Forthcoming in Management Science

The Impact of Prior Decisions on Subsequent Valuations in a Costly Contemplation Model

Elie Ofek and Muhamet Yildiz with Ernan Haruvy*

October 2006

*Elie Ofek: Harvard Business School, Soldiers Field, Boston, MA 02163 (eofek@hbs.edu). Muhamet Yildiz Department of Economics, MIT, Cambridge, MA 02142 (myildiz@mit.edu). Ernan Haruvy: University of Texas at Dallas, School of Management, 2601 North Floyd Road, Richardson, TX 75083 (eharuvy@utdallas.edu).
The authors would like to thank Daron Acemoglu, Abhijit Banerjee, Glenn Ellison, Al Roth, and Luc Wathieu for helpful suggestions. The paper has also benefited from the comments of participants at the Behavioral Research Council conference (Great Barrington, MA, July 2002), and seminar participants at Duke University, UC Berkeley, and UCLA. The authors also thank the Associate Editor and three anonymous reviewers for insightful comments that helped improve the paper.

The Impact of Prior Decisions on Subsequent Valuations in a Costly Contemplation Model

Abstract

This paper develops and tests a model of how recall of information from past decisions affects subsequent related decisions. A boundedly rational individual has to determine her willingness to pay for a good that she previously considered purchasing at a given price, or provide valuations for a set of goods that she previously ranked in order of preference. The individual is ex-ante uncertain about her utility from consumption of the goods and can exert costly cognitive effort to reduce this uncertainty. We show that incorporating information from a prior decision has three primary effects. (a) Valuations are expected to exhibit higher variance; in particular, the spread of valuations between the most and least preferred alternatives increases. (b) Decision makers will, in expectation, exert more effort during the valuation phase. And (c) the relative impact of prior decisions on valuation spread increases the more each attribute contributes to overall utility. The model predictions are then tested in a series of controlled lab experiments.

(*Sequential Decisions; Preference Uncertainty; Effort-Accuracy Tradeoff; Willingness to Pay; Recall; Bounded Rationality*)

1 Introduction

Past decisions often serve as input to subsequent related decisions. Specifically, the conclusion reached in a previous decision can potentially shed light on how to approach the decision at hand. For example, an individual who several weeks ago evaluated a certain bouquet of flowers and was in favor of purchasing it at a given price now has to decide whether to buy a similar bouquet of flowers at a higher price. A football fan who couldn't secure a ticket for a sold-out game at regular price now has to decide how much to bid on e-Bay for an auctioned ticket. A young professional who recently accepted a job offer in a large city, rejecting a similar but higher paying job located in a small suburban town, is currently deciding how much to bid on a house located in the suburbs. In all these cases, the question arises as to how knowledge of the prior decision (willingness to purchase the bouquet at a certain price; desire to buy the ticket at regular price; rejection of the job located in a suburban town) affects willingness to pay for the option currently being considered (the value to assign to the new bouquet of flowers; the amount to bid on tickets online; the amount to bid on the suburban house).

When individuals are certain of their preferences, or can establish them effortlessly, previous evaluations should not influence subsequent decisions. In reality, however, individuals may face considerable uncertainty regarding the value of a good to them, or how they should trade off different product attributes. For example, in negotiating payment on a particular floral arrangement for an upcoming dinner party, the host may wonder about the importance of bright vs. dark colors, the length of the stems, or the type of flowers in the arrangement (tulips, orchids, etc.). The relative importance of each floral arrangement attribute can depend on a number of aspects relevant for the consumption occasion—such as who will be coming to the party, what activities are planned for the party, or where the flowers will be displayed.

Given that the individual is not sure about her preferences, she may try to reduce the uncertainty in several ways. First she can exert effort, in the form of cognitive thinking or time-consuming research, to ascertain the influence of the different aspects. For instance, the host could reflect on the guest list to determine how the floral arrangement under consideration might impress each of them, and she could contemplate on how the arrangement would appear in different locations in the house. Second, the decision-maker might invoke information from past related decisions. For example, the host may recall her conclusion regarding a similar bouquet she saw a few weeks ago in a floral shop. In the context of such sequences of decisions, we ask: how are the incentives to expend effort to determine willingness to pay affected by recalled information from a prior decision? How does incorporating information from a previous decision affect the final valuations used to make future decisions?

The goal of this paper is to answer these questions theoretically and experimentally. We construct a model of individual decision-making that has three central features: (a) Individuals are uncertain about their preferences, in the sense of how much utility will be derived from consuming a good or how much each attribute will contribute to overall utility. (b) Costly cognitive effort can be expended to reduce the uncertainty. And (c) input from a previous decision can be incorporated. In other words, we investigate how boundedly rational individuals use their own prior decisions as a source of information about their utility structure when making subsequent decisions. We analyze the case of an individual who needs to determine her willingness to pay for a single good that she previously considered purchasing at a given price. We also examine the robustness of the primary forces at work when two alternatives are evaluated and a prior choice between them is taken into account.

