaction to earnings announcements (Hartzmark and Shue 2018), dating markets (Bhargava and Fisman 2014), and context dependent willingness to pay (Thaler 1985, Mazar, Koszegi and Ariely 2014). And yet our choices in many situations are stable, pointing to boundaries for such contrast effects. Our model sheds light on the drivers of background contrast effect and yields new predictions on when this effect should subside.

To map our analysis to existing experimental and field evidence, we hold fixed the price p_t observed by the consumer in the second stage and vary either the database $F(p, c)$, or the contextual cue c_t. This corresponds to varying only the price norm $p^n(p_t, c_t)$ in the DV Equation (11). Proposition 3 implies the following result.

Corollary 3. *Suppose that the consumer faces a price* p_t *below the average price norm* $p^n(\overline{p}, c_t)$*. Then* $DV\big(p_t, p^n(p_t, c_t)\big) < p_t$ *if and only if:*

$$p_t < \underline{p} \equiv \alpha p^n(\overline{p}, c_t),$$

where $\alpha < 1$ *is a constant that decreases in prior price variability* π^2*. The case of high prices,* $p_t > p^n(\overline{p}, c_t)$*, is symmetric with* $\alpha > 1$ *that increases in* π^2*.*

The background contrast effect arises if p_t is sufficiently different from the price norm $p^n(\overline{p}, c_t)$ prevailing in context c_t. Unlike in reference dependent models such as Prospect Theory, which always entail contrast relative to the reference, in our model contrast arises only if i) the current price p_t is very different from past experiences $\overline{p}$, and crucially if ii) context c_t induces selective recall of prices that are different from the current p_t, creating an artificial surprise. Cueing airport prices raises the norm of a shopper who is downtown, causing him to appreciate the current price of water.

Contextual cueing helps understand the experimental tests of the background contrast effect. These experiments typically consider familiar goods, so why would the first-stage prices affect the consumer's norm in the second stage? Our model suggests that first stage prices are selectively retrieved due to contextual similarity to the second stage, driven both by the stability of the laboratory environment and by recency. Thus, first stage prices interfere with recall of the broader database $\overline{p}$, and create artificial surprise. The structure of memory offers an additional prediction: if the goods are familiar and have been experienced with high price volatility π^2, the contrast effect should be harder to obtain, since in this case the norm would be closer to the current price p_t.

The idea that background contrast arises only if the current stimulus is far from past experiences is echoed in the large psychology literature on contrast versus assimilation effects (for a review, see Stapel and Suls 2011). When judging stimuli such as length, size, or loudness, assimilation of the current stimulus to a cued reference or past experience prevails when the discrepancy is moderate while contrast occurs when the discrepancy is large (e.g. Herr, Sherman and Fazio 1983).[23] Similar findings arise in more abstract judgments, such as anchoring and priming experiments. A fox is judged to be more aggressive when subjects are cued with a wolf, and less aggressive when subjects are cued with a tiger, than when judged in isolation (Strack, Bahnik, Mussweiler 2016). Priming an extremely hostile person (e.g. Hitler) causes subjects to rate a target person less hostile, while priming a moderately hostile person (e.g. Joe Frazier) generates assimilation (Herr 1989). There is also some evidence on prices. Consumers primed with a moderately high car price judge the price of an unknown car brand as more expensive, creating assimilation (Herr 1989) while extreme prices promote contrast.[24]

To map more precisely our model to some puzzles, and highlight its new predictions, it is useful to consider how prior experience and contextual cues affect the consumer's willingness to pay p_{WTP} for a good of quality q. This is the highest price at which the decision value in Equation (11) is not greater than quality q. Because the DV of price increases in p_t, p_{WTP} is implicitly defined by:

---

[23] The evidence also shows that large departures from the adaptation level lead to overshooting (or what psychologists call "after-effects"): a given level of a stimulus, such as brightness or temperature, is underestimated after exposure to high levels of that stimulus (Brigell and Uhlarick 1979). In our model, overshooting comes from strong attention to surprise, $\sigma(p_t, p^n) > 1$. The existence of contrast effects, defined as the case in which the valuation of an attribute is contrasted away from the norm akin to a negative derivative in Equation (12), does not require overshooting. It only requires that $\sigma(x, x^n)$ be steep enough with respect to the norm.

[24] These effects are closely related to the role of categorization in perception, which suggests that perception emphasizes differences in stimuli from different categories, while dampening differences in stimuli from the same category. Such effects arise in learned categories about musical pitches, brightness, size of anodine objects, etc. (Goldstone 1994). Experiments suggest that categories may arise spontaneously along the presented stimuli, such as size or loudness. In our model, categorization is triggered by surprise. As in Figure 3, there is a normal price range when surprise is small, and two abnormal price ranges (high and low) when surprise is large. The latter drive contrast effects.

$$p^n(p_{WTP}, c_t) + \sigma\big(p_{WTP}, p^n(p_{WTP}, c_t)\big) \cdot \big(p_{WTP} - p^n(p_{WTP}, c_t)\big) = q$$

Note that the norm used for valuation is shaped by two cues: context and p_{WTP} itself. Willingness to pay is a function of context and of the price database. We denote this function by $p_{WTP}(\overline{p}, c_t)$.

Consider the willingness to pay for a good of quality q in context c of two consumers with different price histories. One consumer has on average experienced a high price context c_h, associated with a higher average price $\overline{p}_h$ in the database. Another consumer has instead on average experienced a low price context c_l, so the average price in his database is low, $\overline{p}_l < \overline{p}_h$. The two consumers have identical preferences, so in a rational world they would have the same willingness to pay q. This is not necessarily so when memory and attention affect valuation. The WTP of the two consumers can be characterized as follows.

Proposition 5. *Suppose the two consumers evaluate the good in the same context $c \in (c_l, c_h)$. Then, under the conditions of Proposition 3, there are two thresholds $p_l(c, c_l)$ and $p_h(c, c_h)$, such that $p_{WTP}(\overline{p}_h, c) > q > p_{WTP}(\overline{p}_l, c)$ if $\overline{p}_h > p_h(c, c_h)$ and $\overline{p}_l < p_l(c, c_l)$. These conditions are more likely to hold if context is stable ($|c - c_i|$ is low for $i = l, h$) or if price variability π^2 experienced by the two consumers is low.*

As in Corollary 3, artificial surprise arises when the consumer has seen extreme prices in the past and current context is sufficiently close to these past experiences that these extreme prices are retrieved. The key new result here is that, in these cases, willingness to pay over-reacts, moving toward extreme past experiences: consumer h reports a higher willingness to pay than l despite the fact that they have identical objective valuations of the good. We next discuss some evidence bearing on this mechanism.

