# Memory, Attention, and Choice

Pedro Bordalo, Nicola Gennaioli, Andrei Shleifer[1]

Revised May 7, 2019 (original version 2015)

**Abstract.** Building on the textbook description of associative memory (Kahana 2012), we present a model of choice in which options cue recall of similar past experiences. Recall shapes valuation and choice 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 provides a unified account of many well documented choice puzzles including experience effects, projection and attribution biases, background contrast effects, and context- dependent willingness to pay. The results suggest that well-established psychological processes – memory-based norms and attention to surprising features – are key to understanding decision-making.

[1] The authors are from University of Oxford, Universita Bocconi, and Harvard University, respectively. We are grateful to Dan Benjamin, Paulo Costa, Ben Enke, Matt Gentzkow, Sam Gershman, Thomas Graeber, Michael Kahana, Spencer Kwon, George Loewenstein, Sendhil Mullainathan, Josh Schwartzstein, Jesse Shapiro, Jann Spiess, Linh To, and Pierre-Luc Vautrey for valuable comments. Shleifer thanks the Sloan Foundation and the Pershing Square Venture Fund for Research on the Foundations of Human Behavior for financial support.

## 1. Introduction

Memory appears to play 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 helps predict decision values, even before the consumer enters the shop and sees the product. It serves as an *anchor* for valuation.

But memory does more than that. After the traveler enters the shop, seeing a specific brand and price triggers retrieval of more specific memories of similar past experiences with this brand, including prices. In the neoclassical model of choice, these memories do not influence valuation – neither the expected utility from drinking water nor the disutility of paying the price – so at this point memory no longer shapes behavior. Reality is arguably different. If the traveler is a first-time traveler to an airport, he is likely shocked by the price, which is much higher than what he is used to paying downtown. 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. Memories here act as references, helping valuation to *adjust* to reality: they determine whether reality is normal or surprising.

But what information is recalled? How does it shape valuation? We adapt to an economic setting the textbook model of associative memory (Kahana 2012) and propose a new model of choice based on the following idea. When we consider buying a good, our brain spontaneously forms a memory-based norm of its price and quality, by retrieving past experiences that are similar to the current choice in either hedonic or contextual dimensions. If there is little or no discrepancy between the choice and the norm, current valuation is assimilated to that of retrieved memories, with minimal adjustment. This captures the role of memory as an anchor. Both the first time traveler who thinks about water outside the shop and the seasoned traveler who sees a price close to normal are instances of this mechanism. If instead the discrepancy between reality and the norm is large, valuation is adjusted away from the norm. This captures the role of memory as a reference. The first-time traveler shocked by the abnormally high price at the shop falls in this case.

In this approach, valuation and choice based on objective attributes are inextricably linked to recalled experiences. Memories are automatically retrieved in virtually every choice situation. Valuation sometimes assimilates choice options to those memories, and sometimes exaggerates the difference between them. Valuation (and choice) is a two-stage process, with memory controlling the retrieval of norms, and attention shaping the adjustment of valuation to deviations from these norms.

This two-step approach to choice has antecedents in both psychology and neuroscience. Most directly, it captures the psychological judgment of normality and surprise that Kahneman and Miller (1986) describe in "Norm Theory". In their view, events spontaneously bring to mind norms by "acting as a reminder of similar events in the past". Such memory based norms then shape how the event is assessed. Kahneman and Miller treat surprise as an event that is different from the norm it elicits. In our model, surprise draws attention and shapes valuation. These ideas are also related to an emerging paradigm in the neuroscience of sensory perception, in which the brain spontaneously forms predictions of the environment known as mental representations (e.g., Hawkins and Blakeslee 2004; Clark 2013). Such predictions are necessary to perform everyday activities: when walking our brain predicts the slope of the ground, when picking up a bag it predicts its weight, when opening a door it predicts the door knob's temperature. Typically, these predictions are not conscious. For example, when the temperature of the doorknob is close to the usual, the experience is perceived as normal and the pre-existing mental prediction offers a parsimonious guide to action with little need for conscious attention. When instead the prediction is violated, as when the doorknob is unusually hot, we are surprised: our attention is engaged and our initial representation is adjusted. This approach, which we refer to as the predictive brain model, has not been evaluated for higher-level cognitive tasks such as economic choice, but it offers a useful guide for our analysis of norms and surprise-driven adjustment.

To formalize memory-based norms, we follow the literature in psychology which views episodic memory as a database of past experiences. This approach highlights two crucial features of this database. First, experiences consist not only of intrinsic values (quality and price in our case) but are also tagged by the spatio-temporal context in which they occurred. Second, a task (a choice problem in our case) provides cues that trigger automatic retrieval of past experiences from the database. This recall process is driven by similarity between the cue and past experiences: intrinsic and contextual cues activate similar past experiences and block recall of dissimilar past experiences. We follow the literature by assuming similarity to be Euclidean distance in attribute space. Empirically, it can be estimated from similarity judgments of multidimensional objects (Pantelis et al 2008). Decades of memory research show that this structure captures robust empirical regularities about recall, and in particular its selective nature.

To formalize attention to surprise, we use the Salience Theory from Bordalo, Gennaioli, and Shleifer (BGS 2012, 2013). Salience Theory holds that, when evaluating a choice option, we overweigh features that are very different from the norm, while neglecting features that are close to it. This idea captures strong regularities in sensory perception such as Weber's law. The theory helps explain a wide range of evidence regarding instability of risk preferences and preference reversals in consumer choice. Here we view recalled experiences as the spontaneous background against which most of our evaluations take place. Salience can then capture both anchoring to the norm and strong adjustment to large surprises.

Anchoring to recalled experiences and adjustment to surprises help explain and unify several well documented puzzles in choice. The recalled norms based on extensive experience are typically well-adapted to reality, entailing minimal surprise and close to rational behavior in a wide range of situations. Distortions in valuation and choice arise when a database of selected experiences or a misleading contextual cue causes retrieval of past instances that are different from the current choice, creating an artificial surprise. This mechanism yields several predictions, and extends the role of memory in choice significantly beyond earlier models such as Gilboa and Schmeidler (1995) and Mullainathan (2002). We present four principal findings.

First, memory's role as an anchor for valuation accounts for many choice puzzles involving the valuation of future benefits or costs. Many of these phenomena are driven by contextual similarity. One example is experience effects: when thinking of a stock market investment, we draw on memories of similar investing experiences rather than the statistics of returns (Malmendier and Nagel 2011). Another example is the projection bias: when predicting the utility of a convertible car, we are unduly influenced by the cue of current weather (Busse et al 2015). As in the case of stock investing, contextual similarity is key, but here it activates recall of driving in similar weather. Likewise, when choosing a product in a supermarket aisle, some price components such as taxes are not visible, and both the tax-free price and the aisle location cue recall of similar experiences with tax free prices (Chetty, Looney and Kroft 2009). Crucially, such estimates obey well known mechanisms of recall, which deviate in in a systematic way from optimal forecasts. These deviations reflect differential accessibility of information from memory, not ignorance (as for example in Schwartzstein 2014). In fact, when weather conditions change, or if we are cued about sales taxes, we retrieve those instances from memory and change our valuation accordingly. The structure of similarity helps assess what information is retrieved when we make decisions.

Second, memory also acts as a background reference when evaluating *observable* attributes. In this way, it can help clarify the evidence on reference points documented in previous work. Dependence on past experiences stored in the database causes memory-based norms to look like adaptive reference points (Kahneman and Tversky 1979, Della Vigna et al. 2017). At the same time, similarity causes these norms to change across contexts, as in rational expectations reference dependent choice models (Koszegi and Rabin 2006). As the seasoned traveler selectively recalls similar airport experiences, he forms a high price airport norm that looks rational at the airport. Crucially, similarity creates even more flexibility: as the traveler sees the actual brand and price, he retrieves further memories that are even more similar to this specific realization. These memories then form the ex-post norm, which displaces (at least in part) his ex-ante thoughts. When prices are highly volatile, extreme realized prices look more normal, consistent with evidence from marketing (e.g., Niedrich, Sharma and Wedell 2001). In fact, similarity-based recall can yield very accurate ex-post norms, accounting for rational-looking behavior even in uncertain environments.

Third, by shaping the adjustment of valuation, retrieved norms help unify choice puzzles involving the valuation of observable attributes such as quality and price. When the observed attributes are close to the retrieved norm, the adjustment is small. This mechanism yields inattention to small price changes: seeing a price close to frequently-observed prices retrieves many prices similar to it, and valuation adjusts little to these small differences. This can help explain evidence from marketing that demand is highly insensitive around normal prices (Koschate-Fischer and Wullner 2017), why firms often implement frequent small price increases (Nakamura and Steinsson 2008), why price dispersion is large and increases with price levels (Pratt, Wise and Zeckhauser 1979), as well as evidence of inertia in consumer choice (Chandra, Handel and Schwartzstein 2018). In contrast, we predict that observing an unusual and extreme price generates a large adjustment, because the norm remains anchored to the moderate modal price we are used to. This effect may help explain why firms tend to implement large temporary sales (Nakamura and Steinsson 2008) and why during sales events regular prices are often visibly displayed.

More broadly, conventional contrast effects arise as a special case in which context creates superficial similarity across fundamentally different situations. Superficial contextual similarity triggers the recall of inaccurate norms, creating artificial surprise and large adjustment. In background contrast experiments (Simonson and Tversky 1992), seeing high prices in the first stage increases willingness to buy in the second stage. Our explanation is that contextual similarity across experimental stages causes subjects to retrieve the high first-stage prices, so that any moderate price looks surprisingly low in the second stage. Likewise, superficial similarity in apartment hunting causes movers to a cheaper city to recall high rental prices in the previous city of residence, causing them to spend more on rent than comparable locals (Simonsohn and Loewenstein 2006). The same mechanism works through priming: cueing a beachgoer with the context "resort" as opposed to the context "shack" retrieves a higher price norm that increases his willingness to pay for the very same beer to be consumed at the beach (Thaler 1985). Memory research has developed a rich set of tools to measure contextual similarity and its influence on recall. Our model can guide the use of these tools in the analysis of judgment of normality and economic choice, and in particular provides a systematic account of multiple documented instances of instability of choice.

Fourth, memory sheds light on fairness norms, and in particular on the observed instability in what is perceived as fair behavior. Kahneman, Knetsch, and Thaler (KKT 1986) argue that behavior is assessed as fair at least partially "because it is normal, not because it is just" in some abstract sense. Why does the mere exposure to generous offers in the ultimatum game makes reasonable offers look stingy, so that subjects reject them (Herz and Taubinsky 2017)? In our view, after seeing high offers, contextual similarity across experimental stages causes retrieval of a high fairness norm, as in the background contrast effect. Why do people get upset by wage cuts, but less so by bonus cuts, or why do they get upset if prices of snow-shovels increase during a snow storm but not if prices of flowers rise on Valentine's Day? Again, part of the answer may be that, despite being fundamentally alike, these situations trigger different norms because of superficial similarity to different past experiences. The norm for what is fair changes with experience of context, and behavior follows.

The paper studies in detail the implications 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 well-understood structure, which in turn creates empirical regularities. Using these insights to understand economic choice provides unification, clarity, and new testable predictions.

We develop our analysis in seven steps. In Section 2, we briefly summarize the research on the psychology of memory and attention that motivates our approach, and present the central equation that describes valuation as a two-stage process of anchoring and adjustment. In Section 3, we introduce our model of norms retrieved from memory. In Section 4, we study the implications for choice and discuss how our model accounts for a range of puzzles. In Section 5, we show how contextual similarity yields contrast effects. In Section 6 we discuss fairness norms. In Section 7, we connect our approach to previous research on memory and choice and to the psychology literature on assimilation and contrast, two phenomena that naturally emerge from our model. Section 8 concludes with a discussion of some open problems.

## 2. An Overview of the Model.

We illustrate the logic of our model with an example, and introduce the motivating psychological evidence. Consider a thirsty traveler arriving at an airport and deciding whether to buy a bottle of water. First, the traveler must decide whether to look for water. Second, after entering the airport shop the traveler must decide whether to buy the available water at its current price. In the standard economic model, the traveler forms his best prediction of the water's quality and price and goes to the shop if the former exceeds the latter (assuming he has linear utility, as we do throughout this paper, and neglecting any transaction costs). If the traveler is already at the shop, he sees quality $q$ and price $p$, and buys water if $q > p$. We next describe how, in both cases, memory colors the traveler's evaluation and choice.

### 2.1 Anchoring to Norms and Adjustment

When deciding whether to look for water, the traveler's estimates of price and quality come from memory. That is, they obey well known mechanisms of recall which, as we show in Section 2.2, differ systematically from rational inference. Formally, denote the quality and price norm for water by $(q^n, p^n)$. The traveler goes to the shop if and only if $q^n \geq p^n$. This decision is based entirely on the recalled norm.

But even a traveler who sees the brand and price of water $(q, p)$ at the shop relies on memory to assess whether it is worth buying. Of course, it matters how thirsty he is, but memory enables him to retrieve the pleasure of drinking this water brand and the past experiences with paying for it. These recollections anchor the value of the water in front of him, and help him assess whether the current price is excessive compared to what he is used to. Formally, the stimulus $(q, p)$ cues recall of a more specific norm which, abusing notation, we also denote by $(q^n, p^n)$. We then assume that valuation is described by:

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

Valuation is anchored to the norm and then adjusted in direction of the errors, or surprises, between the stimulus and the norm. The weighting of errors is modulated by a salience function $\sigma(x, x^n) \geq 0$, where $x$ is a generic attribute. In line with our previous work (BGS 2012, 2013), this assumption captures the idea that our attention is drawn to large deviations from to a background memory state. In particular, the salience function $\sigma(x, x^n)$ increases in the proportional difference between $x$ and $x^n$ so that, in line with the Weber-Fechner law, the perception of differences increases in log terms.