We identify the broad conditions under which three central results hold. First, we show that incorporating recall of a prior decision outcome leads to more extreme valuations, that is, valuations that deviate considerably from the individual's ex-ante mean. When multiple alternatives are under consideration, taking into account a prior decision will result in a greater spread of valuations between the goods. Second, we find that this increased spread is more pronounced when the decision stakes are higher (that is, when the goods are of higher ex-ante expected value). Lastly, we find that the expected amount of effort expended in valuing alternatives is greater when prior decisions are taken into account. This last result is especially intriguing, given that effort has already been expended on making the previous decisions.

The main intuition giving rise to these results is that prior decisions not only convey information about the value of the good in and of themselves, but also affect the incentives to expend effort to gain more information in subsequent related decisions. We show that the objective function that relates costly cognitive effort expended to the amount of variance explained, the "effort-accuracy" relationship, is altered by past decision outcomes. Specifically, the differing outcomes of prior decisions have an *asymmetric* impact on how an individual trades off current effort with the desire to arrive at a more accurate valuation. Depending on whether a good was chosen or rejected in a previous decision, the individual considers the good to be of greater or lesser value respectively, relative to the ex-ante mean. In the former case (greater value), the individual perceives a very high marginal return from effort that overshadows the lower marginal return from effort in the latter case (lesser value).

Consequently, incorporating prior decisions is likely to yield final valuations that are more informed, capture more of the variance associated with uncertain preference parameters, and that hence differ considerably from their (uninformative) ex-ante mean.

We conducted a series of experiments using actual prizes in a familiar product category, dining at local restaurants. The empirical results largely confirmed the main implications of the theory: a prior decision increased the spread of valuations between subjects' most and least preferred alternatives, and, on average, subjects who made relevant prior decisions took significantly longer time than those who had not to determine valuations. Furthermore, the impact of a prior decision on valuation spread became more pronounced as the average value of prizes increased.

The rest of the paper is organized as follows: Section 2 relates our work to relevant literature. Section 3 develops a theoretical framework for modeling the impact of prior decisions on subsequent valuations. Section 4 formulates the key findings of the model as hypotheses, which are tested in a series of controlled lab experiments. Section 5 concludes. All proofs are given in the Appendix.

2 Literature Review

In a series of papers, Fischer et al. (2000a, 2000b) posit that decision makers can be uncertain about their preferences. Their use of random attribute weights is similar to our approach in Section 3.6. The major difference between our work and theirs is that we allow individuals to reduce preference uncertainty by exerting cognitive effort and by recalling relevant prior decisions. By contrast, in Fischer et al. uncertainty is exogenously fixed. Our work is thus more in line with behavioral decision theory, which demonstrates that individuals face "effort-accuracy" trade-offs in making decisions (Payne et al. 1993). Moreover, we find it highly plausible that if an individual makes prior decisions regarding an alternative, those will be taken into account in subsequent decisions.[1]

Hsee (1996) and Hsee et al. (1999) study a specific form of preference uncertainty, whereby an attribute whose importance is ex-ante uncertain will receive less weight when an alternative containing the attribute is evaluated separately than when it is evaluated jointly with another alternative. This can lead to preference reversals across the two evaluation modes, but has only been demonstrated between subjects. By contrast, our focus is on the implications for final valuations of the *same*

[1] Khan and Meyer (1991) have examined the issue of uncertain preference weights focusing on the framing of attribute levels as above or below an acceptable threshold. Eliashberg and Hauser (1985) incorporate a random utility component that is assumed to arise from the researcher's measurement error and not from decision makers' uncertainty about their own preferences. These two papers as well ignore effort to reduce uncertainty and recall of past decisions.

individual intertemporally *combining* information from a sequence of decisions. Hence, our analysis is more appropriate for dynamic contexts where decisions are not a one shot task.

Although we model an individual who incorporates information from her *own* prior decision, our analysis would be similar if the previous decision were made by some *other* individual, as long as their preferences and incentives are identical. This links our model to the literature on social learning (e.g., Banerjee 1992, Bikhchandani et. al 1992, Gale 1996), where each agent receives a private and independent signal for the value of undertaking some behavior and agents' decisions are sequential and observable. Our analysis differs from this literature as follows. First, individuals in that literature always face the same kind of decision while we consider two different types of related decisions (a purchase decision at a given price and a willingness to pay assessment). Second, and more importantly, in the social learning literature signal precision is exogenous (with all individuals receiving equally accurate signals), while we focus on the case where the individual endogenously determines the precision of each signal. Hence, our approach allows examining the impact of prior decisions on the incentives to invest in the accuracy of future signals.[2]