Housing Choices of Movers. Simonsohn and Loewenstein (2006) show that the home rental decisions of movers to a new city are influenced, controlling for household characteristics, by the price of housing in the origin city. When moving to a cheaper city, say from San Francisco to Pittsburgh, consumers spend more on housing than comparable locals, and the reverse holds when moving to a more expensive city, say from Atlanta to New York. In subsequent renting decisions the movers converge to locals: they switch to cheaper rents in Pittsburgh and to more expensive ones in New York. This is puzzling for the neoclassical model, including one with search costs, but also for models of rational expectations reference points since such reference should adjust immediately, as the mover acquires information about the destination city.

Instead, the evidence is consistent with the logic of Proposition 5. The “house hunt” database of a mover from San Francisco to Pittsburgh reflects the high San Francisco rents $\overline{p}_h$. Of course, the Pittsburgh context c is not identical to the San Francisco context c_h, but – with a database mostly shaped by San Francisco rents and by very few experiences of house hunting in Pittsburgh (i.e., π^2 is low) – the search for an apartment in Pittsburgh cues recall of the apartment rents in San Francisco. Recalling these high rents makes current rents look surprisingly cheap, causing the mover to have an inflated willingness to pay for housing.[25] Context explains the subsequent adaptation. As the renter’s database becomes populated by Pittsburgh rents, the Pittsburgh context c retrieves local rents and interferes with recall of San Francisco rents (because $c \neq c_h$ and now π^2 is high), so the rent norm of a second time mover drops to the norm of locals.

Besides accounting for the evidence, our model yields new predictions. If the San Francisco mover at some point lived in a city with similar characteristics and rents as Pittsburgh, say Atlanta, the current context c would cue memories of that experience. These memories displace the more recent San Francisco rents. Bordalo, Gennaioli, and Shleifer (2019) replicate Loewenstein and Simonsohn (2006)’s evidence using 20 additional years of data, and confirm this prediction: movers who have past experiences with rents close to the destination city are less influenced by city of origin rents.

Beer on the Beach. Context-driven artificial surprise also underlies Thaler’s (1985) famous beer-at-the-beach experiment. Subjects state a higher willingness to pay for a given beer, to be consumed on the beach, when the beer is described as being bought from a nearby resort rather than from a nearby run-down shack.[26] Unlike in Proposition 5, here the two consumers may have the same underlying database of past experiences, but one is cued by a high price context c_h, “resort”, while another is primed by a low price context c_l, “shack”. The first consumer is cued to recall high prices, while the second recalls low prices. Provided the normal resort price is sufficiently high, the resort prime induces a higher willingness to pay than the shack prime,

[25] The same reasoning holds true in the quality space. The mover coming from San Francisco is accustomed to live in small apartment. When seeing the large apartments in Pittsburgh he is positively surprised and over-reacts, becoming more willing to pay a higher price for the same quality relative to a Pittsburgh resident.

[26] In related experimental evidence, consumers who are primed with “Walmart” spend less on a subsequent choice than consumers primed with “Nordstrom” (Chartrand et al 2008). Relative to Thaler (1985), this design removes a confound in that the purchase decision is exactly the same.

$p_{WTP}(\overline{p}, c_h) > p_{WTP}(\overline{p}, c_l)$. Cued prices are extreme, so the consumer perceives even moderately high prices as low at the resort and moderately low prices as too expensive at the shack.

Here choice distortions arise because the consumer focuses on superficial similarity between the resort prime and high prices, neglecting the fundamental factor that the beer is consumed on the same beach. This is a perverse case, but in many other circumstances the very same mechanism helps us form accurate norms and make better decisions. We may adapt to paying higher prices at resorts than at corner stores because, under more typical circumstances, beer purchased at a resort would be consumed there, enabling a consumer to enjoy the comfort and service that are not available at the corner store. In this and other cases, contextual similarity fosters benign adaptation, improving decisions.

Conflicting primes in the Laboratory. The role of memory, and in particular of contextual similarity, in driving the background contrast effect is further illustrated by Mazar, Koszegi and Ariely (2014). In incentivized experiments, they reproduce the gap in WTP when in a first stage subjects are exposed to high versus low price contexts c_h and c_l. They find that the gap is eliminated if in the second stage subjects are encouraged to think about a "reasonable" price. In this case, the first stage experience is less influential due to the presence of an additional "reasonableness" cue. They also find that the gap is eliminated if in the second stage subjects are initially exposed to both contexts, knowing that only one context would materialize. Reminding subjects of both contexts reduces interference, and allows both high and low prices to be recalled. As the authors argue, this evidence is inconsistent with rational expectations reference points, which only depend on the context that actually materializes, and not on the other one.

Broken Dishes. We conclude with a puzzle entailing quality rather than price assessments: Hsee's (1996) broken dishes experiment. Subjects reported a higher willingness to pay for a set of 10 intact dishes than for a set of 13 dishes of which 11 are intact and two are broken. This behavior violates monotonicity. Our model accounts for this choice by viewing the number of dishes as a contextual variable that retrieves an immediate quality norm from memory. In this norm, no dish is broken because we never experience buying broken dishes. The consumer is negatively surprised. His valuation overshoots, and falls below the rational valuation of a set of 10 intact dishes, for which there is no negative surprise.[27]

5. Conclusion

We present a new theory of choice based on the idea that valuation is a two-stage process: an automatic estimation of value based on cued recall, followed by an adjustment in the direction of any discrepancy between the estimated and observed attributes. We use a biologically founded, textbook model of memory (Kahana 2012) to build a model of memory-based norms, and combine it with the salience theory of choice, which is a natural way to incorporate the notions of surprise, and of reaction to surprise, that are featured in Kahneman and Miller's Norm Theory.