Equation (1) has an intuitive interpretation. The norm $(q^n, p^n)$ is a spontaneous representation of the stimulus $(q, p)$ that cues it, capturing the ex-post "normal" quality and price. If the norm is close to the stimulus, the situation is "normal" and the consumer essentially chooses based on the initial representation $(q^n, p^n)$. The price of water may be a bit different, but these differences are small and the traveler is inattentive to them, so that the adjustment $\sigma(x, x^n)(x - x^n)$ is small. If instead the norm is very different from the stimulus, we say that the surprise is large. For instance, the bottle may be warm, or the price may be much higher than the consumer normally pays. The consumer's attention is driven to the surprising features, leading to a strong adjustment of the initial representation $(q^n, p^n)$ in the direction of the stimulus $(q, p)$. In this case $\sigma(x, x^n)(x - x^n)$ is large. As in Norm Theory, we react strongly to large surprises.

A related interpretation, in line with the psychology of assimilation and contrast (which we review in Section 7) is that Equation (1) reflects the use of mental categories. Here 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 it, as in categorization models where category members are assimilated toward the category exemplar (e.g. Mullainathan 2002, Fryer and Jackson 2008). If instead the current stimulus is far from the normal category, it is placed in the "cheap" or "expensive" category, so its valuation is shaped by contrast.

### 2.2 Memory

To close the model we use the psychology of memory to specify how norms form based on past experience. Since the 1880s, a large body of experimental work has examined episodic memory, or memory of past experiences. The evidence shows that recall is a spontaneous and subconscious process in which a current experience, the cue, stimulates retrieval. As described by Kahana (2012), recall obeys two principles:

1. It is associative, driven by *similarity*: presenting a stimulus facilitates recall of items from memory that are similar to that stimulus.
2. It is subject to *interference*: recall of items of given similarity to the stimulus is weakened or blocked entirely in the presence of more similar items in working memory.

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 tests show the role of similarity because i) the probability of recall is higher for items that belong on the list (that is, when 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 relevant word pairs. These tests reveal the role of interference: if the cued word appears in many word pairs (so it is associated with many other words), recall of each association is less likely. This is known as the fan effect (Anderson and Reder 1999). Interference effects are stronger for items that are more similar, so similarity shapes cued recall as well.

In Section 3 we present a model in which a consumer considering a good automatically forms a memory based prediction, which we denote $(q^n, p^n)$, through a retrieval process that follows these principles. To this end, we adapt to an economic setting the standard model of similarity based recall (Kahana 2012). In this model, past experiences are encoded as vectors of hedonic and contextual attributes. For instance, when we buy water downtown, our memory records its brand and price, which are intrinsic or hedonic attributes, but also contextual conditions such as the time and location of the experience. Episodic memory is then described as an *event x attribute* matrix database of past experiences.[2]

In this model, one cue for retrieval is context. Thirst and the downtown location prime selective recall of past water buying experiences in similar locations. But intrinsic attributes such as quality or price also act as cues. Seeing water for sale at a low price spontaneously cues recall of past experiences of similar low prices. Similarity has two important implications. First, recall of similar past prices or qualities is a form of adaptation of norms towards the current choice, favoring an accurate representation. Second, superficial contextual similarity may bring to mind past prices or qualities that differ from the present choice, creating an inappropriate norm and an artificial surprise. This mechanism plays a key role in background contrast effects.

[2] We abstract from semantic memory, a term that covers functional associations and rule based thinking (e.g., recalling "glass" after "milk"). We return to these issues in the conclusion.

## 3. A Model of Memory

We first describe the memory database, the process of cued recall and the formation of memory-based norms. We then show that, even in settings where the good's attributes are unobserved, this model can shed light on a range of well-documented memory effects in choice, such as experience effects, projection bias, attribution bias, and neglect of shrouded attributes.

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, whereas $c$ 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.[3] For simplicity, we abstract from rehearsal of past options (Mullainathan 2002). We assume context $c$ to be cardinal and restrict it to the most relevant dimension for the application at hand. In reality, $c$ is multi-dimensional and categorical. The model can be extended to capture this richness.[4]

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 bottled water may prime a substitute category such as soft drinks. Price may bring to mind other goods one can purchase

[3] 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.
[4] Cardinality of $c$ is used for exposition; to define a notion of similarity (or inverse psychological distance) between contexts, see Definition 1. At the end of this Section, we discuss how to measure similarity across contexts more broadly.

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).[5]

### 3.1 Cues, Similarity, and Norms

The current choice, which we denote with subscript $t$, starts with a cue $\kappa_t$, which identifies the good (e.g., "evaluating the utility of a bottle of water") and thus the database $F(e)$ of relevant experiences, along with additional contextual detail. 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 that $\kappa_t = c_t$. This is the traveler thinking whether to look for the shop, where only the "airport" cue is available.

A generic cue $\kappa_t$ stimulates recall of similar past experiences, where similarity acts on the set of attributes defining the cue. The cue is written $\kappa_t \equiv \left(\lambda_{q,t} q_t, \lambda_{p,t} p_t, c_t\right)$ where the indicator $\lambda_{x,t}$ takes value one if attribute $x$ is observed at time $t$, and takes value zero otherwise (it is natural to assume that context is always observed, including the category of the good under consideration). Following the standard assumption, we define similarity in terms of the geometric distance in attribute space.

**Definition 1** *The similarity of a past experience* $e \equiv (q, p, c)$ *to the cue* $\kappa_t \equiv \left(\lambda_{q,t} q_t, \lambda_{p,t} p_t, c_t\right)$ *is given by the multiplicatively separable distance:*

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

*where* $S_k: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ *is decreasing for* $k = 1,2,3$. *The exponential specification takes the form:*

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

[5] Although easier to recall, surprising events are hardly a driving force of norms. A meal at an extraordinary restaurant is memorable, but it does not alter our norm for restaurant meals, which is based on recall of more ordinary meals.

*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.[6] Past experiences are activated to different degrees, depending on similarity with $\kappa_t$. Formally, we view 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 $\kappa_t$ is given by:*

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

*The quality and price norms for cue $\kappa_t$ are the quality and price components of the average experience weighted by relative similarity to the cue:*

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

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 $\kappa_t$, because $w(e, \kappa_t)$ denotes relative similarity. This captures interference, whereby more similar memories block less similar ones (Kahana 2012).[7]

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

[6] 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.

[7] Equation (3) 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.

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. If price and context are uncorrelated in $F(e)$, the observed context is irrelevant for norms. Moreover, 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 price norms. Consider a consumer purchasing water downtown. His experiences accumulate in contexts that differ in a myriad of attributes such as brand, location, layout, etc. This bewildering variability across downtown locations, however, does not influence his price norm, because it 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, norms coincide with rational expectations.

Suppose that the consumer sees both context $c_t$ and price $p_t$. Because context is unrelated to price, it is again irrelevant. Observing price is however important: it triggers recall of similar past prices. The norm still depends on the modal price $\overline{p}$, which is experienced most frequently, but due to similarity is also adjusts toward the current price $p_t$. Similarity is thus a mechanism of "on the fly" adaptation to the current context. Ex post adaptation can be transparently characterized in the following special case of Proposition 1.

**Corollary 1** *If the marginal distribution $dF(p)$ is Gaussian with mean $\overline{p}$ and variance $\pi^2$, 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 1, where the green line depicts the norm conditional on price $p$.

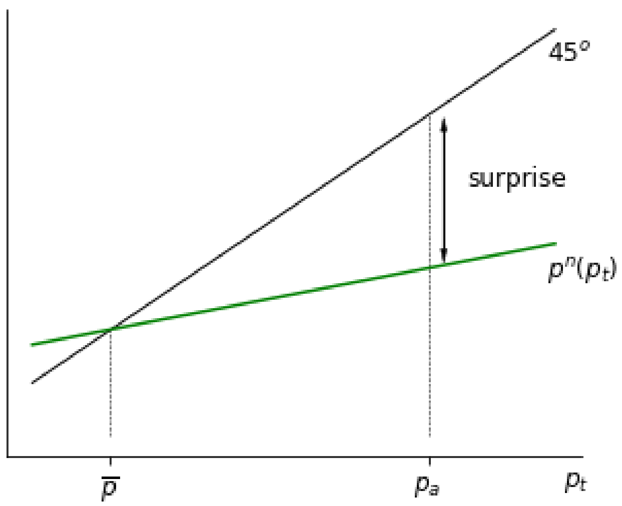


**Figure 1.** The figure plots the norm $p^n$ (in green) as a function of observed price $p_t$ when the database is Gaussian and context is uncorrelated with price (Equation 6).

The norm increases linearly with the observed price, but less than one for one. This is due to the disproportionate recall of the modal past price $\overline{p}$. When seeing the high price of water at the airport $p_a$ for the first time, recall of the modal downtown price $\overline{p}$ brings the norm down. This entails a big surprise, as shown in Figure 1.

The norm adapts better to the observed price, including the high airport price, when $\delta$ is higher: a stronger role of similarity in recall eases retrieval of similar past prices and interferes with retrieval of different, if frequent, prices. In Figure 1, this means that the gray line is closer to the $45^o$ line that goes through $\overline{p}$. Evidently, a steeper norm in Figure 1 reduces the surprise associated with seeing a large price $p_t$. But adaptation is also stronger if historical price variability $\pi^2$ is higher, which causes the norm to be a steeper function of the current price. The intuition for this effect is simple. When price variability is higher, the consumer seeing a high price recalls many instances of similarly high prices, which interfere with retrieval of low prices. Adaptation is thus stronger and surprise is contained.

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

**Proposition 2** *Suppose that $F(e)$ is Gaussian with mean $(\overline{p}, \overline{c})$, variances of price and context $\pi^2$and $\gamma^2$, and correlation $\rho$ between them. Then, 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 are adjusted to take the correlation $\rho$ between price and context into account. This adjustment occurs through the term $\mathbb{E}_F(p|c_t)$, which denotes the average price observed at context $c_t$ in the database $F(e)$. Contextual similarity creates a form of statistical conditioning. After experiencing an association of a context with high prices, this context primes a high price norm. A traveler entering the airport recalls past airport experiences and the high price associated with them, because in the airport context $\mathbb{E}_F(p|c_t)$ is higher (Equation 7). When the price is also a cue, it interacts with context (Equation 8). The price norm at the airport adapts to the airport context as described in Figure 2.

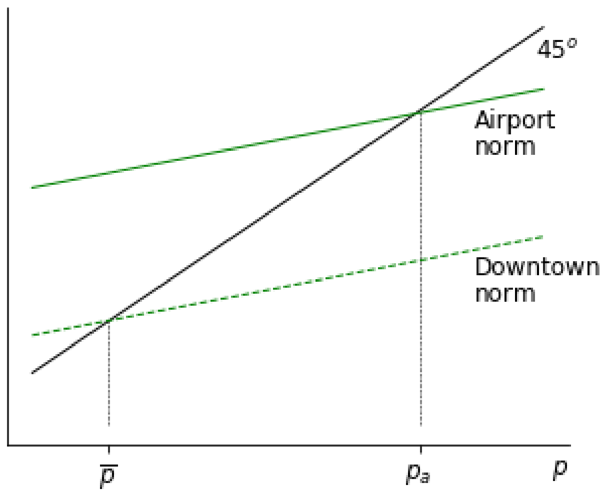


**Figure 2.** The figure plots the norm $p^n$ as a function of observed price $p$ and context $c \in \{c_{downtown}, c_{airport}\}$ (in green, dotted and solid respectively) when the database is Gaussian and context is correlated with price (Equation 8).

The airport context reminds the consumer of the higher prices associated with it, which causes an upward shift of the norm to the bold green line. This implies that a price that would be surprisingly high downtown is no longer surprising at the airport, as shown in Figure 2. This adjustment is reminiscent of rational expectations, with three key differences. First, the airport norm remains anchored to the modal experience $\overline{p}$, as shown in Equation (8). Second, the norm also adapts to the current observed price $p_t$, as evidenced by the positive slope of the dark gray line in Figure 2.

Third, and crucially, context is a malleable cue that can induce retrieval of inappropriate norms. Reminding a friend of an airport experience (or, in Thaler's example, reminding a beachgoer that the beer comes from a resort) can trigger the retrieval of a high price norm, making the current price feel normal or even surprisingly low. Recency effects are another example of the power and malleability of contextual cues. A recently experienced price or wage automatically comes to mind, influencing our judgments even if normatively irrelevant (Della Vigna et al. 2017). This effect can be captured in our model by including calendar time as a dimension of context.[8] In fact, the role of contextual cues is even broader: they can be triggered by many features (location, mode of presentation, explicit allusions to other contexts, etc) provided those features correlate with intrinsic attributes of the good. In this respect, context is a double-edged sword: it promotes adaptation to the environment but it can superficially cue misleading norms. Our model of memory sheds light on such effects through the structure of similarity, showing that many puzzling behaviors in the field and in experiments can be understood as a by-product of this mechanism.

Memory-based norms unify aspects of existing theories of reference points, because they combine a backward-looking component due to recall of past experiences with a forward-looking component due to selective adaptation to context.[9] In stable situations in which the database is populated by a few frequent experiences with no clear correlation between price and context, as in Corollary 1, memory-based norms

[8] This can be formalized using the contextual drift model (Kahana 2012) by assuming context at $t$ is the combination $\alpha c_t + (1-\alpha)c_{t-1}$ with $\alpha < 1$. Mullainathan (2002) is an early discussion of recency effects in economic decisions.

[9] Consider again buying water at the airport. During his second visit at the airport, the consumer remembers that water is more expensive there. Memory-based adaptation is more sluggish than rational adaptation (which is immediate), but it is faster and more selective than under mechanically adaptive reference points. Indeed, there are two "fair" prices of water: one downtown, another at the airport. With enough experience in each context, this process causes memory-based norms to look like rationally expected prices both downtown and at the airport.

subsume the intuition of the classic Kahneman and Tversky's (1979) "status quo" reference point. [10] Crucially, however, memory norms are not rigid, but adapt to changing conditions.