In our model, individuals may recall only partial details of a past decision. Hirshleifer and Welch (2002) show that amnesic decision makers, who recall only previous actions but not previous signals, may follow their current signal more often than full recall individuals, who remember all previous actions and signals.[3] This happens when there is considerable environmental uncertainty that can change an alternative's true value. We do not incorporate such environmental uncertainty. However, relative to our model, they treat signal accuracy as exogenously fixed, which may be a strong assumption given their managerial focus. In Dow (1991), individuals sequentially search two sellers of a good for the lowest price. Consumers partition prices into categories and remember only the category a previously encountered price belongs to. Chen et al. (2005) go a step further by analyzing a market in which competing firms incorporate such memory limitations in their price-setting strategies. In our analysis, consumers are uncertain about their own valuations but not about prices encountered. Our

[2] We believe that individuals who face the kind of consumption decisions modeled in this paper will likely exert effort to reduce uncertainty. Furthermore, for the goods we consider there is typically considerable heterogeneity in consumer tastes, making the observation of others' behavior less relevant for own preferences. At some level, one can consider combining our approach with that of social learning to analyze certain market phenomena, such as the patterns of bids in successive auctions. In such cases, an individual may see the prices at which previous auctions closed (but not the individual bids) and believe that her preferences are correlated with those of previous bidders.

[3] Recalling finer details makes the effects we find even more pronounced (Section 3.5). We also note work by Mehta et al. (2004) who study how consumers form successive (costless) evaluations when past evaluations are recalled imperfectly. Our model is better suited to handle cases where the effort exerted to resolve uncertainty plays a key role in the decision.

analysis is hence more relevant when past prices are known (e.g., when price is part of the product description itself, or the individual took note of the prices). Moreover, there is evidence in psychology (Engelkamp 1998, Lingle and Ostrom 1979) that individuals can reliably recall their past actions and the decisions they faced but not the underlying reasoning that led to their actions. This suggests that it is likely that the decision to "buy or forgo a good at a given price" is recalled but that the informational content of aspects considered and that led to the action are less reliably recalled.

Cognitive effort in our model entails a cost, which can be relevant for a broad set of preference formation problems (Ergin, 2003).[4] Shugan (1980) offers a methodology for quantifying the thinking cost of comparing different alternatives. In his model the individual knows her utility function and uncertainty arises from having to examine ex-ante unknown attribute levels. Our model is more appropriate when product attribute levels are known (e.g., the specs of a computer) but the individual is uncertain how much each attribute or feature will contribute to her utility.

3 Model and Theoretical Results

This section presents our main theory and results. We first analyze the case of a decision maker who considers a single good of uncertain value to her (Sections 3.1-3.4). We start by describing the model's characteristics. In particular, we explain the relationship between the amount of effort expended and the information gained about the value of the good. We also describe the process through which new information from subsequent effort is integrated. Next, we establish the optimal effort the individual would choose to incur: a) when asked to decide between purchasing and rejecting the good at a given price, and b) when asked to provide her monetary value for the good. Lastly, we analyze how recall of a past decision regarding the purchase of the good would impact a subsequent valuation of the good, in terms of the effort expended and the properties of the final valuation. In Section 3.6, we extend our framework to examine the case of multiple goods that possess multiple attributes.

3.1 Model Setup: Single Good Case

Consider a consumer who is contemplating whether to purchase a good of ex-ante unknown value v to her, which is offered at a given price p. The consumer can make a more informed decision by engaging

[4] Zwick et al. (2003) show experimentally that including costs impacts the amount of search individuals perform relative to the amount predicted by theory. The importance of incorporating thinking costs in decision making can also be gleaned from experiments on strategic behavior in coordination games (e.g., Ho and Weigelt 1996).

in introspection about her utility. Introspection reveals information on various consumption aspects that affect her utility from the good, and can be regarded as a mental cost accompanied by disutility.[5] The consumer is thus rational in that she wishes to make an optimal decision that maximizes expected utility, but she is constrained (or bounded) by the costliness of effort needed to acquire information relevant for resolving uncertainty about her utility.[6]

Given the consumer's uncertainty about the value of the good, we treat v as a random variable and assume without loss of generality that its expected value is $E[v] = 1$. The consumer may expend effort c (measured in terms of utility) to reduce the uncertainty associated with v. If, based on the information gained from the effort, she chooses to purchase the good at price p, her utility will be

$$v - p - c.$$

If she rejects the good after incurring effort c, her utility will be $(-c)$.

One can envision a situation where the value v of the object is a function of many uncertain aspects that are stochastically independent, and that the more effort the individual exerts the more aspects she learns about. Specifically, let

$$v = y_0 y_1 \cdots y_n \cdots , \tag{1}$$

where y_n are independent random variables with mean 1 that reflect the uncertain aspects. By expending effort c, the individual learns the values of some of these random variables. We write $N(c)$ for the set of aspects she learns, so that the individual's information about v is $\hat{v} = \prod_{n \in N(c)} y_n$. Her remaining uncertainty about v is $\varepsilon = \prod_{n \notin N(c)} y_n$, where $\hat{v}$ and ε are independent. Clearly, we have

$$v = \hat{v}\varepsilon. \tag{2}$$

We will write $F(\cdot; c)$ for the cumulative distribution function (CDF) of $\hat{v}$ and $V(c)$ for the variance of $\hat{v}$. In our analysis, we assume for simplicity that information is divisible so that c can take on a

[5] The consumer can also expend costly resources to obtain information from *external* sources, such as by speaking to friends, getting an expert opinion, or surveying relevant literature.