The broad principle emerging from our analysis is that similarity-based recall tends to retrieve norms that are well adapted to current conditions, thus creating stability of choice, unless normatively irrelevant cues bring to mind different past experiences. The attribution bias, the projection bias, shrouded attributes, and contrast effects, can all be viewed as specific manifestations of this general process. The structure of selective recall yields new predictions as to when these effects should prevail and how they depend on contextual cues.

Future work may use our results to shed light on several issues. One natural application is judgments of fairness, which have been shown to influence economic decisions in the field and the lab (see Fehr and Schmidt 1999). These studies often rely on two ingredients: a "fair" allocation, and "social preferences" that make departures from that allocation personally costly. In a paper published in the same year as Norm Theory, Kahneman, Knetsch and Thaler (KKT 1986) equate the fair allocation with a so-called "reference transaction", which "provides a basis for fairness judgments because it is normal, not necessarily because it is just." This suggests that a fair allocation is like a norm. It is retrieved from memory, it largely reflects custom, it is malleable, and it adapts to change. As KKT put it, "terms of exchange that are initially seen as unfair may in

[27] When evaluating the set with 13 dishes, $c = 13$, subjects immediately retrieve from memory the quality q_{13} of 13 intact dishes. With two broken dishes, however, the quality of the set is only equal to the value q_{11} of 11 intact dishes and valuation is given by $q_{13} - \sigma(q_{11}, q_{13}) \cdot (q_{13} - q_{11})$, which drops below q_{10} only if $\sigma(q_{11}, q_{13}) > 1$.

time acquire the status of a reference transaction (...) at least in the sense that alternatives to it no longer come to mind."

Selective memory can also prove useful for thinking about expectations. A number of recent papers build on Kahneman and Tversky's (1972) representativeness heuristic to explore how selective memory shapes beliefs. In this approach, recall is selected toward features that are most diagnostic of, or similar to, a group in contrast to a comparison group (Gennaioli and Shleifer 2010, Bordalo et al. 2016), with important implications for expectations in social, financial, and macroeconomic domains. Bordalo et al. (2019) offer a similarity based memory model that generates the representativeness heuristic, and test the predictions of that model in the lab by measuring both recall and probabilistic judgments. More can be done to connect memory and beliefs, particularly in terms of assessing whether selective memory can also account for biases associated with other judgment heuristics such as availability and anchoring.

An extension of our model to the choice between two or more goods may be useful. Besides being help with applications, such an extension may provide insight into narrow framing and mental accounting (Thaler 1985), because available cues may enhance or inhibit the consumer's thinking about alternative goods or uses of funds. The choice of the grade of gas to fill the gas tank retrieves past instances of the same choice and past gas prices, but is unlikely to retrieve the consumer's other choices, leading to a neglect of opportunity cost (Frederick et al 2009, Shah, Shafir and Mullainathan 2015). Nor are gas prices likely to bring to mind the consumer's broader financial situation including his current income, leading to a breakdown of fungibility of money (Hastings and Shapiro 2013). If instead income is directly associated to a consumption choice, which happens with targeted vouchers (Abeler and Marklein 2016) or SNAP benefits (Hastings and Shapiro 2018), then the same memory-based mechanism may entail excess sensitivity of consumption to the windfall. Viewing the choice setting as a cue may explain how choice is both anchored to similar past decisions but at the same time isolated from other decisions, providing a foundation for narrow framing.

Our analysis indicates that, to evaluate our mechanism, choice or valuation data should be ideally paired with recall data and similarity judgments. Research on memory provides some guidance for eliciting similarity judgments from subjects and using them to predict recall. For example, Pantelis et al (2008) elicit similarity judgments between synthetic faces that vary in measurable attributes, and use this data to fit the similarity function in attribute space. Their subjects are more likely to confuse names assigned to faces that have higher similarity (either stated or predicted), in line with a model of noisy, similarity based recall.[28] Future work may build on these methods to predict both recall of past experiences and economic behavior in complex, multi-attribute situations.

While many questions remain open, it seems clear that textbook models of memory offer an opportunity to unify many behavioral models and to improve their empirical testability, and at a deeper level to understand how decision makers represent and make choices.

University of Oxford
Bocconi University and IGIER
Harvard University

[28] This as well as other work shows that similarity judgment are importantly shaped by widely shared deep seated intuition, so that similarity measurement can be used out of sample. Out of sample here means two things. First, similarity judgments of a given subject pool on a certain set of objects can be used to predict the similarity judgments of the same subject pool on a new set of objects. Second, similarity judgments elicited from subject pool A can proxy for those of subject pool B, so the judgments of A can be used to shed light on the choices of B. The latter procedure can enrich existing datasets that do not contain measures of similarity.

References

Abeler, Johannes and Felix Marklein, “Fungibiligy, Labels and Consumption,” Journal of the European Economic Association, 15 (2016), 99 – 127.

Anderson, John and Lynne Reder, “The Fan Effect: New Results and New Theories,” Journal of Experimental Psychology, 128 (1999), 186 – 197.

Azeredo da Silveira, Rava and Michael Woodford, “Noisy Memory and Over-Reaction to News”, NBER working paper w25478, 2019.

Bhargava, Saurabh, and Ray Fisman, "Contrast Effects in Sequential Decisions: Evidence from Speed Dating," Review of Economics and Statistics, 96 (2014), 444-457.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer, “Stereotypes," Quarterly Journal of Economics, 131 (2016), 1753 -- 1794.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, “Salience Theory of Choice under Risk," Quarterly Journal of Economics, 127 (2012), 1243 -- 1285.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, “Salience and Consumer Choice," Journal of Political Economy, 121 (2013), 803 -- 843.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, Andrei Shleifer, Frederik Schwerter, “Memory and Representativeness,” NBER working paper w25692, 2019.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, “Memory and Reference Prices: an Application to Rental Choice,” American Economic Review Papers and Proceedings, 109 (2019), 572–576.

Bornstein, Aaron and Kenneth Norman, “Reinstated episodic context guides sampling-based decisions for reward,” Nature Neuroscience, 20 (2017), 997–1003.

Bushong, Benjamin and Tristan Gagnon-Bartsch, “Learning with Misattribution of Reference Dependence,” Unpublished Working Paper, 2019.

Bushong, Benjamin, Matthew Rabin and Joshua Schwartzstein, “A Model of Relative Thinking,” Unpublished Working Paper, 2017.