Theories of adaptive reference points include the Della Vigna et al. (2017) model with mechanically adaptive, backward-looking expectations, and Koszegi and Rabin's (2006) theory of forward-looking rational expectations.[11] Like mechanically adaptive reference points, memory-based norms are influenced by frequent past experiences $\overline{p}$, as evident from Equation $(6)$. At the same time, similarly to rational expectations reference points, memory norms selectively adjust to different contexts, as in Equation $(8)$. In fact, because memory-based norms adapt ex-post to the current realization, they are even more flexible than rational expectations. Seeing a high price reminds the consumer of similarly high past prices. In an environment with highly volatile prices, there can be full adaptation to the realized payoff. In such environments, memory-based norms generate reference points that match realized prices and thus entail undistorted valuations and rational decisions. This stands in contrast to models with reference points based on rational expectations, in which the reference coincides with the price realization (and hence the decision is rational) only when the environment is deterministic.

We close this discussion with a remark on taking the model to field data. Equation (8) shows that context in our model is very tractable: its impact on the norm is captured by $\mathbb{E}_F(p|c_t)$, which can in principle be observed. The main difficulty is to measure context itself, and in particular to measure similarity across categorical contexts such as locations. At some level, certain categories and thus similarity relations are self-evident. For example, it is natural to expect travelers at La Guardia to condition on "airport'' context as opposed to "close to New York City'', so that La Guardia is judged more similar to other airports than to downtown Manhattan. In other cases, similarity relations may be harder to grasp. For instance, is an airport in a poor country more similar to La Guardia or to the developing country itself? In these cases, our model offers guidance on how to measure similarity relations in light of the data.

---

[10] In most theories, the reference point is triggered by the situation, not by choice options. For instance, it consists of the decision maker's expected choice (Koszegi and Rabin 2006), or by his consumption (Kahneman and Tversky 1979).

[11] Bell (1985) also identifies the reference price as the rational expected price (see also Gul 1991). Barberis and Huang (2001) and Barberis and Xiong (2009) take a related approach in asset pricing with respect to the expected risk free rate.

In particular, we defined similarity as a weighted Euclidean distance along a set of observable attributes. This distance function can be estimated using data on observable attributes together with individual similarity judgments. A large literature in Psychology elicits similarity judgments directly from subjects, under the well-founded assumption that such judgments follow some deep-seated intuition and are widely shared. This work shows that the similarity judgments thus obtained predict subsequent judgments or decisions. Kahneman and Tversky (1972) show that stated similarity between a type (e.g., a feminist bank teller) and a group (e.g., women who were activist in college) aligns with the assessed likelihood of that type in that group. In our approach, a similar strategy can be applied to predicting consumer choice.

Pantelis et al (2008) go further: they elicit similarity judgments between synthetic faces that vary in several measurable attributes, and use this data to fit the similarity function in attribute space. They find that closer similarity (either stated or predicted) predicts a higher likelihood of confusion between the names associated with those faces. This evidence suggests that similarity across contexts can be elicited from subjects, that similarity predicts recall in line with memory models, and finally that similarity judgments allow one to fit similarity functions that can be used out of sample. These methods offer new opportunities to test the role of memory in economic choice.

## 4. Memory and Choice

We study memory and choice in two cases. In the first case, the consumer does not observe hedonic attributes. He assesses them by retrieving their norms from memory. For instance, the consumer considers investing in stocks and retrieves from memory a norm for returns. In the second case, the consumer observes the hedonic attributes of the good, but evaluates them in light of the retrieved norms. For instance, the consumer sees the brand and price of water but evaluates them in light of recalled experiences. There is also a third, mixed case in which one attribute is observed while the other is not (e.g., the purchase of a durable good). We do not formally study this case, but it is a straightforward combination of the first two.

### 4.1 Unobserved Hedonic Attributes

When neither quality nor price 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 utility is $q^n(c_t) - p^n(c_t)$. To illustrate, consider the valuation of quality only. From Equation (7) it follows that:

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

Memory influences valuation via: i) the average quality $\overline{q}$ experienced by the individual, and ii) the average quality $\mathbb{E}_F(q|c_t)$ associated with the current context. These two factors control how memory norms of quality depart from the rational expectations benchmark. We next show how they jointly account for a range of disparate evidence on estimation of quality, including experience effects, attribution bias, and projection bias. The same mechanism, applied to prices, sheds light on the neglect of shrouded attributes.

*Biased Databases: Experience Effects, Attribution Bias*

We start with the role of average past experience $\overline{q}$ in driving valuation. Malmendier and Nagel (2011) show that 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.

Our model traces this phenomenon to individual databases that contain a biased subset of past realizations. Suppose investor $i$ has had better experiences on average than $j$, namely $\overline{q}_i > \overline{q}_j$. The average valuation of stocks reported by these investors across different contexts will then differ, with investor $i$

reporting a higher valuation of stocks $\overline{q}_i$ than $\overline{q}_j$ reported by $j$.[12] This is due to the fact that when thinking about stocks investor $i$ retrieves better average experiences that are more common in his database.

This experience effect is very different from the valuation gap that would arise under rational learning. This is illustrated by the closely related phenomenon of the "attribution bias", whereby the valuation of a good to be consumed in the future is unduly influenced by the context of past experiences with it. Haggag et al. (2018) find that consumers expect higher utility from going to an amusement park if during a past visit to the park the weather was good. If consumer $i$ experienced better weather during his visit to the amusement park than consumer *j*, his memory-based anchor for valuation is higher, $\overline{q}_i > \overline{q}_j$. This distortion in valuation is due to the fact that *i'*s experiences are disproportionally formed in weather contexts $c_t$ in which valuation is higher. Even if consumers know that weather affects the enjoyment of the amusement park, they take their memories at "face value" because of the spontaneous association between the park and the pleasure of a past visit. Among other things, this implies, unlike with rational expectations, that beliefs can be changed by cueing specific weather conditions. We return to this issue below.

*Distorted Retrieval: Projection Bias and "Back of Mind" Inattention*

The difference between memory-based and rational valuation is made evident by another set of puzzles in which selective retrieval plays a role. In terms of Equation (9), these effects act through the conditioning term $\mathbb{E}_F(q|c_t)$ rather than through the unconditional frequency term $\overline{q}$ at play above.

A prominent example is the "projection bias". Conlin et al (2007) show that catalogue 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 better on the day they test-drive it, even if the consumer already owned a convertible. Using data from a large Chinese insurance provider, Chang et al. (2016) find that on days in which air pollution is high there is

[12] This prediction rests on the plausible assumption that memory is disproportionately populated with own experiences. While memory surely stores other information as well, we do not model explicitly the information acquisition process.

a spike in the purchase of health insurance, but that consumers subsequently cancel their contracts if during the free cancellation period air quality improves. In these examples, consumers appear to unduly weigh current conditions when estimating future benefits, and remarkably they do so even in cases where they have sufficient experience to know the value of the good across different conditions.

Equation (9) can account for this behavior. Upon seeing an extreme context $c_t$, say cold weather, memory automatically retrieves an extreme prediction $\mathbb{E}_F(q|c_t)$ of the utility of warm clothes, which raises valuation.[13] In principle, the consumer has enough experience to form a stable and accurate valuation of clothes. 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.

A similar mechanism helps shed light on inattention to shrouded product attributes (Gabaix and Laibson 2004), particularly those involving repeated neglect, for example of add-on fees or taxes, which may reflect memory limitations. Chetty, Looney and Kraft (2009) find that displaying full prices, inclusive of sales taxes, in supermarket aisles reduces demand. Crucially, consumers correctly recall sales taxes across a range of products when asked directly about them. This suggests that many consumers do not have sales taxes in mind when choosing goods, even though they regularly pay these taxes at the counter and can recall them accurately when cued appropriately.[14] One reason for this might be that, due to interference, the pre-tax price that the shoppers observe does not cue retrieval of the full price. From Equation (8), the total normal cost $p^n(c_t, p_t)$ retrieved in the context $c_t = aisle$ when the before tax price $p_t$ is observed is given by:

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

[13] A related phenomenon is the anchoring heuristic (KT 1973), whereby judgments of unfamiliar quantities, such as the length of the river Amazon, 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).

[14] This mechanism is distinct from that in Gabaix and Laibson (2004) where consumers learn about products with hidden attributes and sellers compete on the visible price but raise the hidden costs. To take their running example, once a consumer buys an expensive cartridge he remembers it, even when next buying a printer (since it is a sticker shock).

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 or when checking the receipt. Clearly, $\overline{p}$ insufficiently accounts for the tax. But there are two other sources of tax neglect. First, the current price $p_t$ seen on the aisle is also before tax, and cues recall of similarly lower prices. Second, the aisle context also retrieves a before tax average aisle price $\mathbb{E}_F(p|aisle)$. These forces, which create neglect of non-visible sales taxes, are ultimately due to the contextual dissociation between the price seen in the aisle and the full price at the counter. This dissociation biases the database and hinders recall of taxes in the aisle context. This perspective is consistent with a broader literature that suggests that payment methods that decouple buying and paying may 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.

In our model, the projection bias, the attribution bias, the experience effects, and shrouded attributes are not independent phenomena. Instead, they are manifestations of the same memory substrate, reflecting a biased database and a selective retrieval process. In particular, because these effects are a product of memory-based norms, they can be modulated by contextual cues, as we discuss next.

*Reminders and Selective Retrieval*

Instances that are underrepresented in the memory database can still be incorporated in the norm if decision makers are cued to think about them. According to Equation (9), providing a contextual cue $c$ tilts the norm towards the conditionally expected quality $\mathbb{E}_F(q|c)$. In the above example of the amusement park, reminding vacation planners of the possibility of bad weather helps them disproportionately retrieve past experiences of vacationing in bad weather. If such experiences are underrepresented in the norm, a reminder helps de-bias the norm and to dampen the attribution bias. Similarly, reminding pessimistic investors of good times where returns on the market are high may dampen experience effects. The same logic extends to the case of the projection bias, where errors in valuation are due to selected retrieval cued by the current context. In this case, biases could be compensated by providing a cue that retrieves memories suppressed in the current situation, such as reminding the prospective buyer of a convertible about snowy winter days.[15] A similar mechanism may also work for shrouded attributes, where reminding consumers of sales taxes would raise their assessed price.

Contextual cues, aptly referred to as reminders, can have rich effects. First, by reminding people of similar events and interfering with recall of different weather conditions, context can dampen the role of differences in past experiences and cause a convergence of norms. Second, by triggering retrieval of different past experiences, the same cue can enhance disagreement. To see the first effect, consider again the example of the amusement park. Suppose that consumers $i$ and $j$ are identical in all respects except that $i$ has experienced on average better weather at the amusement park. This is captured by a higher marginal frequency of $c_t = good\ weather$ in $i$'s database. Absent cues, the consumers display a valuation difference of $\overline{q}_i - \overline{q}_j > 0$. The cue $c_{gw} = good\ weather$ shifts both consumers' norms by putting the same weight on $\mathbb{E}_F(q|c_{gw})$, dampening their valuation difference to:

$$q_i^n(c_{gw}) - q_j^n(c_{gw}) = \frac{\overline{q}_i - \overline{q}_j}{1 + 2\delta\gamma^2} < \overline{q}_i - \overline{q}_j.$$

This prediction helps distinguish the role of recall from that of coarse encoding, such as in Schwartzstein (2014), in which memories would not encode attributes that are deemed less relevant.

Consider now the opposite effect, whereby providing common contextual information can cause valuations to diverge further. This effect occurs when the cue has a disproportionate impact on the recall of some consumers relative to others. Suppose that consumer $i$ experienced a higher variance of weather than consumer $j$ ($\gamma_i^2 > \gamma_j^2$), so consumer $i$ has more experiences of both good and bad weather than consumer $j$. This implies that, when cued with $c_t = good\ weather$, consumer $i$ retrieves more similar good weather experiences than consumer $j$ and hence forms a higher norm (Equation 9). Even if both consumers agree on

[15] Loewenstein, O'Donoghue and Rabin (2003) model projection bias as a perception of excess persistence in valuation. Such an error would be impervious to reminders about the weather.

the fact that the park is more pleasurable in good weather, namely they agree on the term $\mathbb{E}_F(q|c_{gw})$, the cue can make consumer $i$ much more optimistic than $j$ if quality with good weather is sufficiently high. To see this in the starkest possible way, suppose that consumers $i$ and $j$ have identical average past experiences: $\overline{q}_i = \overline{q}_j = \overline{q}$. From Equation (9) it is immediate to see that the same good weather cue creates divergence among them when:

$$q_i^n(c_{gw}) - q_j^n(c_{gw}) > 0 \Leftrightarrow \mathbb{E}_F(q|c_{gw}) > \overline{q},$$

i.e., provided that good weather is associated with above average quality. Similarly, an instance of poor stock market performance may depress the valuation of an investor with bad experiences much more than that of an investor with good experiences. In this sense, measurable differences in past experiences may help explain markedly different reactions to the same information.

### 4.2 Observed Hedonic Attributes

We now consider a decision maker observing hedonic attributes, $\kappa_t = (q_t, p_t, c_t)$. This cue triggers spontaneous retrieval of price and quality norms $p^n(\kappa_t)$ and $q^n(\kappa_t)$. These memories provide the reference for valuation, shaping surprise and attention. Equation (1) can be restated more precisely as:

**Definition 3.** *Given cue $\kappa_t = (\lambda_{q,t}q_t, \lambda_{p,t}p_t, c_t)$, the utility from good $(q_t, p_t)$ is given by:*

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

To formalize the Weber-Fechner law of perception, we assume that the salience function $\sigma(x, y) \geq 0$ is symmetric, homogeneous of degree zero, and increasing in $x/y$ for $x \geq y > 0$. These assumptions follow our previous work (BGS 2012, 2013) and imply that salience increases in the proportional difference between the attribute and its norm. Equation (10) introduces two significant modifications relative to our prior work. First, salience weights are attached to deviations from the norm, not to attribute levels themselves as in BGS (2013). This formulation captures the intuition – central to Norm Theory and the predictive brain hypothesis – that valuation is a two-stage process of anchoring and adjustment to error, but also preserves the key properties of the original salience model.[16] Second, in Equation (10) separate attention is paid to price and to quality, while in BGS (2013) what matters is the relative salience of the two attributes. This modification captures the intuition that little attention is devoted to price if it is close to the norm, and much attention is devoted to price if it is surprising, regardless of attention paid to quality. For this purpose, we let $\sigma(y, y) = 0$ and $\lim_{x/y \to +\infty} \sigma(x/y, 1) = \sigma > 1$. This ensures that salience can capture both inattention to small errors and (bounded) overreaction to large surprises.