[6] Contemplation or deliberation costs are considered to be a central aspect of bounded rationality in the literature (Rubinstein 1998, Conlisk 1996), as are issues related to uncertain preferences (March 1978). We believe that treating decision-makers as having perfect knowledge of their preferences for all goods at all times without having to expend any effort to figure them out is an unrealistic idealization. In this paper we propose an "as if" representation of how individuals approach costly evaluations to determine their preferences. In this representation, individuals understand the effort-accuracy tradeoff involved and select an optimal amount to think. Thus, we envision a world in which individuals have prior experience regarding the procedural aspects of thinking, but the details for given situations are left to be resolved when they arise. Otherwise, individuals must also think about the amount of thinking to do, the amount of thinking to solve this problem and so on, leading to an infinite regress (Lipman, 1991).

continuum of values, denoted by $[0, \bar{c}]$, where $\bar{c}$ is the cost of learning all the aspects. In the Appendix we provide details on how this is consistent with the structure in (1).

Note that the means of $\hat{v}$ and ε are 1, and the conditional expectation of v given $\hat{v}$ is $E[v|\hat{v}] = \hat{v}$. $Var(\hat{v})$ measures how much of the original variance is explained by $\hat{v}$. The more aspects the individual learns (i.e., the larger the set N), the greater the variance of $\hat{v}$ and the lower the variance of ε. If she does not expend any effort $(c = 0)$ then no information is obtained, in which case $\hat{v}$ captures none of the variance and $E[v|\hat{v}] = 1$ is essentially the initial expectation. If, on the other hand, she expends maximal effort $(c = \bar{c})$ and learns all the aspects, then $\hat{v}$ captures all the variance in v.

V is the function that maps a given thinking effort into the amount of variance explained by the information gained from that effort. We make the following straightforward assumptions on V.

Assumption 1 *V is a strictly increasing, strictly concave, and three-times continuously differentiable function of c.*

The condition that V is increasing $(V' > 0)$ implies that as the consumer exerts more cognitive effort (or thinks longer), she generates a more precise idea about v. Indeed, in our model the individual will try to maximize the variance of $\hat{v}$ subject to the cost of doing so. Therefore, for any given cost level c if the individual chooses the set $N(c)$ optimally then $V(c) = Var(\hat{v})$ will be increasing in c. The concavity of V $(V'' < 0)$ ensures that the first-order condition is sufficient for optimization. This will be consistent with our restricted attention to interior solutions, in which there is some remaining uncertainty when the individual makes her decision. Since $V'' < 0$, the first derivative V' has an inverse, which is denoted by $h(\cdot)$.

We make the following assumption regarding the CDF of $\hat{v}$ resulting from effort c.

Assumption 2 *The function $F(\cdot; c)$ is continuously differentiable, and there exists some $\bar{v} > 0$ such that $F(\bar{v}; c) = 1$ for each $c \in [0, \bar{c}]$.*

The assumption that the value of the good to the consumer is bounded from above (by some $\bar{v}$) simplifies our analysis and will be consistent with our experimental setting.

3.1.1 Subsequent Effort, New Information, and the Updated Valuation of the Good

As our focus is on how prior decisions about goods influence subsequent valuations of these goods, we need to specify how the individual *combines* different pieces of information arising from separate decisions. Through initial effort c_0, the individual learns the values of some set N_0 of aspects. What she knows is summarized by $\hat{v}_0 = \prod_{n \in N_0} y_n$. If she then expends c_1 units of effort to learn some of the aspects that are not already incorporated into $\hat{v}_0$, she will get new information $\hat{v}_1 = \prod_{n \in N_1} y_n$, where $N_0 \cap N_1 = \varnothing$. Writing $\varepsilon_1 = \prod_{n \notin N_0 \cup N_1} y_n$ for the remaining uncertainty, we have

$$v = \hat{v}_0 \hat{v}_1 \varepsilon_1. \tag{3}$$

Here, $\hat{v}_0$ and $\hat{v}_1$ and the remaining error ε_1 are stochastically independent, but they are all dependent to v. Note that $E[\hat{v}_0] = E[\hat{v}_1] = E[\varepsilon_1] = 1$. Note also that $Var(\hat{v}_0) = V(c_0)$ and $Var(\hat{v}_1) = V(c_1)$. The relationship in (3) implies that the conditional expectation of v given $\hat{v}_0$ and $\hat{v}_1$ is $E[v|\hat{v}_0, \hat{v}_1] = \hat{v}_0\hat{v}_1$, which is clearly stochastically dependent to $\hat{v}_0$, but the extra information from new effort will be stochastically independent from the original information.[7]