Busse, Meghan, Devin Pope, Jaren Pope and Jorge Silva-Risso, “The Psychological Effect of Weather on Car Purchases,” Quarterly Journal of Economics 130 (2015), 371 – 414.

Cadena, Ximena and Antionette Schoar, “Remembering to Pay? Reminders vs. Financial Incentives for Loan Payments.” Unpublished Working Paper, 2011.

Calzolari, Giacomo and Mattia Nardotto, “Effective Reminders,” Management Science 63 (2016), 2915 – 2932.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia, "Inflation Expectations, Learning, and Supermarket Prices: Evidence from Survey Experiments," American Economic Journal: Macroeconomics, 9 (2017), 1-35.

Casado, Esteban and Juan-Carlos Ferrer, “Consumer Price Sensitivity in the Retail Industry: Latitude of Acceptance with Heterogeneous Demand,” European Journal of Operational Research, 228 (2013), 418 – 426.

Chandra, Amitabh, Benjamin Handel, and Joshua Schwartzstein, “Behavioral Economics and Health-Care Markets,” In Handbook of Behavioral Economics: Foundations and Applications 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson. Amsterdam: Elsevier/North-Holland, 2019.

Chang, Tom, Wei Huang and Yongxiang Wang, “Something in the Air: Pollution and the Demand for Health Insurance,” Review of Economic Studies, 85 (2018), 1609–1634.

Chartrand, Tanya, Joel Huber, Baba Shiv, and Robin Tanner, “Nonconscious Goals and Consumer Choice,” Journal of Consumer Research, 35 (2008), 189–201.

Chatterjee, Promothesh and Randall Rose, “Do Payment Mechanisms Change the Way Consumers Perceive Products?” Journal of Consumer Research, 38 (2011), 1129–1139.

Cheng, Lilian and Ken Monroe, “An Appraisal of Behavioral Price Research (Part 1): Price as a Physical Stimulus,” Academy of Marketing Science Review, 3 (2013), 103 – 129.

Chetty, Raj, Adam Looney, and Kory Kroft, “Salience and Taxation: Theory and Evidence,” American Economic Review, 99 (2009), 1145 – 1177.

Clark, Andy, ”Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science,” Behavioral and Brain Sciences, 36 (2013), 181 – 204.

Conlin, Michael, Ted O’Donoghue, and Timothy Vogelsang, “Projection Bias in Catalogue Orders,” American Economic Review, 97 (2007), 1217 – 1249.

Cunningham, Tom, "Comparisons and Choice," Unpublished Working Paper, 2013.

Cunha, Marcus and Jeffrey Shulman, “Assimilation and Contrast in Price Evaluations,” Journal of Consumer Research, 37 (2011), 822 – 835.

Dehaene, Stanislas. The Number Sense. Oxford University Press, New York NY, 1997.

DellaVigna, Stefano, Attila Lindner, Balazs Reizer, and Johannes Schmieder, “Reference-Dependent Job Search: Evidence from Hungary," Quarterly Journal of Economics, 132 (2017): 1969 – 2018.

Dickson, Peter and Alan Sawyer, “The Price Knowledge and Search of Supermarket Shoppers,” Journal of Marketing, 54 (1990), 42—53.

Enke, Benjamin, Frederik Schwerter, and Florian Zimmermann, "Associative Memory and Belief Formation," Unpublished Working Paper, 2019.

Ericson, Keith, “On the Interaction of Memory and Procrastination: Implications for Reminders,Deadlines and Empirical Estimation,” Journal of the European Economic Association, 15 (2017), 692-719.

Fehr, Ernst and Klaus Schmidt, “A Theory of Fairness, Competition, and Cooperation,” Quarterly Journal of Economics, 114 (1999), 817 – 868.

Finkelstein, Amy, "E-ztax: Tax Salience and Tax Rates," Quarterly Journal of Economics, 124 (2009), 969-1010.

Frederick, Shane, Nathan Novemsky, Jim Wang, Ravi Dhar, Stephen Nowlis, “Opportunity Cost Neglect” Journal of Consumer Research, 36 (2009), 553 – 561.

Frydman, Cary and Lawrence Jin, “Efficient Coding and Risky Choice,” Unpublished Working Paper, 2018.

Fryer, Roland and Matthew Jackson, “A Categorical Model of Cognition and Biased Decision Making,” The B.E. Journal of Theoretical Economics, 8 (2008): article 6.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," Quarterly Journal of Economics, 129 (2014), 1661-1710.

Gabaix, Xavier and David Laibson, “Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets,” Quarterly Journal of Economics, 121 (2004), 505–540.

Gennaioli, Nicola and Andrei Shleifer, “What Comes to Mind," Quarterly Journal of Economics, 125 (2010), 1399 -- 1433.

Georganas, Sotiris, Paul Healy, and Nan Li, "Frequency Bias in Consumers′ Perceptions of Inflation: An Experimental Study," European Economic Review, 67 (2014), 144-158.

Gallagher, Justin, "Learning About an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States," American Economic Journal: Applied Economics, 6 (2014), 206-233.

Gilboa, Itzhak and David Schmeidler, “Case-Based Decision Theory," Quarterly Journal of Economics, 110 (1995), 605 -- 639.

Goldstone, Robert, "The Role of Similarity in Categorization: Providing a Groundwork," Cognition, 52 (1994), 125-157.

Grubb, Michael, and Matthew Osborne, "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock," American Economic Review, 105 (2015), 234-71.

Haggag, Kareem, Devin Pope, Kinsey Bryant-Lees, Maarten Bos, “Attribution Bias in Consumer Choice,” Review of Economic Studies, 86 (2018), 2136-2183.

Hartzmark, Samuel, and Kelly Shue, "A Tough Act to Follow: Contrast Effects in Financial Markets," Journal of Finance, 73 (2018), 1567-1613.

Hastings, Justine and Jesse Shapiro, “Fungibility and Consumer Choice: Evidence from Commodity Price Shocks,” Quarterly Journal of Economics, 128 (2013), 1449 – 1498.

Hastings, Justine and Jesse Shapiro, “How are SNAP Benefits Spent? Evidence from a Retail Panel,” American Economic Review, 108 (2018), 3493 – 3540.