To examine Equation (10), consider the simplest case in which quality is deterministic and correctly perceived at $q$. Here memory and attention only influence price valuation:

$$V\big(p_t, p^n(p_t, c_t)\big) = p^n(p_t, c_t) + \sigma\big(p_t, p^n(p_t, c_t)\big)[p_t - p^n(p_t, c_t)]. \tag{11}$$

Equation (11) is illustrated in Figure 3 below.

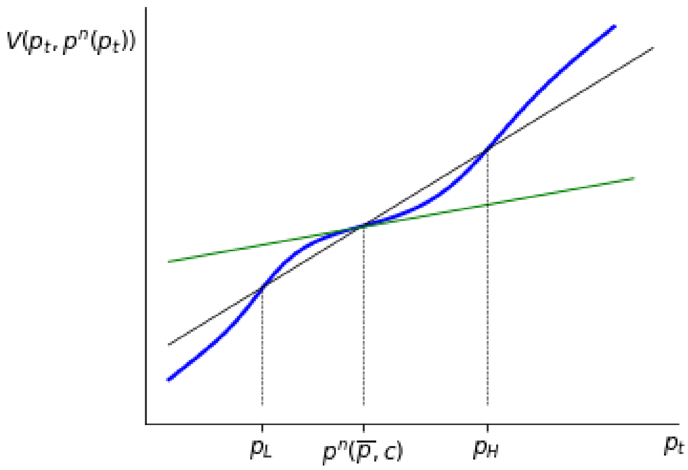


**Figure 3.** The figure plots the utility of price, $p^n + \sigma(p_t, p^n)(p_t - p^n)$ (in blue) and the norm $p^n$ (in green) as a function of observed price $p_t$ when the database is Gaussian and context is uncorrelated with price (Equation 6). The salience function is of the form $\sigma(x, 1)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

The flat line plots the price norm conditional on $p_t$. For price levels around $p^n(\bar{p}, c)$, the norm is close to the current price. Because the discrepancy between the price and the norm is small, error is not salient

[16] One advantage of the current formulation is that valuation is monotonic in price, so that a salient low price always increases valuation. In BGS (2013), by contrast, price salience reduces valuation of all goods, but reduces it more for more expensive goods. We are not aware of evidence favoring one prediction or the other. They are not critical for Salience Theory, but monotonicity creates an intuitive symmetry between the effects of high quality and low price.

and receives little attention. Valuation (blue curve) then only partially adjusts to departures from $p^n(\bar{p}, c)$, and it is flatter than the $45^o$ line (in black). In this case, we say that the current price is assimilated to the recalled norm. 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 relative to past airport experiences.

For price levels far enough from the average, the retrieved norm is far from the actual price. The surprise is big and salient, so the consumer pays a lot of attention to it. Valuation now adjusts sharply, becoming steeper than the $45^o$ line. In this case, we say that the current price is contrasted with the norm. 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 pain of paying, and refuses to buy. This is a standard contrast effect.

The precise shape of valuation depends on the salience function. We offer a characterization below.

**Proposition 3**. *Suppose that the database $F(p, c)$ is Gaussian and the similarity function follows Equation (3). Then for any context $c_t$ such that $p^n(\overline{p}, c_t) > 0$, there exist two price thresholds $p_L, p_H$ satisfying $0 < p_L < p^n(\overline{p}, c_t) < p_H$ such that price is correctly valued if and only If $p_t \in \{p_L, p^n(\overline{p}, c_t), p_H\}$. Furthermore:*

*1) The valuation function of price $V\left(p_t, p^n(p_t, c_t)\right)$ is monotonically increasing in $p_t$. Valuation of price is too high for $p_t < p_L$ and for $p_t \in (p^n(\overline{p}, c_t), p_H)$, and too low for $p_t \in \left(p_L, p^n(\overline{p}, c_t)\right)$ and for $p_t > p_H$.*

*2) If $2\sigma'(x, 1) + \sigma''(x, 1)(x - 1)$ is monotonically decreasing for $x > 1$ and negative at $x = \lim_{p_t \to +\infty} \frac{p_t}{p^n(p_t, c_t)}$, the valuation function has three inflection points. It is convex for $p_t < p_L$, and concave for $p_t > p_H$.*

Property 1) holds for any salience function that satisfies ordering and homogeneity of degree zero. Intuitively, valuation of price increases in the stimulus $p_t$. Most important, the patterns of departures from rational valuation mimic Figure 3. At intermediate prices, discrepancies with the norm are not salient, and valuation is "too flat". That is, both the gains from prices moderately below the norm, and the losses from prices moderately above the norm, are dampened. As price gets far enough from the average, discrepancies with the norm are salient, and valuations of prices become exaggerated. However, due to the diminishing sensitivity of salience, valuation is less and less sensitive to marginally larger surprises and eventually turns parallel to the 45-degree line.[17] As in Proposition 3, the valuation function has exactly three inflection points.

These features of the valuation function relate to assimilation and contrast effects in psychology, and to categorization more broadly. As we discuss in Section 7, a large literature documents that when judging stimuli such as length, loudness, and others, assimilation of the current stimulus to a cued reference prevails when the discrepancy is moderate while contrast occurs when the discrepancy is large (Brigell and Uhlarick 1979, Herr, Sherman and Fazio 1983, Pelham and Wachsmuth 1995). These effects are closely related to evidence on 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. For example, experiments using speech, musical pitch and other stimuli where categories are learned show higher discriminability for stimuli differences that straddle category boundaries (Burns and Ward 1978, Logan, Lively and Pisoni 1991). Two stimuli with the same objective distance are more easily distinguished when they belong to two different categories, but only for subjects who have learned the categories (that is, speakers of the language under test or expert musicians). Goldstone (1994) shows that such effects arise even for the perception of stimuli under arbitrary categories defined in an experiment (in this case, stimuli defined by the brightness and size of anodine objects).

Proposition 3 is broadly consistent with this evidence. First, it defines price categories against the background of a price norm: a ``normal'' price range, $p_t \in (p_L, p_H)$, in which valuation is assimilated to the norm, and ``cheap'' ($p_t < p_L$) and ``expensive'' ($p_t > p_H$) price ranges where differences relative to the norm are exaggerated. Second, price discriminability as measured by the slope of valuation is lowest close to the norm and highest close to category boundaries. Finally, and importantly, the model suggests that the categories themselves (that is, the boundaries $p_L, p_H$) are not fixed, but rather depend on measurable features of the memory database. These properties yield several predictions.

[17] Markowitz (1956) proposed a utility of wealth that, similar to Figure 1, exhibits convexity for intermediate wealth ranges but concavity for extreme wealth. Such a function could help reconcile risk loving and risk averse choices within individuals. Here, the valuation function is not fixed; rather, it is a result of recall and varies with past experience.

First, consider the prediction that valuation is flat for small price changes around the norm. Insensitivity to small price changes is consistent with a large literature in both economics and marketing. Chandra, Handel and Schwartzstein (2018) review a large literature showing that households do not react to changes in the price of their health plans, which are typically small. Marketing scholars have used the term "just noticeable difference" (Monroe 1973) to capture the idea that consumers neglect small price changes. Examining aggregate demand across dozens of retail product categories and different retailers, research in marketing suggests that demand is rigid with respect to small price changes, but shows greater elasticity for larger price changes in either direction (Pauwels et al 2007, 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.

For 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. Selective memory and the valuation function in Figure 3 could help explain this phenomenon. Small price increases are barely noticed, so they allow retailers to increase the price without reducing demand much. Of course, the cumulative cost to the consumer of many neglected price increases can be substantial. At the same time, large temporary sales can harness strong demand, and potentially mask more permanent prices increases.

In economics, a conventional approach to this issue is rational inattention (Sims 2003), or the idea that paying attention to all price changes is too costly (Gabaix 2014). This approach can account for limited sensitivity to prices but has a hard time capturing over-reaction. In our model, the coexistence of an inattention regime for small changes with an overreaction regime for large changes may help explain the rigidity of behavior observed in some economic settings, but also reconcile it with the high elasticity observed in others. Note that in our model, the valuation errors entailed by inattention are naturally bounded because they are confined to small ex post errors.

We now turn to the determinants of the price categories. A central property of our model is that that price sensitivity depends on the accuracy of norms, and in particular on the memory database and the strength of similarity. To show this, we focus on the role of the past distribution of prices and context in shaping the intermediate price range where inattention prevails.

**Proposition 4.** *Denote by* $(\tilde{p}_L, \tilde{p}_H)$ *the range of intermediate prices in which valuation is "too flat", namely where* $\partial_{p_t} V\big(p_t, p^n(p_t, c_t)\big) < 1$. *Our model yields the following comparative statics:*

*1) The range of under-reaction* $(\tilde{p}_L, \tilde{p}_H)$ *moves up and expands for a consumer who has experienced higher prices on average, formally* $\frac{\partial \tilde{p}_L}{\partial \overline{p}} > 0$, $\frac{\partial \tilde{p}_H}{\partial \overline{p}} > 0$, *and* $\frac{\partial (\tilde{p}_H - \tilde{p}_L)}{\partial \overline{p}} > 0$. *The same occurs when the consumer is primed with a context* $c_t$ *associated with higher prices.*

*2) When context is uncorrelated with prices,* $\rho = 0$, *higher volatility of prices* $\pi^2$ *expands the under-reaction range* $(\tilde{p}_L, \tilde{p}_H)$ *symmetrically around* $\overline{p}$, *formally* $\frac{\partial \tilde{p}_L}{\partial \pi^2} < 0, \frac{\partial \tilde{p}_H}{\partial \pi^2} > 0$.

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 again 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 in the consideration of price: the threshold at which a good is perceived as too expensive grows proportionally with the average price level of that good, generating a "latitude of price acceptance" which also grows with price (Koschate-Fischer and Wullner 2017).

The fact that the normal price range is wider for more expensive goods 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, Zeckhauser 1979; see Ratchford 2009 for a review). Price dispersion may arise due to imperfect competition, as a function of search costs (Pratt, Wise, Zeckhauser 1979). Yet both the level of price dispersion, and its increase with price levels, 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.[18]

Proposition 4 also implies that inattention to prices is context-dependent because memory is associative. A cue that induces recall of 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. Once back home, the consumer again adapts to the usual prices, but cues that bring the previous experience to mind jolt him to see the usual prices as being very high. Contextual similarity brings past price norms to mind, switching attention to prices on and off.

Finally, price volatility also expands the range of inattention. Exposure to a higher $\pi^2$ makes our mental prediction less rigid, dampening surprise. As a result, differences between the choice and the norm become more likely to go unnoticed. In models of rational inattention, greater ex-ante price variability should induce the agent to pay more attention to prices. In our model there are instead two conflicting effects. On the one hand, a higher $\pi^2$ reduces prices sensitivity by fostering more flexible norms. On the other hand, a higher $\pi^2$ increases price sensitivity by making it more likely that a surprising price is realized.

In an experiment on price valuation, 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.

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. The marketing literature argues that frequent shallow sales lower consumers' "internal reference price" much more dramatically than

[18] Kahneman and Tversky (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).

do infrequent deep sales (Chang 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. When the regular price is posted, recall of sales prices makes the current price look like a bad deal. When the product is on sale, recall of past sales prices makes the current discount look normal, so it is barely noticed.

We conclude this section with a comparison of the value function in our model, as described in Figure 3 and Proposition 3, and that in Prospect Theory. Taking the norm $p^n(\overline{p}, c_t)$ as the reference point, the Prospect Theory value function: i) is steepest around the reference, ii) features aversion to small price increases, and iii) displays diminishing sensitivity. In our model, properties i) and ii) do not hold. Valuation is flat around the norm, reflecting inattention to small changes, and there is no loss aversion kink at the reference point. If anything, in the space of prices the opposite phenomenon arises. In our model, Weber's law implies that consumers would be more sensitive to a price cut from \$10 to \$5 than to a price increase from \$10 to \$15 (just as they would be more sensitive to a drop than to a commensurate increase in quality). At the same time, our model yields the diminishing sensitivity property iii) outside of the normal price range.

As we showed in BGS (2012), loss aversion may be added to our model, but it is not an integral part of our approach so we abstract from it here. With respect to the steepness of the value function at the origin, our model replaces it with maximal steepness at category boundaries. That is, our valuation function is steepest as the price transitions away from the normal category. As we just saw in Proposition 4, the size of the normal category increases with price variability. As a consequence, valuation is steepest around the origin when prices are highly stable, namely $\pi^2$ is very small. In this special case, our model approximates the Prospect Theory value function with a norm equivalent to the "status quo" price.

## 5. Assimilation and Contrast

Our model helps account for and unify several puzzles on the instability of preferences. The most standard ones arise from background contrast experiments (e.g. Simonson and Tversky 1992). In these experiments, in a first stage subjects are presented with goods characterized by high or low prices. In a second stage, then, subjects choose whether or not to buy similar goods at an intermediate price. The finding is that exposure to high prices in the first stage renders buying the good more attractive in the second stage. The reverse occurs after prior exposure to low prices. Similar effects arise in field evidence from housing choice (Simonsohn and Loewenstein 2006) or context dependent willingness to pay (Thaler 1985).