Example 1 For a concrete illustration of information acquisition and updating in our model, consider an individual who is invited to a dinner party along with $n + 1$ other guests. The value of attending the party clearly depends on how much she will enjoy the company of all other guests. Assume also that there are synergies between the guests. In particular, the value of going to the party is $v = y_0 y_1 \cdots y_n$, where y_i summarizes how the idiosyncrasies of guest i will affect her enjoyment (for example, guest i may inspire interesting conversation, tell jokes, introduce her to new people). The parameters $y_0, \ldots, y_n$ are independently distributed with mean 1. By thinking about guest i, she gets an idea about i's impact on the party. Imagine that by expending effort c_0 she learns about the first k_0 guests. In that case, $\hat{v}_0 = y_0 \cdots y_{k_0}$ and $\varepsilon_0 = y_{k_0+1} \cdots y_n$. Suppose that upon thinking more she now learns about guests $k_0 + 1, \ldots, k_1$ by expending effort c_1. She thus obtains new information $\hat{v}_1 = y_{k_0+1} \cdots y_{k_1}$. In total, she now knows the value of $\hat{v}_0\hat{v}_1 = y_0 \cdots y_{k_1}$ and has remaining uncertainty captured by $\varepsilon_1 = y_{k_1+1} \cdots y_n$.[8]

Note that in this example when the individual learns about k aspects through effort c, the variance of her information is $E\left[y_i^2\right]^k - 1$. If acquiring this information costs $c = C^k - 1$ for some $C > E\left[y_i^2\right]$, then the variance can be written as $V(c) = (1 + c)^{\alpha} - 1$, where $\alpha = \log\left(E\left[y_i^2\right]\right) / \log C < 1$. Clearly, V is an increasing and concave function of c and satisfies all of our assumptions.[9]

[7] This specification is consistent with literature on information acquisition (e.g., McCardle 1985), where independent signals are purchased and combined to form an estimate, and literature on forecasting (e.g., Gul and Lundholm 1995), where the value of a project depends on combining independent signals.

[8] Note that information cannot be acquired indefinitely as the total amount is fixed. We make parametric assumptions under which not all the information is exhausted to ensure interior solutions for effort level (see Footnote 11).

[9] If the cost were linearly proportional to k, i.e., $c = Ck$, then $V(c) = (\beta^c - 1)$, where $\beta = E\left[y_i^2\right]^{1/C}$, and V would

Remark 1 We wish to point out that our theory holds broadly in any model in which effort $c_0 \in [0, \bar{c}]$ yields information $\hat{v}_0$ that satisfies (2) and in which additional effort yields new information $\hat{v}_1$ that satisfies (3); along with the other assumptions on V and F. The structure in (1) and Example 1 reflect a detailed information acquisition process that can conform to the characteristics of the broader model.

In what follows, we explore the implications of our model set-up. Though we have made several stylized assumptions, we believe our model captures the key elements of the phenomenon under study and is well suited for the type of Bayesian analysis conducted here. In a separate Technical Appendix (available on the journal website) we provide details of the following extensions: a) using an additive rather than multiplicative updating process for (3), b) allowing correlation between new and prior information, c) an alternative specification where the information gained each time is a holistic measurement of the good's value (as in econometric modeling).

3.2 Purchase of a Good at a Given Price

A consumer faces a decision between purchasing and rejecting a good at a given price p_0. The consumer needs to determine the amount of cognitive effort c_0 to incur in making her decision. Consistent with previous notation, upon thinking c_0 units the consumer obtains information contained in $\hat{v}_0$. She buys the good if and only if she finds that $\hat{v}_0 \geq p_0$; this happens with probability $\Pr(\hat{v}_0 \geq p_0)$. When she finds that $\hat{v}_0 < p_0$, the consumer forgoes purchasing the good but has still incurred the cognitive effort c_0. Therefore, her expected utility U_0 from thinking c_0 units is

$$U_0(c_0; p_0) = E[v - p_0 | \hat{v}_0 \geq p_0] \Pr(\hat{v}_0 \geq p_0) - c_0. \tag{4}$$

The consumer optimizes (4) with respect to c_0. Note that the expectation operator E depends on c_0 through $F(\cdot; c_0)$. Proposition 1 establishes several characteristics of the optimal effort level.

Proposition 1 *Let* $\hat{c}_0 \equiv \arg\max_{c_0 \in [0,\bar{c}]} U_0(c_0; p_0)$. *Then,*

(1) $\hat{c}_0$ *is a single-peaked function of* p_0 *and is maximized at* $p_0 = 1$ *if* $\frac{\partial F(v;c)}{\partial c} \geq 0$ *at each* $v < 1$*, and* $\frac{\partial F(v;c)}{\partial c} \leq 0$ *at each* $v > 1$*;*

(2) $\hat{c}_0$ *increases whenever both* v *and* p_0 *are multiplied by a known constant* $\lambda > 1$*.*

be a convex function of c. For V to be concave then, $E\left[y_i^2\right]$ should be sufficiently decreasing in i. In that case, if the individual is forgetful, then she may have an incentive to contemplate on aspects she previously thought about; introducing correlation between information from successive decisions.