Hawkins, Jeff and Sandra Blakeslee. On Intelligence. Times Books, New York NY, 2004.

Herr, Paul, “Priming Price: Prior Knowledge and Context Effects,” Journal of Consumer Research, 16 (1989), 67–75.

Herr, Paul, Steven Sherman, and Russell Fazio, “On the Consequences of Priming: Assimilation and Contrast Effects,” Journal of Experimental Social Psychology, 19 (1983), 323-340.

Hsee, Christopher, “The Evaluability Hypothesis: An Explanation for Preference Reversals between Joint and Separate Evaluations of Alternatives,” Organizational Behavior and Human Decision Processes, 67 (1996), 247-257.

Kahana, Michael. Foundation of Human Memory. Oxford University Press, Oxford UK, 2012.

Kahneman, Daniel, Barbara Fredrickson, Charles Schreiber, and Donald Redelmeier, "When More Pain Is Preferred to Less: Adding a Better End." Psychological Science, 4 (1993), 401 – 405.

Kahneman, Daniel, Jack Knetsch and Richard Thaler, “Fairness as a Constraint on Profit Seeking: Entitlements in the Market,” American Economic Review, 76 (1986), 728-741.

Kahneman, Daniel and Dale Miller, “Norm Theory: Comparing Reality to its Alternatives,” Psychological Review, 93 (1986), 136-153.

Kahneman, Daniel and Amos Tversky, “Subjective Probability: A Judgment of Representativeness,” Cognitive Psychology, 3 (1972), 430–454.

Kahneman, Daniel and Amos Tversky, “Prospect Theory: an Analysis of Decision under Risk," Econometrica, 47 (1979), 263 – 292.

Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman, “Getting to the Top of Mind: How Reminders Increase Saving.” Management Science, 62 (2016), 3393–3411.

Koschate-Fischer, Nicole and Katharina Wullner, “New Developments in Behavioral Pricing Research,” Journal of Business Economics, 87 (2017), 809–875.

Koszegi, Botond and Matthew Rabin, “A Model of Reference-Dependent Preferences," Quarterly Journal of Economics, 121 (2006), 1133 -- 1165.

Koszegi, Botond and Adam Szeidl, "A Theory of Focusing in Economic Choice," Quarterly Journal of Economics, 128 (2012), 53-104.

Kuchler, Theresa, and Basit Zafar, "Personal Experiences and Expectations about Aggregate Outcomes," Journal of Finance, 74 (2019), 2491 – 2542.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," Quarterly Journal of Economics, 118 (2003), 1209–1248.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?" Quarterly Journal of Economics, 126 (2011), 373 -- 416.

Malmendier, Ulrike and Stefan Nagel, "Learning from Inflation Experiences," Quarterly Journal of Economics, 131 (2016), 53-87.

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

Mullainathan, Sendhil, "Thinking through Categories," Unpublished Working Paper, 2000.

Mullainathan, Sendhil, "Memory-Based Model of Bounded Rationality," Quarterly Journal of Economics, 117 (2002), 735-774.

Mussweiler, Thomas and Fritz Strack, "Considering the Impossible: Explaining the Effects of Implausible Anchors," Social Cognition, 19 (2001), 145 – 160.

Nagel, Stefan and Zhengyang Xu, "Asset Pricing with Fading Memory," Unpublished Working Paper, 2018.

Nakamura Emi and Jon Steinsson, "Five Facts about Prices: A Reevaluation of Menu Cost Models," Quarterly Journal of Economics, 123 (2008), 1415–1464.

Ortmeyer, Gwen, John Quelch, and Walter Salmon, "Restoring Credibility to Retail Pricing." Sloan Management Review, 33 (1991), 55–66.

Niedrich, Ronald, Subhash Sharma and Douglas Wedell, "Reference Price and Price Perceptions: A Comparison of Alternative Models," Journal of Consumer Research, 28 (2001), 339-354.

Pantelis, Peter, Marieke Van Vugt, Robert Sekuler, Hugh Wilson, and Michael Kahana, "Why Are Some People's Names Easier to Learn Than Others? The Effects of Face Similarity on Memory for Face-Name Associations," Memory & Cognition, 36 (2008), 1182-1195.

Pratt, John, David Wise and Richard Zeckhauser, "Price Differences in Almost Competitive Markets," Quarterly Journal of Economics, 93 (1979), 189–211.

Schacter, Daniel, Donna Addis, and Randy Buckner, "Remembering the Past to Imagine the Future: the Prospective Brain," Nature Reviews Neuroscience, 8 (2007), 657 -- 661.

Schwartzstein, Joshua, "Selective Attention and Learning," Journal of the European Economic Association, 12 (2014), 1423–1452.

Shah, Anuj, Eldar Shafir, and Sendhil Mullainathan, "Scarcity Frames Value." Psychological Science, 26 (2015), 402 – 412.

Shapiro, Jesse, "A 'Memory-Jamming' Theory of Advertising," Unpublished Working Paper, 2006.

Simonsohn, Uri and George Loewenstein, "Mistake #37: The Effect of Previously Encountered Prices on Current Housing Demand," Economic Journal, 508 (2006), 175-199.

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

Sims, Christopher, "Implications of Rational Inattention," Journal of Monetary Economics, 50 (2003), 665–690.

Slovic, Paul, Howard Kunreuther, and Gilbert White, “Decision Processes, Rationality, and Adjustment to Natural Hazards: a Review of Some Hypotheses,” in Gilbert White, editor, Natural Hazards, Local, National and Global, Oxford University Press, 1974.

Soman, Dilip and John Gourville, “Transaction Decoupling: How Price Bundling Affects the Decision to Consumer,” Journal of Marketing Research, 38 (2001), 30 – 44.

Stapel, Diederik and Jerry Suls, Assimilation and Contrast in Social Psychology. Psychology Press, New York NY, 2011.

Strack, Fritz, Stepan Bahnik, Thomas Mussweiler, “Anchoring: Accessibility as a Cause of Judgmental Assimilation,” Current Opinion in Psychology, 12 (2016), 67–70.

Taubinsky, Dmitry, “From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice," Unpublished Working Paper, 2014.

Thaler, Richard, “Mental Accounting and Consumer Choice,” Marketing Science, 4 (1985), 199 – 214.