To capture this 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 $c_t$. By so doing, we assess how the valuation of $p_t$ changes with different retrieved norms. As before, we discuss only the case of price because valuation of quality works in the exact same way. In this setting, the anchoring and adjustment logic implies that contrast effects should arise when the retrieved norm is very different from the observed price $p_t$, so that the latter is surprising and valuation overreacts.

But how can a process of similarity-based recall produce norms starkly different from the observed price? This occurs in two cases. First, this occurs when the observed price $p_t$ is very different from the prices stored in the database, in which case surprise is unavoidable, as when the traveler goes for the first time to the airport. But a second and crucial case is when context cues selective recall of different prices, creating an artificial surprise. Because similarity operates on both intrinsic and contextual attributes, similarity along the contextual dimension may induce recall of dissimilar prices, creating surprise and contrast.

The latter effect is key. In experiments, the first stage populates the database $F(p, c)$ with prices that are different from the second stage ones. The observed second stage price is indeed different from the prices added to the database in the first stage. But there is also strong contextual similarity between the two stages. The goods are similar, the laboratory environment is the same, and the time of the experience is similar. Such similarities induce retrieval of first stage prices, making the 2nd stage price surprising. Absent this contextual but normatively irrelevant similarity, subjects would arguably recall more frequent past experiences, perhaps outside of the laboratory setting. In this precise sense, context is an engine of artificial surprise. This general intuition sheds light on several examples of contrast effect.

To proceed, note that the valuation of price $p_t$ based on a norm $p^n$ is given by:

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

We call assimilation the case in which the valuation of $p_t$ increases in the norm $p^n$. We call contrast the case in which the valuation of $p_t$ decreases in $p^n$. Thus, assimilation arises when:

$$\frac{\partial V(p_t, p^n)}{\partial p^n} = 1 + \frac{\partial \sigma(p_t, p^n)}{\partial p^n} \cdot (p_t - p^n) - \sigma(p_t, p^n) > 0, \tag{12}$$

while contrast arises when the reverse holds. Whether valuation assimilates to the norm or contrasts away from it depends on the anchoring and adjustment mechanism. When the error is small, $p^n(p_t) \approx p_t$, the right-hand side of Equation (12) is positive and assimilation prevails. In this case, little attention is paid to the error and judgment is anchored to the mental representation $p^n$. When instead the error is large, it attracts attention and adjustment is large. Equation (12) turns negative, reflecting a contrast with the norm.

Contrast and assimilation as defined here are intimately related to the normal and extreme price ranges discussed in Section 4.2 in the context of inattention and over-reaction to prices. Assimilation occurs when the norm is close to the actual price, contrast occurs when they are far apart. In fact, the ranges for assimilation and contrast and those for inattention and over-reaction to prices largely overlap.

Price valuation $V(p_t, p^n)$ for a fixed realized price is depicted below, under the assumption of Proposition 3, point 2.

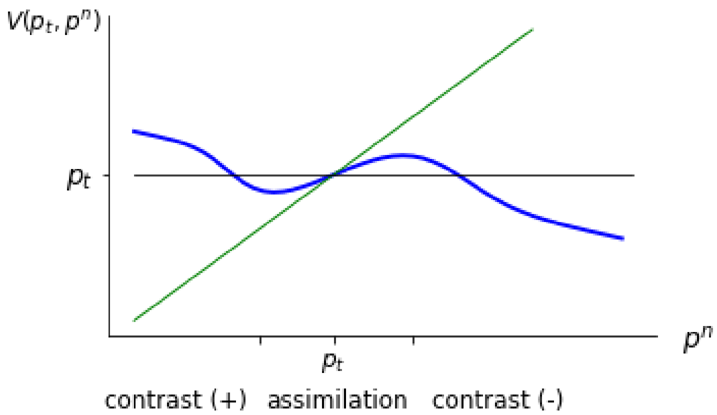


**Figure 4.** The figure plots the utility of price, $p^n + \sigma(p_t, p^n)[p_t - p^n]$ (in blue) for a given observed price $p_t$ as the norm $p^n$ (in green) is exogenously varied, when the database is Gaussian and context is uncorrelated with price. The salience function is of the form $\sigma(x, 1) = \sigma\frac{e^{\theta x}}{1+e^{\theta x}}$, where $\sigma > 1$ and $\theta > 0$. The rational benchmark $p_t$ is shown in black.

Assimilation – an intermediate valuation between the price and the norm – prevails when the norm $p^n$ is sufficiently close to $p_t$. Contrast prevails when the norm is extreme. In the contrast (+) region the price norm is so low that the price looks very high relative to the norm. In the contrast (-) region the norm is so high that the price looks very low relative to the norm.

Consistent with Figure 4, many experiments in psychology show assimilation to moderate cues and contrast to extreme cues. Herr (1989) shows that 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. In price judgments, Herr (1989) shows that priming a consumer with a moderately high car price causes him to judge the price of an unknown car brand as more expensive, creating assimilation. The background contrast effects described before can also be explained using Figure 4. Seeing high prices in the first stage causes a high price norm, so that a moderate price looks cheaper. The reverse effect obtains after seeing low prices in the first stage.

To formalize this mechanism, and to address several choice puzzles, we focus on the willingness to pay $p_{WTP}$ for quality $q$. We investigate how experienced prices $\bar{p}$ and contextual cues $c_t$ shape willingness to pay by shaping norms. According to Equation (8), the extremity of the norm depends on the price $p_t$ itself, and on the past experiences that the choice context brings to mind. The consumer prefers buying the good at price $p_t$ over not buying it (0,0) if and only if:

$$q - p^n(p_t, c_t) - \sigma\big(p_t, p^n(p_t, c_t)\big) \cdot \big(p_t - p^n(p_t, c_t)\big) \geq 0, \tag{13}$$

where the price norm satisfies Equation (8). The consumer's willingness to pay $p_{WTP}$ is the highest price such that Equation (13) holds. Because the valuation of the good decreases in $p_t$, $p_{WTP}$ is implicitly defined by:

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

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. We denote this function by $p_{WTP}(c_t)$.

The consumer has a database of memories $F(p, c)$ of a good. To capture background contrast experiments, we assume that this database contains prices the consumer has seen in the past for this good, but is also shaped by the first experimental stage. That stage is described by a specific experimental context $c_1$, with associated average experienced price $p_1$. Consider two consumers who are identical except that they experienced different first stage prices. One consumer sees a high price $p_{1h}$ while the other sees a low price $p_{1l}$ in the same context. To separately denote the two first stage treatments, we label by $c_{1h}$ the experimental context for the first consumer and $c_{1l}$ the identical context for the other consumer. In the second stage, both consumers are cued by the same experimental context of the first stage. To simplify, we assume that similarity is strong, but the same qualitative result obtains under milder assumptions.

**Proposition 5.** *Under the conditions of Proposition 3 and $\delta \to \infty$, there are two thresholds $p_L^n$ and $p_H^n$, with $p_L^n < q < p_H^n$, such that $p_{WTP}(c_{1h}) > q > p_{WTP}(c_{1l})$ if and only if and $p_{1l} < p_L^n < p_H^n < p_{1h}$.*

Contrast effects such as the background contrast effects require context to cue recall of extreme prices. Stability of context implies that the second stage cues recall of first stage prices, to the exclusion of any other price in the database (recall that $\delta \to \infty$). When first stage prices are high, the resulting price norm is high, so that current prices look "cheap" by comparison. This causes an expansion of the range of prices at which the consumer is willing to buy, as reflected by a higher than rational $p_{WTP}$. By the same logic, when first stage prices are very low, stability in context cues low first stage prices and blocks recall of other prices in the database. Now the second stage norm is low, so that the current price looks "expensive". The range of prices at which the consumer is willing to buy shrinks, as reflected by a lower than rational $p_{WTP}$.

Stability in context and extreme first stage prices are both necessary to obtain this effect. If context is not stable, the consumer samples the entire database $F(p, c)$ and price norms are unrelated to the first stage experiences. If instead first stage prices are not extreme, then the norm retrieved in the second stage is accurate. As a result, the consumer is not surprised and his willingness to pay is close to rational. In fact, for first stage prices in the range $(p_L^n, p_H^n)$, assimilation to memory norms prevails, which causes willingness to pay to *decrease* in first stage prices, opposite to contrast effects.

These results offer a unified explanation of several choice anomalies. We start with Simonsohn and Loewenstein (2006), who present evidence 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, but they spend less than comparable locals when moving to a more expensive city, say from Atlanta to New York. In subsequent renting decisions, however, the movers converge to locals: they switch to cheaper rents in Pittsburgh and to more expensive ones in New York. This behavior is puzzling for the neoclassical model, including one with search costs, but also for models of reference dependence where the reference is the rational expectation of price, since in these models the reference should adjust immediately.

This evidence is however accounted for by Proposition 5. For movers from San Francisco to Pittsburgh, their first stage prices $p_{1h}$ are determined by their past experience in San Francisco. Of course, the second stage in Pittsburgh does not have identical context $c_{1h}$ as the first stage. San Francisco differs from Pittsburgh in many dimensions. Yet, there is an important contextual similarity between the two stages: they both involve an apartment hunt. The context of looking for an apartment in Pittsburgh superficially cues the mover to recall the apartment prices in San Francisco. Because the recalled norm is so much higher than Pittsburgh rents, the latter look surprisingly cheap. Movers, then, have a high willingness to pay for apartment quality $q$, similarly to the experimental subjects who saw high first stage prices.

The role of context also explains the subsequent adaptation. Living in Pittsburgh, the renter's database becomes progressively populated by Pittsburgh prices and by Pittsburgh context that become increasingly available in memory when searching for the next apartment. As a result, if the renter switches to a new apartment within Pittsburgh, contextual similarity with past Pittsburgh experiences brings Pittsburgh prices to mind, inhibiting recall of San Francisco prices. This implies that the price norm of a second time mover drops to the norm of locals. As a result, the willingness to pay for quality $q$ drops relative to the initial move, and the behavior of these households is now indistinguishable from that of long-term locals.

The logic of similarity generates new predictions. If the San Francisco mover at some point lived in a city with similar characteristics and rents as Pittsburgh, say Atlanta, memories of such experience would displace the more recent San Francisco rents. This previous experience would be more similar, both in rents and in context, to the current Pittsburgh experience, interfering with recall of San Francisco. 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 prices close to the destination city are less influenced by city of origin prices.

Context-induced 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. [19] Naturally, the resort cue brings to mind high prices, while the shack cue brings to mind low prices. In the terminology of Proposition 5, the consumer primed with "resort" is cued with a context $c_{1h}$ associated with high prices. The consumer primed with "shack" is instead cued with a context $c_{1l}$ associated with low prices. As a result, the former consumer reports a high willingness to pay while the latter reports a low willingness to pay. Cued prices are extreme, so the consumer contrasts reasonable prices away from them.

In these examples, the contextual cue distorts behavior because it does not reflect the actual choice situation: renting is not in San Francisco but in Pittsburgh, and the beer is not drunk at the resort but on the beach. Critically, similarity-based recall is automatic and as such does not distinguish between superficial and fundamental similarity. These experiments highlight a perverse role of context: superficial contextual similarity can create artificial surprise and distort decisions.[20]

---

[19] 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.

[20] Contextual similarity generates normatively irrelevant experience effects, but it also circumscribes them. For instance, the experience of high prices in an experiment should not affect subjects' behavior after the experiment, because the experimental context is so different from a naturalistic one.

But the very same logic implies that, as discussed in Section 3, contextual similarity can also have a beneficial effect: it typically helps us form accurate norms and make better decisions. This is the case also in the settings we just discussed. Context allows movers to eventually adapt to the behavior of locals. Even in the Thaler beer experiment, context can foster adaptation. For the consumer deciding how much to pay to drink his beer not at the beach but at the resort itself, recalling the context-incongruent shack prices would draw attention to the higher resort prices. This could prevent him from enjoying the comfortable atmosphere and the attentive service of the resort. In this case, interference of the resort location with the recall of past low prices reduces surprise, improving decisions.

We conclude this Section by discussing cases in which surprise is big enough that valuation overshoots, namely $\sigma(x, x^n) > 1$. In psychology, cases of overshooting are called "aftereffects" and explain why after being exposed to high levels of stimuli, such as size or temperature, we systematically underestimate the strength of a given target stimulus, and the reverse if exposed to low levels of stimuli (Brigell and Uhlarick 1979). The mere existence of contrast effects – the phenomenon that the valuation of an available attribute is contrasted away from past experiences -- does not require overshooting. It only requires that $\sigma(x, x^n)$ be steep enough with respect to the norm.

However, some of the phenomena we considered do rely on overshooting. Take for instance the Simonsohn and Loewenstein evidence on movers. Here some findings do not require overshooting, others do. For instance, the finding that movers coming from more expensive cities are willing to pay more for rent than movers coming from cheaper cities does not rely on overshooting. It only requires that $\sigma(p, p^n)$ is steep enough with respect to the price norm. Overshooting is however required to explain why movers from expensive cities are willing to pay more than locals. The latter are arguably well adapted to local prices, in the sense that they perceive them correctly. As a result, explaining this evidence amounts to requiring that movers from, say, San Francisco undervalue Pittsburgh prices in absolute terms. That is:

$$V(p_t, p^n) \equiv p^n + \sigma(p_t, p^n) \cdot (p_t - p^n) < p_t.$$

When the movers come from an expensive city $p^n > p_t$, this condition is equivalent to $\sigma(p, p^n) > 1$. Overshooting is also necessary in the case of "misleading sales", defined as sales strategies whereby regular prices are highly inflated in order to lure consumers into buying during artificial sales events (BGS 2013).

Another dramatic instance of overshooting, this time in the quality space, is 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 with 11 intact ones and two 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 quality norm, no dish is broken because we never experience buying broken dishes. The consumer is very negatively surprised by the broken dishes. His valuation overshoots, and falls below the rational valuation of a set of 10 intact dishes, for which there is no negative surprise. [21]

In sum, this Section makes two points. First, the structure of Equation (10) helps account for the role of norms in creating assimilation and contrast effects in evaluating perceptual stimuli. Second, associative memory puts structure on the norms that are used to evaluate these stimuli in different cases. In particular, contextual similarity sometimes superficially activates poor representations which create artificial surprise, and thus distort decisions. Several choice anomalies can be explained in a unified way with this mechanism.