Proposition 1 tells us that the optimal effort ($\hat{c}_0$) increases as the price p_0 is closer to the expected value $E[v] = 1$. This makes intuitive sense: the individual thinks more when she is ex-ante indifferent between having the good and keeping her money; that is, when the decision is a "close call". She thinks less when p_0 moves away from 1 in either direction, because the ex-ante dominance of the good or the money makes the decision "easier." The condition stated in part (1) of the proposition ensures that as the individual exerts more effort, she obtains more information (in the sense of second-order stochastic dominance). In addition, when the incentives or stakes are higher—that is, when the surplus $(v - p_0)$ is multiplied by some constant λ—the individual is induced to think more.

3.3 Valuation of a Good

In a valuation task, the consumer decides on the highest amount she would be willing to pay for a good. To set the benchmark for comparison, this section analyzes such a task assuming no prior information is available to her. The individual must first determine the amount of cognitive effort, c_1, to expend. Contemplation effort generates information $\hat{v}_1$, which, in the absence of additional information, is her valuation of the good (because $E[v|\hat{v}_1] = \hat{v}_1$). Thinking is costly, hence the amount of effort to expend on forming and then submitting a truthful valuation depends on how the payoff is determined. We assume the following mechanism: after the individual submits her willingness to pay for the good, she will face a randomly drawn price of p_1 and will buy the object if the price is lower than or equal to her submitted value. The price p_1 is drawn from a uniform distribution on the interval $[0, m/2]$ with $m/2 > \bar{v}$. This mechanism ensures that the valuation the individual provides has a real consequence, in terms of whether the individual gets the good or keeps her money. In our experiments, we create an identical payoff scheme using the well-known BDM mechanism proposed by Becker et al. (1964).

We will write $U_1(c_1)$ for her expected utility when she expends c_1 units of effort and faces the mechanism described above.

Proposition 2 *Given any $c_1 \in [0, \bar{c}]$, we have*

$$U_1(c_1) = V(c_1)/m + 1/m - c_1. \tag{5}$$

This proposition tells us how the individual trades off the cost of thinking with the benefit of getting more information regarding v in a valuation task. One can interpret $V(c_1)/m$ as the benefit of incurring c_1 units of effort. The first-order condition for optimal effort that maximizes (5) is

$$U_1'(\hat{c}_1) = V'(\hat{c}_1)/m - 1 = 0. \tag{6}$$

Since V is strictly concave, the solution to (6) is unique. Since V is also an increasing function, the optimal cognitive effort in valuation, $\hat{c}_1$, is increasing with $1/m$.[10] Defining the inverse function $h = (V')^{-1}$, from (6) we obtain

$$\hat{c}_1 \equiv \arg \max_{c_1 \in [0, \bar{c}]} U_1 (c_1) = h (m) . \tag{7}$$

In our setup, the individual expends cognitive effort to learn about aspects that determine the value of a good to her. She provides her best estimate for the value before learning the price, and she buys the good if the realized price is less than this value. Clearly, this captures many important situations, such as bidding in an auction, posting an ask price in a financial market, determining one's upper price limit prior to entering a bargaining negotiation (say on a used car), or instructing an agent at what maximal price to buy a good in a spot market. In all these cases, an individual typically only has a probability distribution for the final price (others' bids, or others' selling value) and this price is realized only after her willingness to pay is determined. The specific level of cognitive effort the individual expends will quantitatively depend on the distribution of the price. In our paper, we consider a simple uniform distribution that leads to (5). In general, if price is distributed with density f, then the relevant function for the benefit of thinking c units would be $V^f (c) = E \left[\int_{p \leq v\varepsilon} (v - p) f (p) \, dp \right]$. As long as we can specify a task in which there is a utility consequence for how close willingness to pay estimates are to the true value, that is, accuracy matters, then our results will carry through.

3.4 Forming Valuations with Recall of a Prior Decision

How does the consumer determine her willingness to pay for a good *that she has previously contemplated purchasing at price* p_0? The individual knows that when she previously considered purchasing the good, she obtained information $\hat{v}_0$. We assume that she at least recalls whether $\hat{v}_0$ is greater or lower than p_0, because she remembers whether or not she decided that the good was worth purchasing at that price. In our context, remembering whether $\hat{v}_0 < p_0$ or $\hat{v}_0 \geq p_0$ represents a minimal partition of the valuation space for $\hat{v}_0$ relative to the price p_0. (In Section 3.5 we will show that our main findings are even more pronounced when the individual recalls finer information). We write I to signify the information the individual recalls about a previous decision. The individual selects an optimal amount of effort to expend on valuation, taking into account the information I. As described