Torgerson, Warren, Theory and Methods of Scaling. John Wiley & Sons, Inc, New York NY, 1958.

Tversky, Amos and Daniel Kahneman, “Availability: A Heuristic for Judging Frequency and Probability," Cognitive Psychology, 5 (1973), 207-232.

Tversky, Amos, and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," Science, 211 (1981), 453-458.

Tversky, Amos, “Features of Similarity,” Psychological Review, 84 (1977), 327-352.

Wachter, Jessica, and Michael Kahana, "A Retrieved-Context Theory of Financial Decisions," Working paper, 2019.

Woodford, Michael, “Prospect Theory as Efficient Perception Distortion,” American Economic Review, 102 (2012), 41 – 46.

Online Appendix for "Memory, Attention, and Choice"

Bordalo, Gennaioli, Shleifer

A. Proofs

Proposition 1. After observing a cue c_t, the norm retrieved from memory is equal to:

$$p^n(c_t) = \iint p \frac{S(|c - c_t|)}{\int S(|c - c_t|)dF(c)} dF(c,p)$$

$$= \int p \left[\int \frac{S(|c - c_t|)}{\int S(|c - c_t|)dF(c)} dF(c|p) \right] dF(p)$$

If context and price are uncorrelated, $dF(c|p) = dF(c)$, so the term in square brackets equals 1 and $p^n(c_t) = \overline{p}$. If the cue includes context and price, then the norm is equal to:

$$p^n(p_t, c_t) = \int p \left\{ \frac{S(|p - p_t|) \int S(|c - c_t|)\, dF(c|p)}{\int S(|p - p_t|)[\int S(|c - c_t|)dF(c|p)]dF(p)} \right\} dF(p)$$

If context and price are uncorrelated, this can be rewritten as:

$$p^n(p_t, c_t) = \int p \left\{ \frac{S(|p - p_t|) \int S(|c - c_t|)\, dF(c)}{\int S(|p - p_t|)dF(p) \int S(|c - c_t|)dF(c)} \right\} dF(p)$$

$$= \int p \left\{ \frac{S(|p - p_t|)}{\int S(|p - p_t|)dF(p)} \right\} dF(p) = p^n(p_t).$$

Finally, consider the case where $F(p)$ is symmetric and unimodal. We show this implies that $p^n(p_t)$ lies between p_t and $\overline{p}$. To do so, consider the claim that $p^n(p_t) \leq p_t$. This condition can be written as:

$$p_t - p^n(p_t) = \int_{-\infty}^{+\infty} (p_t - p)\tilde{S}(|p - p_t|; p_t)dF(p)$$

where $\tilde{S}(|p - p_t|; p_t)$ is a shorthand for $\frac{S(|p-p_t|)}{\int S(|p-p_t|)dF(p)}$. Thus, $p_t - p^n(p_t)$ becomes:

$$\int_{-\infty}^{p_t} (p_t - p)\tilde{S}(|p - p_t|; p_t)dF(p) + \int_{p_t}^{+\infty} (p_t - p)\tilde{S}(|p - p_t|; p_t)dF(p)$$

$$= \int_0^{+\infty} u\tilde{S}(u; p_t)dF(p_t - u) - \int_0^{+\infty} u\tilde{S}(u; p_t)dF(p_t + u)$$

$$= \int_0^{+\infty} u\tilde{S}(u; p_t)[dF(p_t - u) - dF(p_t + u)] \geq 0. \qquad (A.1)$$

Because $dF(p)$ is symmetric and unimodal around $\overline{p}$, Equation (A.1) holds if $|p_t - u - \overline{p}| \leq |p_t + u - \overline{p}|$ for any $u \geq 0$. This is true if and only if $p_t \geq \overline{p}$. A similar argument shows that $p^n(p_t) \geq \overline{p}$ if and only if $p_t \geq \overline{p}$, and in particular $p^n(p_t) = \overline{p}$ if and only if $p_t = \overline{p}$.

Corollary 1. With multiplicative similarity and zero correlation between price and context, the price norm $p^n(p_t, c_t)$ is the average price under the distorted measure $\frac{S(|p-p_t|)}{\int S(|p-p_t|)dF(p)} dF(p)$. When similarity is given by Equation (3) and the price distribution is normal with variance π^2, the distorted density is proportional to $e^{-\delta(p_t-p)^2} e^{-\frac{(p-\overline{p})^2}{2\pi^2}}$. This is a normal distribution with mean $\frac{\overline{p}+2\delta\pi^2 p_t}{1+2\delta\pi^2}$ and variance $\frac{\pi^2}{1+2\delta\pi^2}$.

Proposition 2. When the cue is context alone, the price norm $p^n(c_t)$ is the average price under the distorted measure $\left[\int \frac{S(|c-c_t|)}{\int S(|c-c_t|)dF(c)} dF(c|p)\right] dF(p)$. When the cue is (p_t, c_t), then with multiplicative similarity between price and context, the price norm $p^n(p_t, c_t)$ is the average price under the distorted measure $\frac{S(|p-p_t|) \int S(|c-c_t|)dF(c|p)}{\int S(|p-p_t|)dF(p) \int S(|c-c_t|)dF(c)} dF(p)$. We derive the price norm under the assumptions of the Proposition and using similarity function (3). Denote $\delta_p = \delta\lambda_{p,t}$ the strength of similarity in recall given a price cue, so that the first case where there is no price cue arises for $\delta_p = 0$.