## 6. Norms, Fairness and Social Preferences

A large body of work challenges the assumption of individual self-interest by showing that fairness considerations influences economic decisions in the field and the lab (see Fehr and Schmidt 2006). These studies often rely on two ingredients: a "fair" allocation, and "social preferences" that make departures from that allocation personally costly. Several models of social preferences have been developed. For instance, Fehr and Schmidt (1999) assume that an agent is altruistic toward players whose payoff is below the fair

[21] 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 $\sigma(q_{11}, q_{13}) > 1$.

allocation and envious about the others. Bolton and Ockenfels (2000) assume that an agent's utility is maximal when he receives his "fair" payoff share.[22]

Less attention has been devoted to the question of where the fair allocation itself comes from. Often, it is assumed to be equal split of payoffs, but this is conceptually and empirically unsatisfactory. Here we follow a different approach, which connects fairness norms to memory. We already showed some examples of this logic in the context of "fair prices". The first-time traveler at the airport may be shocked, and refuse to buy, because he deems the high price of water unfair. But why does the high airport price become fair with experience, so that the seasoned traveler is willing to buy? Here we argue that progress on this and other questions can be made by viewing fairness norms as coming from memory, because this approach can help explain different, at times puzzling, instances of fairness judgments and behavior. We stress that our focus here is on the determination of the fair allocation. We have nothing new to say on how social preferences attach utilities to departures from such a transaction.

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", described as a norm. The authors state "the reference transaction provides a basis for fairness judgments because it is normal, not necessarily because it is just." That is, the fair allocation does not necessarily follow from abstract principles. It conforms to a norm based on experience. We would say that these norms are retrieved from memory, they largely reflect custom, they are malleable, and they adapt to change. As KKT put it, "terms of exchange that are initially seen as unfair may in time acquire the status of a reference transaction (...) at least in the sense that alternatives to it no longer come to mind." Here fairness and availability in memory nearly coincide.

To see the implications of our model for these issues, consider the following KKT experiment on the instability in judgments of fairness. The authors describe an employee who has worked at a small shop for six months and earns $9 an hour. There is some unemployment in the area so similar shops are hiring reliable

[22] In Rabin's (1993) model of reciprocity, fair behavior of a player is defined with respect to a counterfactual action the same player may have taken. An unequal split is not considered unfair unless a more equitable split is easily obtained. Charness and Rabin (2002) offer experimental evidence consistent with this view.

workers for $7 an hour. Survey respondents are asked to judge the fairness of the firm's actions in two scenarios. In scenario 1, the firm cuts the employee's wage down to the market level. In scenario 2, the employee quits and the firm hires a new worker for $7. A vast majority of respondents find the first scenario unfair, while the same respondents find the second scenario to be fair. A small modification of the problem leads to drastically different fairness judgements. Why?

We think that memory is at work in the KKT experiments. Scenario 1 induces subjects to selectively recall similar past instances of employed workers facing fluctuations in the market wage. Because these memories mostly feature wages that are downward rigid, recall generates a norm for transactions with an employed worker in which wages do not drop. Subjects then see a wage reduction as unfair. By equating normality with fairness as in KKT, a violation of the norm that harms the worker is perceived as unfair. By looking at Scenario 1 in isolation one may think that subjects' notion of fairness is based on distributional justice, favoring an equitable distribution of the surplus. Scenario 2 shows that this is not the case. Here, letting the firm reap the entire surplus vis-à-vis a new entrant is judged to be fair. Again, it is easy to make sense of this with memory: Scenario 2 causes subjects to recall similar past experiences of new hires and their wages. The normal wage for new hires is the market wage. Again, equating fairness with normality as in KKT implies that it is fair for the firm to only pay the new hire the going market wage.[23]

Direct evidence on the role of memory in judgments of fairness is provided by the Herz and Taubinsky's (2017) ultimatum game experiment. In the first stage of the experiment, the authors randomize subjects across treatments in which they receive high versus low offers.[24] In the second stage, all subjects play standard (one-on-one) ultimatum games. They find that receivers exposed to the high offers in the first stage reject more generous offers in the second stage, compared to receivers who were exposed to low first

[23] KKT discuss a third scenario in which a worker moves to a new task in the firm, for which pay is 7$ rather than 9$. Also in this case, paying the worker the 7$ wage is viewed as fair. Similarity based recall can account for this evidence: task-similarity primes subjects to recall the 7$ wage, deeming it fair. This is akin to judging higher prices fair at resorts.

[24] In the ultimatum game, a proposer is given an endowment $e$ and offers an allocation $(e - x, x)$ to a receiver, who chooses whether to accept or reject it. If the offer is rejected, both players get $0$. The literature robustly finds that small offers ($x < 0.3e$) are often rejected and that, in turn, proposers tend to offer larger shares. In Herz and Taubinsky, the first (priming) stage consists in subjects playing 15 games of either a proposer-competition or a receiver-competition version of the ultimatum game. Consistent with previous work, receivers get a significantly higher share of the endowment when proposers compete (80% vs 30%).

stage offers. Again, we have nothing to say about the social preferences causing subjects to demand a fair split. The point, though, is that what constitutes a fair split depends on retrieved memories from the first experimental stage, as our theory of recalled norms would predict. The connection of the Herz and Taubinsky experiment with our analysis is even deeper. The experimental malleability of fairness judgments is likely due to the same reason why in background contrast experiments first stage prices influence second stage norms: contextual similarity across experimental stages.

In our approach, fairness norms are unstable because they do not reflect universal welfare principles. They are derived on the fly, through memory of past experiences. One implication, already hinted at in KKT's example, is that the practice of certain customs renders them self-reinforcing. Formally, consider a price $p$ that is normally distributed with context $c$ as in Proposition 2. Price $p$ may capture the wage of a worker and $c$ the market wage. Alternatively, $p$ may stand for the price of snow shovels and $c$ for weather (e.g., whether there is a snow storm), or $p$ may stand for the price of flowers and $c$ for a date, such as Valentine's Day. Equation (7) shows how the norm $p^n(c_t, p_t)$ conditions on expected price given context, $\mathbb{E}_F(p|c_t)$. It follows that with a change in context the norm changes as:

$$\frac{\partial p^n(c_t, p_t)}{\partial c_t} \propto 2\delta\rho\pi\gamma.$$

When context changes, a price or wage change is considered fair if: i) similarity influences recall, $\delta > 0$, and ii) price and context have been historically associated, as captured by the correlation parameter $\rho > 0$. The role of similarity is key. When consumers see a higher context $c_t$ (say, Valentine's Day) they recall *specific* instances characterized by high context and price. This drives up the price norm $p^n$ which is accepted as fair.

By this account, raising the price of snow shovels during a storm is considered unfair due to the very custom of not raising that price, $\rho = 0$. Deviating from the constant price norm causes negative surprise and a refusal to buy, which makes the custom itself self-reinforcing. In other settings, it is instead considered fair to raise prices with demand. It is for instance well accepted that flowers are more expensive on Valentine's

Day.[25] Even here the self-reinforcing memory mechanism plays an important role. On Valentine's day, people recall high prices of flowers because of their past Valentine's Day experiences. Likewise, while firms' reluctance to cut wages (Bewley 2007) justifies the unfairness of this practice, firms routinely cut bonuses and other benefits, which is not viewed as unfair. The fact that customs are, at least in part, self-reinforcing implies that they may be actively changed. But doing so takes time. The memory database of consumers must adapt to new practice, until the force of similarity renders the new custom acceptable.

Another instance of instability of fairness norms comes from the experiments on the dictator game. The literature has robustly shown that small modifications in the presentation of the problem can have large implications for behavior even if they are payoff-neutral (List 2007, Krupka Weber 2013). In particular, Krupka and Weber (2013) run experiments on two variants of the Dictator game. In one variant, player 1 is given a total of \$10 to share with player 2; in the other variant, both players are given \$5, and player 1 chooses whether to take up to, or give up to, \$5. The two versions are strategically equivalent. Yet in the experiments player 2 ends up with more in the second variant than in the first, and is significantly more likely to end up with \$5 in the second treatment.

Krupka and Weber interpret the evidence by saying that the first variant brings to mind a sharing frame whereas the second brings to mind a taking frame, which is socially sanctioned. This interpretation feels right and accords with memory-based norms. In our model, the context of the second variant (taking) filters in the memory database real world situations in which there is a choice of how much to take. Because taking is often discouraged or unlawful, this context mechanically evokes a low norm, encouraging player 1 to be content with a payoff closer to the initial endowment. The sharing variant evokes giving, which brings to mind situations in which something, but not a lot, is given. Behavior then changes across treatments because the description of the problem cues retrieval of different representations of choice.

[25] In terms of first principles, one may even argue that since snow shovels can be purchased in advance by consumers while Valentines' day flowers cannot, it would be fairer to have inflexible flower rather than snow shovel prices.

**7. Related Literature**

Here we connect several concepts in the literature to our anchoring and adjustment model of valuation. We consider previous work on memory in economics models, as well as other approaches to judgment, including assimilation, contrast, and categorization effects in psychology, as well as limited attention and efficient coding in economics.

We start with a discussion of previous work on memory in economics. In case-based decision theory (CBDT, Gilboa and Schmeidler 1995), decision makers assess risky choices according to their similarity with past choices. The role of similarity is related to our model but there are two main differences. First, in CBDT similarity is characterized axiomatically, not psychologically. For example, CBDT does not allow for contextual attributes to influence recall, a crucial feature both of the standard model of memory and of our approach, and one that is critical for explaining a lot of data. Second, CBDT does not consider the allocation of attention on the basis of surprise. In our terminology, in CBDT assessment is anchored to memories, but there are no reference effects.[26]

Recent work takes a more psychological approach. In Mullainathan (2002), limited memory distorts Bayesian updating and forecasting of an economic variable. Similarity influences recall, but there is no role either for context or for interference. Taubinsky (2014) studies optimal reminders in a model where memory is imperfect and mental rehearsal promotes recall. Ericson (2016) studies the interaction of forgetting and procrastination, drawing implications for the demand for reminders. Bushong and Gagnon-Bartsch (2019) consider the distorted encoding (rather than the selective retrieval) of experiences and derive implications for biases in valuation. Following adaptation level theory, the marketing literature discusses the reference role of past choice, but these models focus on recency effects (Cheng and Monroe 2013). Nagel and Xu (2018) model expectation formation under limited memory, where past data is forgotten. Azeredo da Silveira

[26] Other papers explore optimal storage of information (e.g., Dow 1991, Wilson 2014) or decision problems with exogenous imperfect recall (Piccione and Rubinstein 1997). Rubinstein (1998) summarizes some of this work.

and Woodford (2019) derive optimal recall under limited memory in a dynamic setting. These models feature recency effects but abstract from similarity and interference.

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 selective toward features that are most diagnostic of, or representative of, a group relative to a comparison group (Gennaioli and Shleifer 2010, Bordalo, Coffman, Gennaioli, Shleifer 2016, Bordalo et al 2018). While these effects could also be at work in the recall of norms, our approach, like Norm Theory, abstracts from them. It is an open question whether the representativeness heuristic and its effect on beliefs are also a product of similarity based recall.

As discussed in the introduction, our concept of adjustment relative to an anchor has antecedents in the Psychology and Neuroscience literature. The fact that adjustment can be incomplete or excessive is consistent with extensive evidence in psychology on "assimilation" vs. "contrast" effects in valuation (Schwarz and Bless 1988, Stapel and Sulz 2011). In assimilation effects, judgments move in the direction of a prime, or anchor. In contrast effects, the reverse happens. Starting with Sherif, Taub and Hovard (1958) many experiments document these effects for judgments of size, brightness, weight or hue (Brigell and Uhlarick 1979, Herr, Sherman and Fazio 1983, Pelham and Wachsmuth 1995).

It has also been previously shown that, as in our model, memory plays a role in creating assimilation and contrast effects (Bless and Schwarz 2010). The conventional explanation for these effects is that assimilation occurs when the target of judgment is perceived to belong to the same category of the primed comparison stimulus while contrast arises because of sharp categorical differentiation between them. More broadly, the literature on categorization suggests that categories coarsen valuation but also distort it, by emphasizing perceived differences across categories and dampening them within categories (Goldstone 1994, Mullanaithan 2000). The mechanisms underlying the enhanced discriminability at category boundaries are not fully understood. One possibility is that attention is differentially allocated to stimulus levels that straddle categories, or to dimensions along which the categories differ (Goldstone 1994).

A growing literature in Economics has modelled the allocation of attention in choice. As discussed in Section 4.2 (see Figure 3), the salience model described here yields a valuation of price that is to some extent categorical, in the sense that realizations close to the norm are assimilated to it (they fall in the ``normal'' category), and realizations far from the norm are contrasted to it (they fall in the ``expensive'' or ``cheap" categories). The valuation of price is the steepest in the transition between categories. In addition to salience theory (BGS 2012, 2013), Koszegi and Szeidl (2012) and Bushong, Rabin and Schwartzstein (2017) also explore the allocation of attention in response to the choice data. Relative to the salience model, they focus on the role of the range of attributes, and in particular on whether larger attribute ranges attract or dampen attention. Because these models do not feature a reference, or normal, price it is difficult to reconcile them with the logic of price norms, and with over- or underreaction to departures from the norm. Furthermore, while these models do not consider memory, our analysis indicates that recall can play a central role in providing a starting point for models of allocation of attention.