[10] For an intuition, suppose the individual knows that $v \in [0, 1]$. Whatever information she obtains about v will only matter if $p \in [0, 1]$. From the uniform distribution, the price will fall in this range with probability $2/m$ ($m > 2$). Hence, the benefit of more information is multiplied by a factor proportional to $1/m$ and the individual has a lower incentive to resolve uncertainty when m is higher (because of the lower likelihood that price will fall in the relevant range).

in Section 3.1.1, subsequent contemplation yields new information, which is now denoted by $\hat{v}_1^I$, that satisfies $v = \hat{v}_0 \hat{v}_1^I \varepsilon_1^I$,where $\hat{v}_0$, $\hat{v}_1^I$, and ε_1^I are stochastically independent conditional on I. The amount of effort expended on forming valuations when prior information is available is denoted by c_1^I. Note that $\hat{v}_1^I$ has CDF $F\left(\cdot; c_1^I\right)$ and variance $Var\left(\hat{v}_1^I\right) = V\left(c_1^I\right)$ and that ε_1^I is the remaining information she does not know. The individual's final valuation of v, which incorporates details from the prior decision and the new information from contemplation in the valuation phase, is

$$\hat{v}_{01} = E[v|I, \hat{v}_1^I] = \hat{v}_1^I E[\hat{v}_0|I]. \tag{8}$$

As in the proof of Proposition 2, we compute the expected utility from thinking c_1^I units to be

$$U_1^I(c_1^I|I) = (E[\hat{v}_0|I])^2 \left[V\left(c_1^I\right) + 1\right] / m - c_1^I. \tag{9}$$

Notice that $(E[\hat{v}_0|I])^2 V\left(c_1^I\right)$ is the variance of $\hat{v}_{01}$ conditional on I. Furthermore, optimal contemplation in the valuation phase satisfies[11]

$$\hat{c}_1^I[I] \equiv \arg\max_{c_1^I \in [0,\bar{c}]} U_1^I(c_1^I|I) = h(m/(E[\hat{v}_0|I])^2). \tag{10}$$

Since $(E[\hat{v}_0|\hat{v}_0 \geq p_0])^2 \geq 1$ and $(E[\hat{v}_0|\hat{v}_0 \leq p_0])^2 \leq 1$, we have the following relationships:

$$\hat{c}_1^I[\hat{v}_0 \geq p_0] = h(m/(E[\hat{v}_0|\hat{v}_0 \geq p_0])^2) \geq h(m) = \hat{c}_1, \tag{11a}$$

$$\hat{c}_1^I[\hat{v}_0 \leq p_0] = h(m/(E[\hat{v}_0|\hat{v}_0 \leq p_0])^2) \leq h(m) = \hat{c}_1. \tag{11b}$$

From (11a), if the individual recalls that she decided to purchase the good ($\hat{v}_0 \geq p_0$), she infers that she must have found the good to be relatively valuable. Hence, the consequences of erring in the true value of v are high, and she has a greater incentive to expend effort to make a more informed decision. Conversely, from (11b), when the individual recalls that she rejected the good at price p_0, she infers that she must have found the good to be less valuable than her ex-ante expectation. Thus, the consequences of erring are low, and she devotes less effort to resolving the remaining uncertainty.[12]

This analysis reveals an important point. The individual faces a different objective function (9) in choosing an optimal effort level in each of the following cases: (a) no prior decision, (b) recall of

[11] To guarantee an interior solution for thinking effort c_1^I, we assume that $V'(0) > m/(E[v|I, \hat{v}_0 < p_0])^2$ and $V'(\tilde{c}) < m/(E[v|I, \hat{v}_0 \geq p_0])^2$, where $V(\tilde{c}) = Var(\varepsilon_0)$.

[12] The following example illustrates these points. Let the expected value of v, conditional on the prior decision implying that $\hat{v}_0 \geq p_0$, be 100. If the final estimator is then determined with a remaining 10 percent error, this could lead to a potential mistake of up to \$10. If, however, the expected value of v conditional on the prior decision implying that $\hat{v}_0 < p_0$ is 20, a 10 percent error would lead to a potential mistake of only \$2.

a prior decision to purchase the good at a given price, and (c) recall of a prior decision to reject the good. In other words, the effort-accuracy trade-off that the individual faces itself varies depending on the nature of the information available from a prior decision. This is vividly apparent in Figure 1: when recall of a prior decision implies that the good would be purchased ($\hat{v}_0 \geq p_0$), the incentives to invest effort in the subsequent valuation phase are high because the variance captured by integrating new information as a function of effort expended has a higher slope than with no prior decision ($\forall c$). The reverse is true if the individual recalls a prior decision to reject the good ($\hat{v}_0 < p_0$).