The distorted distribution is the product of two normal distributions, namely the undistorted database with mean $\mu = [\bar{p}, \bar{c}]$ and variance matrix $\Sigma_F = \begin{bmatrix} \pi^2 & \rho\pi\gamma \\ \rho\pi\gamma & \gamma^2 \end{bmatrix}$ and the similarity distribution with mean $\kappa = [p_t, c_t]$ and variance matrix $\Sigma_{sim} = \begin{bmatrix} \frac{1}{2\delta_p} & 0 \\ 0 & \frac{1}{2\delta_p} \end{bmatrix}$. This product the variance matrix $\Sigma = \left(\Sigma_{sim}^{-1} + \Sigma_F^{-1}\right)^{-1}$, and the price norm is then the top element of the vector $\Sigma\Sigma_F^{-1}\mu + \Sigma\Sigma_{sim}^{-1}\kappa$. Plugging in Σ_F and Σ_{sim} and simplifying, we find:

$$p^n(p_t, c_t) = \frac{\overline{p} + 2\delta\gamma^2\mathbb{E}_F(p|c_t) + \left[2\delta_p\pi^2 + 4\delta_p\delta\gamma^2\pi^2(1-\rho^2)\right]p_t}{1 + 2\delta\gamma^2 + \left[2\delta_p\pi^2 + 4\delta_p\delta\gamma^2\pi^2(1-\rho^2)\right]}$$

where $\mathbb{E}_F(p|c_t) = \overline{p} + \rho\frac{\gamma}{\pi}(c_t - \overline{c})$. The case in which only context acts as a cue can be assessed as the limit case in which $\delta_p \longrightarrow 0$, so that the above equation becomes $\frac{\overline{p}+2\delta\gamma^2\mathbb{E}_F(p|c_t)}{1+2\delta\gamma^2}$, as in Equation (7). When price is also a cue price interference is finite, $\delta_p \longrightarrow \delta$, so that we obtain Equation (8).

Corollary 2. By inspection of Equation (9), together with the fact that $\mathbb{E}_F(q|c_t) = \overline{q} + \rho\frac{\gamma}{\pi}(c_t - \overline{c})$.

Proposition 3. Define

$$DV\big(p_t, p^n(p_t)\big) \equiv p^n(p_t) + \sigma\big(p_t, p^n(p_t)\big)[p_t - p^n(p_t)]$$

Setting $DV\big(p_t, p^n(p_t)\big) = p_t$ becomes $\sigma\big(p_t, p^n(p_t)\big)[p_t - p^n(p_t)] = p_t - p^n(p_t)$, with solutions $p_t = p^n(p_t)$ and $\sigma\big(p_t, p^n(p_t)\big) = 1$. Because, under the assumptions of the Proposition, $p^n(p_t)$ is a linearly increasing function of p_t, the condition $p_t = p^n(p_t)$ has a unique solution. From Equations (7, 8), this solution is given by $p_t = p^n(c_t)$. In turn, $\sigma\big(p_t, p^n(p_t)\big)$ is an increasing function of p_t for $p_t > p^n(c_t)$ and a decreasing function of p_t for $p_t < p^n(c_t)$. To see this, write:

$$\frac{\partial\sigma\big(p_t, p^n(p_t)\big)}{\partial p_t} = \sigma'\left(\frac{M_t}{m_t}, 1\right)\frac{\overline{p}}{m_t^2} sgn\big(p_t - p^n(p_t)\big)$$

where $m_t \equiv \min[p_t, p^n]$, and $M_t \equiv \max[p_t, p^n]$, as can be checked by considering the two cases individually. Because $\lim_{x\to 0,\infty} \sigma(x, 1) = \sigma > 1$, there exist unique price thresholds p_L, p_H satisfying $0 < p_L < p^n(\overline{p}, c_t) < p_H$ such that $DV\big(p_t, p^n(p_t)\big) = p_t$ if and only if $p_t \in \{p_L, p^n(c_t), p_H\}$. Furthermore, it is easy to see that $DV(p_t, p^n)$ is monotonically increasing in p_t, since both $p^n(p_t)$ and $\sigma\big(p_t, p^n(p_t)\big)[p_t - p^n(p_t)]$ are.

Undervaluation $DV(p_t, p^n) < p_t$ occurs when:

$$\big(\sigma\big(p_t, p^n(p_t)\big) - 1\big)[p_t - p^n(p_t)] < 0$$

namely when $p_t < p^n(p_t)$ and $\sigma\left(p_t, p^n(p_t)\right) > 1$ or vice versa. Thus, there is undervaluation for $p_t < p_L$, which guarantees both conditions, but also for $p_t \in (p^n(\overline{p}, c_t), p_H)$ where $p_t > p^n(p_t)$ but $\sigma\left(p_t, p^n(p_t)\right) < 1$. Similar arguments show that in the complementary regions there is overvaluation, $V(p_t, p^n) > p_t$.

We now examine the curvature profile of $DV(p_t, p^n)$. To do so, consider first the region where $p_t > p^n(p_t)$. Denoting $x = \frac{p_t}{p^n} > 1$ and $\alpha = \frac{2\delta\pi^2}{1+2\delta\pi^2}$ we can rewrite $\frac{\partial DV(p_t,p^n)}{\partial p_t}$ as:

$$\frac{\partial DV(p_t, p^n)}{\partial p_t} = \sigma'(x)(x-1)(1-\alpha x) + \sigma(x)(1-\alpha)$$

where $1 - \alpha x = (1-\alpha)\frac{\overline{p}}{p^n} > 0$. It then follows that $\partial^2_{p_t} DV$ is proportional to

$$\sigma''(x)(x-1) + 2\sigma'(x)$$

For x close to 1 this expression is positive, because $\sigma'(x) > 0$. However, as x increases it tends to decrease if $\sigma''(x) < 0$. More generally, provided the expression decreases monotonically and reaches negative values, that is it is negative for $x = \lim\limits_{p_t \to +\infty} \frac{p_t}{p^n(p_t,c_t)}$, then it has exactly one zero for $x > 1$. In this case, $DV(p_t, p^n)$ has exactly one inflection point for $p_t > p^n(p_t)$.

A similar calculation shows that, under the same conditions, $DV(p_t, p^n)$ has exactly one inflection point for $p_t < p^n(p_t)$. Finally, we assume differentiability of $DV(p_t, p^n)$ at $p_t = p^n(p_t)$, which is equivalent to differentiability of $\sigma(x, 1)$ at $x = 1$. Given homogeneity of degree zero, this requires setting $\sigma'(1) = \sigma''(1) = 0$, in which case $p_t = p^n(p_t)$ is the third inflection point.