An alternative approach in economics focuses on the optimal representation of choice stimuli given limited attention or information processing capabilities. Gabaix (2014) builds a model in which decision makers form a sparse representation of their environment using only a subset of the data. This model admits a representation similar to Equation (1) in which a prior about attributes is adjusted using attention weights that do not have the proportional structure of salience and that in the main specification are optimally chosen ex ante. Woodford (2012) builds a model where the sensory system is noisy, and the cognitive apparatus filters the noise optimally. This model also admits a representation in which the prior about attributes is adjusted when a choice option is observed, with a weight equal to the signal to noise ratio of the sensory system. Frydman and Jin (2018) extend this approach by examining the efficient coding hypothesis, whereby cognitive resources are optimally allocated such that perception can differentiate among stimuli that are expected to occur with high probability. Similarly to Prospect Theory, these models predict that valuation is steepest around the prior.

Our model departs from this work in three ways. First, it endogenizes the initial mental prediction building on a well-documented model of associative memory. In particular, the norm is generated ex post by similarity to the stimulus, which implies that the error in the representation is contained. Second, relative to models of inattentive choice, which rationalize rigidity, our model also captures over-reaction to surprise. Over-reaction in our framework is due to a reallocation of attention to surprising data, which cannot be explained using initial mental representations. Third, and related, our model entails a valuation function that is flatter around the norm but exaggerates the differences from the norm for stimuli that are sufficiently different from it. Efficient coding is also able to produce overreaction in the sense that a stimulus is contrasted away from the prior, unlike in the standard Bayesian noisy perception framework (Wei and Stocker 2015, Polania Woodford and Ruff 2019). In contrast to our model, such "anti-Bayesian" behavior typically arises for stimuli close to the norm, where the variation in prior probability is steepest, while the standard Bayesian behavior of assimilation to the prior holds for extreme stimuli. The precise relationship between efficient coding and our model needs to be investigated further, and it also requires consideration of a broader class of stimulus distributions (here we only considered Gaussian stimuli). More broadly, the role of efficient coding in providing a foundation for attention and perception remains an open question.

**8. Conclusion**

In this paper, we propose 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 central feature of the model – recall through similarity – yields many predictions on what comes to mind when decision makers face a stimulus, which have been extensively tested and confirmed in memory research but which also have multiple implications for economic analysis. Due to the central role of similarity in recall, norms tend to be adapted to the specific choice at hand, creating the stability (and even rationality) of choice that is faced often. At the same time, norms can also incorporate normatively irrelevant contextual features, and through this channel lead to unstable and apparently irrational choice. The model sheds light on conditions under which surprise leads to assimilation or contrast, and provides a unified account of a range of evidence documented in the literature, as well as making several novel and testable predictions.

Throughout this paper, we have made a number of specific modeling choices for clarity, many of which can be revisited or relaxed. There are several missing aspects in the basic model of memory, such as the importance of salient memories, the inattention to some aspects of the initial stimulus that may influence recall (Schwartzstein 2014), or even the failure of initial encoding of some experiences. In addition, with some modifications, our model can perhaps also incorporate recall of other types of information from memory, such as goals or information about future events. Similarly, the recall cued by a choice option is not necessarily restricted to instances of the same good, but can include alternative goods or alternative uses of funds, which can shape how decision-makers frame their own decisions.

Our model draws a link between memory-based norms and expectations of value, and this raises two important issues. First, in more complex tasks such as predicting the next realization of a stochastic process, decision makers do not just recall past realizations, but also use forward looking intuitions about the data generating process. To capture this aspect requires broadening the model to include semantic memory. Second, the representativeness heuristic has been shown to be an important driver of expectations in macroeconomic, financial, and social domains. Bordalo et al. (2018) use experimental methods to show that this heuristic influences probability judgments by affecting recall, but further work is necessary to connect the similarity model of recall with those findings. 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.

**References**


Anderson, John and Lynne Reder. 1999. "The Fan Effect: New Results and New Theories," *Journal of Experimental Psychology* 128 (2): 186 – 197.

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

Barberis, Nicholas and Ming Huang. 2001. "Mental Accounting, Loss Aversion, and Individual Stock Returns." *Journal of Finance* 56(4): 1247–1292.

Barberis, Nicholas and Wei Xiong. 2009. "What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation." *Journal of Finance* 64(2): 751–784.

Bell, David. 1985. "Disappointment in Decision Making Under Uncertainty." *Operations Research* 33(1): 1—27.

Bewley, Truman. 2007. "Fairness, Reciprocity, and Wage Rigidity," in *Behavioral Economics and its Applications*, Peter Diamond and Hannu Vartiainen Editors, Princeton University Press, Princeton NJ.

Bless, Herbert and Norbert Schwarz. 2010. "Mental Construal and the Emergence of Assimilation and Contrast Effects: The Inclusion/Exclusion Model," *Advances in Experimental Social Psychology* 42: 319 – 373.

Bolton, Gary and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review* 90(1): 166-193.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *Quarterly Journal of Economics* 131(4): 1753 -- 1794.

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

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

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

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2019. "Memory and Reference Prices: an Application to Rental Choice," *American Economic Review Papers and Proceedings, forthcoming*.

Brigell, Mitchell and John Uhlarick. 1979. "The Relational Determination of Length Illusions and Length Aftereffects" *Perception* 8 (2): 187 – 197.

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

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

Busse Meghan, Devin Pope, Jaren Pope and Jorge Silva-Risso. 2015. "The Psychological Effect of Weather on Car Purchases," *Quarterly Journal of Economics* 130 (1): 371 – 414.

Casado, Esteban and Juan-Carlos Ferrer. 2013. "Consumer Price Sensitivity in the Retail Industry: Latitude of Acceptance with Heterogeneous Demand," *European Journal of Operational Research* 228 (2): 418 – 426.

Chandra, Amitabh, Benjamin Handel, and Joshua Schwartzstein. 2018. "Behavioral Economics and Health-Care Markets," working paper.

Chang, Tom, Wei Huang and Yongxiang Wang. 2018. "Something in the Air: Pollution and the Demand for Health Insurance," *Review of Economic Studies* 85 (3): 1609–1634.

Charness, Gary and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics* 117 (3): 817–869.

Chartrand, Tanya, Joel Huber, Baba Shiv, and Robin Tanner. 2008. "Nonconscious Goals and Consumer Choice," *Journal of Consumer Research* 35(2): 189–201.

Chatterjee, Promothesh and Randall Rose. 2011. "Do Payment Mechanisms Change the Way Consumers Perceive Products?" *Journal of Consumer Research* 38 (6): 1129–1139.

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

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

Clark, Andy. 2013. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science," *Behavioral and Brain Sciences* 36 (3): 181 – 204.

Conlin, Michael, Ted O'Donoghue, and Timothy Vogelsang. 2007. "Projection Bias in Catalogue Orders," *American Economic Review* 97 (4): 1217 – 1249.

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

DellaVigna, Stefano, Attila Lindner, Balazs Reizer, and Johannes Schmieder. 2017. "Reference-Dependent Job Search: Evidence from Hungary," *Quarterly Journal of Economics* 132 (4): 1969 – 2018.

Desimone, Robert and John Duncan. 1995. "Neural Mechanisms of Selective Visual Attention," *Annual Review of Neuroscience* 18: 193 – 222.

Dickson, Peter and Alan Sawyer. 1990. "The Price Knowledge and Search of Supermarket Shoppers," *Journal of Marketing* 54 (3): 42—53.

Dow, James. 1991. "Search Decisions with Limited Memory," *Review of Economic Studies* 58(1): 1-14.

Ericson, Keith. 2016. "On the Interaction of Memory and Procrastination: Implications for Reminders,Deadlines and Empirical Estimation," *Journal of the European Economic Association,* forthcoming.

Fehr, Ernst and Klaus Schmidt. 2006. "The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories," *Handbook of the Economics of Giving, Altruism and Reciprocity,* Serge-Christophe Kolm, Jean Mercier Ythier Editors.

Fehr, Ernst and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics* 114 (3): 817 – 868.

Friston, Karl. 2010. "The Free-Energy Principle: A Unified Brain Theory?" *Nature Reviews Neuroscience* 11: 127–138.

Frydman, Cary and Lawrence Jin. 2018. "Efficient Coding and Risky Choice," working paper.

Fryer, Roland and Matthew Jackson. 2008. "A Categorical Model of Cognition and Biased Decision Making," *The B.E. Journal of Theoretical Economics* 8 (1): article 6.

Gabaix, Xavier. 2014. "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics* 129 (4): 1661-1710.

Gabaix, Xavier and David Laibson. 2004. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics* 121 (2): 505–540.

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

Gilboa, Itzhak and David Schmeidler. 1995. “Case-Based Decision Theory," *Quarterly Journal of Economics* 110(3): 605 -- 639.

Gul, Faruk. 1991. “A Theory of Disappointment Aversion,” *Econometrica* 59 (3): 667-686.

Haggag, Kareem, Devin Pope, Kinsey Bryant-Lees, Maarten Bos. 2018. “Attribution Bias in Consumer Choice,” *Review of Economic Studies*, forthcoming.

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

Herr, Paul. 1989. “Priming Price: Prior Knowledge and Context Effects,” *Journal of Consumer Research* 16(1): 67–75.

Herr, Paul, Steven Sherman, and Russell Fazio. 1983. “On the Consequences of Priming: Assimilation and Contrast Effects,” *Journal of Experimental Social Psychology* 19 (4): 323-340.

Herz, Holger and Dmitry Taubinsky. 2017. “What Makes a Price Fair? An Experimental Study of Transaction Experience and Endogenous Fairness Views,” *Journal of the European Economic Association* 16 (2): 316–352.

Hsee, Christopher. 1996. “The Evaluability Hypothesis: An Explanation for Preference Reversals between Joint and Separate Evaluations of Alternatives,” *Organizational Behavior and Human Decision Processes* 67 (3): 247-257.

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

Kahneman, Daniel. 2003. “Maps of Bounded Rationality: Psychology for Behavioral Economics,” *American Economic Review* 93 (5): 1449 – 1475.

Kahneman, Daniel, Jack Knetsch and Richard Thaler. 1986. “Fairness as a Constraint on Profit Seeking: Entitlements in the Market,” *American Economic Review* 76 (4): 728-741.

Kahneman, Daniel and Dale Miller. 1986. “Norm Theory: Comparing Reality to its Alternatives,” *Psychological Review* 93(2): 136-153.

Kahneman, Daniel, Barbara Fredrickson, Charles Schreiber, and Donald Redelmeier. 1993. “When More Pain Is Preferred to Less: Adding a Better End,” *Psychological Science* 4(6): 401-405.

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

Kahneman, Daniel and Amos Tversky. 1979. “Prospect Theory: an Analysis of Decision under Risk," *Econometrica* 47 (2): 263 -- 292.

Koschate-Fischer, Nicole and Katharina Wullner. 2017. “New Developments in Behavioral Pricing Research,” *Journal of Business Economics* 87 (6): 809–875.

Koszegi, Botond and Matthew Rabin. 2006. “A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121 (4): 1133 -- 1165.

Koszegi, Botond and Adam Szeidl. 2012. “A Theory of Focusing in Economic Choice,” *Quarterly Journal of Economics* 128 (1): 53-104.

Krupka, Erin and Roberto Weber. 2013. “Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?” *Journal of the European Economic Association* 11 (3): 495–52.

List, John. 2007. “On the Interpretation of Giving in Dictator Games,” *Journal of Political Economy* 115(3): 482-494.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection Bias in Predicting Future Utility," *Quarterly Journal of Economics* 118 (4): 1209–1248.

Malmendier, Ulrike and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?" *Quarterly Journal of Economics* 126(1): 373 -- 416.

Malmendier, Ulrike and Stefan Nagel. 2016. "Learning from Inflation Experiences," *Quarterly Journal of Economics* 131(1): 53-87.

Markowitz, Harry. 1952. "The Utility of Wealth," *Journal of Political Economy* 60 (2): 151 – 158.

Monroe, Kent. 1973. "Buyers' Subjective Perceptions of Price," *Journal of Marketing Research* 10 (1): 70-80.

Mullainathan, Sendhil. 2000. "Thinking through Categories," working paper.

Mullainathan, Sendhil. 2002. "Memory-Based Model of Bounded Rationality," *Quarterly Journal of Economics* 117 (3): 735-774.

Mussweiler, Thomas and Fritz Strack 2001. "Considering the Impossible: Explaining the Effects of Implausible Anchors," *Social Cognition* 19 (2): 145 – 160.

Nagel, Stefan and Zhengyang Xu. 2018. "Asset Pricing with Fading Memory," working paper.

Nakamura Emi and Jon Steinsson. 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models," *Quarterly Journal of Economics* 123 (4): 1415–1464.

Pantelis, Peter, Marieke Van Vugt, Robert Sekuler, Hugh Wilson, and Michael Kahana. 2008. "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 (6): 1182-1195.

Pauwels, Kown, Shuba Srinivasan, Philip Franses. 2007. "When Do Price Thresholds Matter in Retail Categories?" *Marketing Science* 26 (1): 83 –100.

Pelham, Brett and Jeff Wachsmuth. 1995. ""The Waxing and Waning of the Social Self: Assimilation and Contrast in Social Comparison," *Journal of Personality and Social Psychology* 69(5): 825–838.

Piccione, Michele, and Ariel Rubinstein. 1997. "On the Interpretation of Decision Problems with Imperfect Recall," *Games and Economic Behavior* 20(1): 3-24.

Pratt, John, David Wise and Richard Zeckhauser. 1979. "Price Differences in Almost Competitive Markets," *Quarterly Journal of Economics* 93 (2): 189–211.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 83 (5): 1281-1302.

Rubinstein, Ariel. 1998. *Modeling bounded rationality*. MIT press, Cambridge MA.

Schwarz, Norbert and Bless. 1992. "Constructing Reality and its Alternatives: Assimilation and Contrast Effects in Social Judgment," In *The Construction of Social Judgments,* L. Martin and A. Tesser editors. Eribaum, Hillsdale NJ.

Schwartzstein, Joshua. 2014. "Selective Attention and Learning," *Journal of the European Economic Association* 12 (6): 1423–1452.