Proposition 4. The value of p_H solves $\sigma\left(p_H/p^n(p_H)\right) = 1$, which entails:

$$p_H = z p^n(p_H, c_t),$$

for $z > 1$ and for contexts such that $p^n(p_H, c_t) > 0$. If $1 + 2\delta\gamma^2 + (1-z)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1-\rho^2)] > 0$, this condition implicitly defines:

$$p_H = \frac{z[\overline{p} + 2\delta\gamma^2 \mathbb{E}_F(p|c_t)]}{[1 + 2\delta\gamma^2 + (1-z)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1-\rho^2)]]},$$

If $1 + 2\delta\gamma^2 + (1 - z)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)] < 0$, then p_H is not defined (there is insensitivity to all prices above the norm). Likewise, the value of p_L solves $\sigma(p^n(p_L)/p_L) = 1$, which entails:

$$zp_L = p^n(p_L, c_t),$$

for $z > 1$ and for contexts such that $p^n(p_L, c_t) > 0$. If $\mathrm{z} + z2\delta\gamma^2 + (z - 1)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)] > 0$, this condition implicitly defines:

$$p_L = \frac{\overline{p} + 2\delta\gamma^2\mathbb{E}_F(p|c_t)}{\mathrm{z} + z2\delta\gamma^2 + (z - 1)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)]}$$

If instead $\mathrm{z} + z2\delta\gamma^2 + (z - 1)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)] < 0$, then p_L is not defined (there is insensitivity to all prices below the norm).

It is evident that both p_H and p_L increase in $\overline{p}$. It is also evident that, given that $z > 1$, higher π^2 increases p_H while reduces p_L. It is also immediate, after some algebra, to see that – again because $z > 1$ – higher $\overline{p}$ or higher $\mathbb{E}_F(p|c_t)$ increase the range $p_H - p_L$.

Corollary 3. Suppose that $p_t < p^n(\overline{p}, c_t)$. Then, it is immediate to see that it also the case $p_t < p^n(p_t, c_t)$. As a result, the decision value of price DV is less than the true price p_t if and only if:

$$(\sigma(p^n(p_t, c_t)/p_t) - 1)[p_t - p^n(p_t, c_t)] < 0,$$

which boils down to $\sigma(p^n(p_t, c_t)/p_t) > 1$, or $p^n(p_t, c_t) > zp_t$. This condition becomes:

$$p_t < p_L = \frac{\overline{p} + 2\delta\gamma^2\mathbb{E}_F(p|c_t)}{\mathrm{z} + z2\delta\gamma^2 + (z - 1)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)]}$$

$$= \frac{1 + 2\delta\gamma^2}{\mathrm{z} + z2\delta\gamma^2 + (z - 1)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)]} p^n(\overline{p}, c_t),$$

which matches the condition in the proposition by the definition:

$$\alpha \equiv \frac{1 + 2\delta\gamma^2}{\mathrm{z} + z2\delta\gamma^2 + (z - 1)[2\delta\pi^2 + 4\delta^2\gamma^2\pi^2(1 - \rho^2)]},$$

where clearly $\alpha < 1$ by $z > 1$ and where α decreases in π^2. Following the same logic one can establish the stated properties for overvaluation of price in the opposite case in which $p_t > p^n(\overline{p}, c_t)$.

Proposition 5. Consumer $i \in \{h, l\}$ has a database $N(\overline{p}_i, \overline{c}_i, \pi^2, \gamma^2, \rho)$. Suppose that δ is finite. Given context c, consumer i's price norm is equal to:

$$p_i^n(c) = \frac{\overline{p}_i + 2\delta\gamma^2 \mathbb{E}_F(p|c)}{1 + 2\delta\gamma^2}.$$

The willingness to pay of consumer i is the solution p to the equation:

$$DV(p, p_i^n(p)) = q \Leftrightarrow \{\sigma[p, v_1 p + v_2 p_i^n(c)] - 1\}[p - v_1 p - v_2 p_i^n(c)] = q - p,$$

where v_1, v_2 are positive coefficients. It is easy to see that the left-hand side of the above equation is monotonically increasing in p. As a result, the following condition describes when consumer i has a WTP p above or below q:

$$p > q \;\Leftrightarrow \{\sigma[q, v_1 q + v_2 p^n(\overline{p}_i, c)] - 1\}[q - v_1 q - v_2 p^n(\overline{p}_i, c)] < 0.$$

We are looking for a condition whereby $p > q$ for $i = h$ and $p < q$ for $i = l$. The left-hand side of the above inequality has the following property. There are three thresholds $p_* < \hat{p} < p^*$ such that:

$$\{\sigma[q, v_1 q + v_2 p^n(\overline{p}_i, c)] - 1\}[q - v_1 q - v_2 p^n(\overline{p}_i, c)] > 0 \;\; for \;\; p^n(\overline{p}_i, c) < p_* \; and \; \hat{p} < p^n(\overline{p}_i, c) < p^*,$$

$$\{\sigma[q, v_1 q + v_2 p^n(\overline{p}_i, c)] - 1\}[q - v_1 q - v_2 p^n(\overline{p}_i, c)] < 0 \;\; for \;\; \hat{p} < p^n(\overline{p}_i, c) < p_* \; and \; p^n(\overline{p}_i, c) > p^*.$$

As a result, given that $p^n(\overline{p}_l, c) < p^n(\overline{p}_h, c)$, a sufficient condition under which the willingness to pay of h is larger than the willingness to pay of l is $p^n(\overline{p}_l, c) < p_*$ and $p^n(\overline{p}_h, c) > p^*$, which is equivalent to:

$$\overline{p}_l + \frac{2\delta\rho\pi^2(c - \overline{c}_l)}{1 + 2\delta\gamma^2} < p_*,$$

$$\overline{p}_h + \frac{2\delta\rho\pi^2(c - \overline{c}_h)}{1 + 2\delta\gamma^2} > p^*,$$

Which can be rewritten as:

$$\overline{p}_l < p_* - \frac{2\delta\rho\pi^2(c - \overline{c}_l)}{1 + 2\delta\gamma^2} \equiv p_l(c, c_l),$$

$$\overline{p}_h > p^* - \frac{2\delta\rho\pi^2(c - \overline{c}_h)}{1 + 2\delta\gamma^2} \equiv p_h(c, c_h),$$

The conditions are satisfied provided $\overline{p}_l$ and $(c - \overline{c}_l) > 0$ are low enough, and provided $\overline{p}_h$ and $(c - \overline{c}_h) < 0$ are high enough. For given $\overline{p}_l$ and $\overline{p}_h$, the condition is more likely to be satisfied when π^2 is low.