Sherif, Muzafer, Taub, Daniel and Hovland, Carl. 1958. "Assimilation and Contrast Effects of Anchoring Stimuli on Judgments" *Journal of Experimental Psychology,* 55(2): 150-155.

Simonsohn, Uri and George Loewenstein, 2006. "Mistake #37: The Effect of Previously Encountered Prices on Current Housing Demand," *Economic Journal* (116)508: 175-199.

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

Sims, Christopher. 2003. “Implications of Rational Inattention,” *Journal of Monetary Economics*, 50(3), 665–690.

Soman, Dilip and John Gourville. 2001. “Transaction Decoupling: How Price Bundling Affects the Decision to Consumer,” *Journal of Marketing Research* 38 (1): 30 – 44.

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

Taubinsky, Dmitry. 2014. “From Intentions to Actions: A Model and Experimental Evidence of Inattentive Choice," Mimeo, Harvard University.

Thaler, Richard. 1985. “Mental Accounting and Consumer Choice,” *Marketing Science*, 4 (3): 199 – 214.

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

Tversky, Amos and Daniel Kahneman. 1973. “Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology* 5(2): 207-232.

Tversky, Amos. 1977. “Features of Similarity,” *Psychological Review* 84(4): 327-352.

Niedrich, Ronald, Subhash Sharma and Douglas Wedell. 2001. “Reference Price and Price Perceptions: A Comparison of Alternative Models,” *Journal of Consumer Research* 28 (3): 339-354.

Wilson, Andrea. 2014. "Bounded memory and biases in information processing," *Econometrica* 82(6): 2257-2294.

Woodford, Michael. 2012. “Prospect Theory as Efficient Perception Distortion,” *American Economic Review* 102 (3): 41 – 46.

## Appendix A. Proofs

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

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

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

$$= \overline{p} + A(c_t) \cdot cov[p, Z(c_t, p)],$$

where $Z(c_t, p) = \mathbb{E}[S(|c - c_t|)|p]$ is the average similarity conditional on price, and $A(c_t) > 0$ depends on the context distribution and the similarity function. If context and price are uncorrelated, $Z(c_t, p) = Z(c_t)$, so that $p^n(c_t) = \overline{p}$.

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, $dF(c|p) = dF(c)$, 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:

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

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

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

$$= \int_0^{+\infty} u\tilde{S}(u; p_t)[dF(p_t - u) - dF(p_t + u)] \geq 0. \qquad (A.2)$$

The distribution $dF(p)$ is symmetric and unimodal around $\overline{p}$. As a result, 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 $p_t \geq \overline{p}$.

Consider now the claim that $p^n(p_t) \geq \overline{p}$. This condition can be written as:

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

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

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

$$= \int_0^{+\infty} u\big[\tilde{S}(|\overline{p} - p_t + u|; p_t) - \tilde{S}(|\overline{p} - p_t - u|; p_t)\big]dF(p_t + u) \geq 0. \qquad (A.2)$$

The function $\tilde{S}(x; p_t)$ decreases in $x$ for $x \geq 0$. As a result, Equation (A.2) holds if $|\overline{p} - p_t + u| \leq |\overline{p} - p_t - u|$ for any $u \geq 0$. This is true if and only if $p_t \geq \overline{p}$. Thus, we either have $p_t \geq p^n(p_t) \geq \overline{p}$ or $p_t \leq p^n(p_t) \leq \overline{p}$, with equality 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 $\tilde{S}(|p - p_t|; p_t)dF(p)$. When similarity is given by Equation (3) and the price distribution is normal with variance $\pi^2$, 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 $\Sigma_{sim} = \begin{bmatrix} \frac{1}{\delta_p} & 0 \\ 0 & \frac{1}{\delta} \end{bmatrix}$. This product the variance matrix $\Sigma = \left(\Sigma_{sim}^{-1} + \Sigma_w^{-1}\right)^{-1}$, and the price norm is then the top element of the vector $\Sigma\Sigma_w^{-1}\mu + \Sigma\Sigma_{sim}^{-1}\kappa$. Plugging in $\Sigma_F$ and $\Sigma_{sim}$ 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 $\mathbb{E}_F(p|c_t) = \overline{p} + \rho\frac{\gamma}{\pi}(c_t - \bar{c})$. As $\delta_p \longrightarrow 0$, this becomes $\frac{\overline{p}+2\delta\gamma^2\mathbb{E}_F(p|c_t)}{1+2\delta\gamma^2}$, while as $\delta_p \longrightarrow \delta$ we obtain Equation (8).

**Proposition 3**. Define

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

Setting $V(p_t, p^n(p_t)) = p_t$ becomes $\sigma(p_t, p^n(p_t))[p_t - p^n(p_t)] = p_t - p^n(p_t)$, with solutions $p_t = p^n(p_t)$ and $\sigma(p_t, p^n(p_t)) = 1$. 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, This is given by $p_t = p^n(\overline{p}, c_t)$. In turn, $\sigma(p_t, p^n(p_t))$ is an increasing function of $p_t$ for $p_t > p^n(\overline{p}, c_t)$ and a decreasing function of $p_t$ for $p_t < p^n(\overline{p}, c_t)$. To see this, write:

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

where $m_t \equiv \min[p_t, p^n]$, and $M_t \equiv \max[p_t, p^n]$. 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 $V\big(p_t, p^n(p_t)\big) = p_t$ if and only if $p_t \in \{p_L, p^n(\overline{p}, c_t), p_H\}$.

Furthermore, setting $p^n = p^n(p_t) = \frac{\overline{p}+2\delta\pi^2 p_t}{1+2\delta\pi^2}$ we find:

$$\frac{\partial V(p_t, p^n)}{\partial p_t} = \frac{2\delta\pi^2}{1+2\delta\pi^2} + \frac{\sigma(p_t, p^n)}{1+2\delta\pi^2} + \frac{\sigma'\left(\frac{M_t}{m_t}\right)\overline{p}(p_t - \overline{p}) sgn\big(p_t - p^n(p_t)\big)}{m_t^2[1+2\delta\pi^2]} > 0$$

so $V(p_t, p^n)$ is monotonically increasing in $p_t$.

Undervaluation $V(p_t, p^n) < p_t$ 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 $\sigma\big(p_t, p^n(p_t)\big) > 1$ 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 $\sigma\big(p_t, p^n(p_t)\big) < 1$. 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 $V(p_t, p^n)$. 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 $\frac{\partial V(p_t, p^n)}{\partial p_t}$ as:

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

where $1 - \alpha x = (1-\alpha)\frac{\overline{p}}{p^n} > 0$. It then follows that $\partial_{p_t}^2 V$ 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 $x = \lim_{p_t \to +\infty} \frac{p_t}{p^n(p_t, c_t)}$, then it has exactly one zero for $x > 1$. In this case, $V(p_t, p^n)$ has exactly one inflection point for $p_t > p^n(p_t)$.

A similar calculation shows that, under the same conditions, $V(p_t, p^n)$ has exactly one inflection point for $p_t < p^n(p_t)$. Finally, we assume differentiability of $V(p_t, p^n)$ 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.** We first characterize the region where valuation is too flat, $\partial_{p_t} V\big(p_t, p^n(p_t, c_t)\big) < 1$. Consider first the region where $p_t > p^n(p_t)$. For simplicity, write $\alpha = \frac{2\delta\pi^2}{1+2\delta\pi^2}$ and again use $x = \frac{p_t}{p^n} > 1$. Then $\partial_{p_t} V < 1$ becomes:

$$\sigma'(x)(x-1)(1-\alpha x) - (1-\alpha)\big(1-\sigma(x)\big) < 0$$

where $1 - \alpha x = \frac{\overline{p}}{\overline{p} + 2\delta\pi^2 p_t} > 0$. The left-hand side of the equation starts out negative for $x = 1$ but it increases monotonically in $x$ provided:

$$\sigma''(x)(x-1) + 2\sigma'(x) > 0$$

for all $x$. Moreover, the left-hand side is positive at $p_t = p_H$, where $\sigma(x) = 1$. Thus, under this condition, there is a unique price level $\tilde{p}_H > \overline{p}$ such that $\partial_{p_t} V = 1$. Moreover, this price satisfies $\frac{\partial \tilde{p}_H}{\partial \overline{p}} > 0$. To see this, write $F(x) = \sigma'(x)(x-1)(1-\alpha x) - (1-\alpha)\big(1-\sigma(x)\big)$. Then

$$\frac{\partial \tilde{p}_H}{\partial \overline{p}} = -\frac{\partial_{\overline{p}} F}{\partial_{\tilde{p}_H} F} = -\frac{\partial_{\overline{p}} x}{\partial_{\tilde{p}_H} x} > 0$$

Because the nominator is negative while the denominator is positive. Consider now the case $p_t < p^n(p_t)$. Similar arguments show that there exists a single price level $\tilde{p}_L < p^n(p_t)$ such that $\partial_{p_t} V = 1$, under the same condition on the curvature of the salience function. Moreover, we find again that $\frac{\partial \tilde{p}_L}{\partial \overline{p}} = -\frac{\partial_{\overline{p}} x}{\partial_{\tilde{p}_H} x} > 0$ for $x = \frac{p^n(\tilde{p}_L, c_t)}{\tilde{p}_L}$.

Finally, using $p^n(\tilde{p}_X, c_t) = \frac{\overline{p}+2\delta\gamma^2\mathbb{E}_F(p|c_t)+[2\delta\pi^2+4\delta^2\gamma^2\pi^2(1-\rho^2)]p_t}{1+2\delta\gamma^2+[2\delta\pi^2+4\delta^2\gamma^2\pi^2(1-\rho^2)]}$ it is straightforward to show that $\frac{\partial \tilde{p}_H}{\partial \overline{p}} - \frac{\partial \tilde{p}_L}{\partial \overline{p}} > 0$. This arises out of the diminishing sensitivity property of salience.

Consider now the case where context is uncorrelated with prices so that $p^n(p_t) = \frac{\overline{p}+2\delta\pi^2 p_t}{1+2\delta\pi^2}$. In this case is is straightforward to characterize the effect of price volatility $\pi^2$ on the range $[\tilde{p}_L, \tilde{p}_H]$. We find:

$$\frac{\partial \tilde{p}_H}{\partial \pi^2} = -\frac{\partial_{\pi^2} F}{\partial_{\tilde{p}_H} F} = -\frac{\partial_{\pi^2} x}{\partial_{\tilde{p}_H} x} > 0$$

because $\partial_{\pi^2} x \propto 2\delta\tilde{p}_H[\overline{p} - \tilde{p}_H] < 0$ (and $\partial_{\tilde{p}_H} x > 0$). In contrast, we find:

$$\frac{\partial \tilde{p}_L}{\partial \pi^2} = -\frac{\partial_{\pi^2} F}{\partial_{\tilde{p}_L} F} = -\frac{\partial_{\pi^2} x}{\partial_{\tilde{p}_L} x} < 0$$

because $x = \frac{\overline{p}+2\delta\pi^2\tilde{p}_L}{\tilde{p}_L(1+2\delta\pi^2)}$ and $\partial_{\pi^2} x \propto 2\delta\tilde{p}_L(\tilde{p}_L - \overline{p}) < 0$ (and $\partial_{\tilde{p}_L} x < 0$).

**Proposition 5.** Because $V(p_t, p^n)$ is monotonically decreasing, it has a unique threshold $p_{WTP}$ such that the consumer buys the product if $p_t \leq p_{WTP}$ and does not buy the product otherwise.

Given the assumption that $\delta \to 0$, so that the price norm is equal to the price that held in the previous occurrence of the contextual cue, the Proposition is equivalent to the following statement: there exist two thresholds $p_L^n$ and $p_H^n$, with $p_L^n < q < p_H^n$ such that $p_{WTP}(p_H^n) > q$ for any $p_{1h} > p_H^n$, and $p_{WTP}(p_L^n) < q$ for any $p_{1l} < p_L^n$. We first show that the thresholds are defined uniquely by $p_{WTP}(p_H^n) = q = p_{WTP}(p_L^n)$ (together with $p_L^n < q < p_H^n$). We then show that $\frac{\partial p_{WTP}(p^n)}{\partial p^n} > 0$ for $p^n > p_H^n$ and for $p^n < p_L^n$.

By the implicit function theorem:

$$\frac{\partial p_W}{\partial p^n} \propto \frac{\partial V(p_W, p^n)}{\partial p^n} = - \cdot \begin{cases} 1 - \sigma(p_{WTP}, p^n) - \sigma'(p_{WTP}, p^n)\left[\frac{p^n}{p_{WTP}} - 1\right], & p^n > p_{WTP} \\ 1 - \sigma(p_{WTP}, p^n) - \sigma'(p_W TP, p^n)\frac{p_{WTP}}{p^n}\left[\frac{p_{WTP}}{p^n} - 1\right], & p^n < p_{WTP} \end{cases}$$

The first step relies on the fact that $\partial V / \partial p_W$ is positive. Because $\frac{\partial p^n}{\partial \overline{p}} = \frac{1}{1+2\delta\pi^2}$ is also positive, it follows that $\frac{\partial p_W}{\partial \overline{p}} > 0$ if and only if

$$\begin{cases} \sigma(p_{WTP}, p^n) + \sigma'(p_{WTP}, p^n)\left[\frac{p^n}{p_{WTP}} - 1\right] > 1, & p^n > p_{WTP} \\ \sigma(p_{WTP}, p^n) + \sigma'(p_{WTP}, p^n)\frac{p_{WTP}}{p^n}\left[\frac{p_{WTP}}{p^n} - 1\right] > 1 & p^n < p_{WTP} \end{cases}$$

Thus, willingness to pay increases with the norm whenever the surprise entailed by $p_{WTP}$ is sufficiently large.

It now suffices to show that these conditions are satisfied for the price norms such that $p_{WTP} = q$. This follows trivially, since $p_{WTP} = q$ implies a correct valuation, $V(q, p_X^n) = q$, which in turn implies $\sigma(q, p^n) = 1$.