The multiplicative noise structure (3) does play a role here. In general, the results in (11a)-(11b) will be true whenever v is a supermodular function of $\hat{v}$ and ε, such as the multiplicative specification. When this does not hold, for example under the additive specification $v = \hat{v}_0 + \hat{v}_1 + \varepsilon$, the effort level for $\hat{v}_1$ will be independent of prior decisions (we formalize this in the Technical Appendix).

Remark 2 (*Consistency and Preference Reversals*) When the individual recalls a previous purchase decision, her final valuation will reflect a "preference for consistency" in the following sense. From (8), knowing that she previously decided to purchase the good, the individual scales up new information obtained ($\hat{v}_1^I$) by $E[\hat{v}_0|\hat{v}_0 \geq p_0] \geq 1$ in formulating her final valuation for v. Similarly, when she recalls that she decided to reject the good, she scales down new information by $E[\hat{v}_0|\hat{v}_0 < p_0] \leq 1$. Even so, preference reversals are possible when contemplation in the valuation phase (c_1^I) yields new information $\hat{v}_1^I$ that sharply contradicts the results of the previous decision. For example, let the individual recall that $\hat{v}_0 \geq p_0$. If $\hat{v}_1^I < p_0/E[\hat{v}_0|\hat{v}_0 \geq p_0]$ then $\hat{v}_{01}\left[\hat{v}_0 \geq p_0\right] < p_0$ and the individual will now refuse to purchase the good even at some price $p < p_0$. The probability of this event is $F\left(p_0/E[\hat{v}_0|\hat{v}_0 \geq p_0]; \hat{c}_1^I[\hat{v}_0 \geq p_0]\right)$. Similarly, when the individual recalls that $\hat{v}_0 < p_0$, but finds $\hat{v}_1^I > p_0/E[\hat{v}_0|\hat{v}_0 < p_0]$, then $\hat{v}_{01}\left[\hat{v}_0 < p_0\right] > p_0$. These reversals do not reflect a systematic bias; they are due to the likelihood that contradictory new information is encountered. Using the above logic, we can analyze how more than two decisions interact (see the Technical Appendix for an illustration).

Remark 3 In our analysis we assume that, conditional on I, $\hat{v}_1^I$ is independent of $\hat{v}_0$. Under broad conditions, this assumption is implied by the optimality of information acquisition when the individual knows which information she contemplated in the past. In particular, if the individual knows that in a prior decision she considered the aspects contained in $\hat{v}_0$ she will subsequently choose to learn different aspects contained in $\hat{v}_1^I$. For instance, in Example 1 take $\hat{v}_0 = y_0$ and I is what she later recalls about $\hat{v}_0$ (e.g., that she should go to the dinner party). In a subsequent contemplation about the party, the individual knows that she has thought about guest y_0, *but not necessarily* y_0 *itself.* If I provides sufficiently precise information about $\hat{v}_0$, then it would be wasteful to think about y_0 again, and she would rather expend effort to learn about y_1. In general, she would prefer to learn the value of y_1 rather than y_0 if $E\left[y_1^2\right] \geq E\left[y_0^2|I\right]/E\left[y_0|I\right]^2$. One can easily check that this condition is satisfied for most distributions, such as the uniform, exponential, and power distributions. In that case, she will have an incentive to acquire information that is independent of $\hat{v}_0$.

3.4.1 The Impact of Prior Decisions on the Variance of the Final Valuation

We now examine how prior decisions affect the variance of the final valuation for the good, namely $\hat{v}_{01}$. Recall that V is the function that maps a given cognitive effort into the amount of variance explained by the information obtained from that specific effort. Here, $-V''/V'$ measures the local concavity of V. Treating V as a VNM utility function from effort, $-V''/V'$ can be regarded as the degree of absolute risk aversion. We will focus on the case in which $-V''/V'$ is nonincreasing.[13]

For convenience, let us write x for the conditional expectation of $\hat{v}_0$ given the information recalled from the previous decision, such that: $x = E[\hat{v}_0|\hat{v}_0 \geq p_0]$ when the individual recalls that she decided to purchase the good at p_0; $x = E[\hat{v}_0|\ \hat{v}_0 < p_0]$ when she recalls that she decided to reject the good; and $x = 1$ when there was no prior decision. From (7) and (10), one can then express the solution for optimal thinking effort during valuation as $\hat{c} = h\left(m/x^2\right)$, with CDF $F\left(\cdot; h\left(m/x^2\right)\right)$. It is convenient to denote the variance of the final valuation the individual has for the good as Ψ, where from (8)

$$\Psi\left(x\right) = Var\left(\hat{v}_{01}|I\right) = (E[\hat{v}_0|I])^2 V\left(c_1^I\right) = x^2 V\left(h\left(m/x^2\right)\right), \forall x > 0. \tag{12}$$

When no prior decision has taken place, $\Psi\left(1\right) = Var(\hat{v}_1)$. If the individual recalls that she would purchase the good at price p_0, $\Psi(E[\hat{v}_0|\hat{v}_0 \geq p_0]) = Var(\hat{v}_{01}|\hat{v}_0 \geq p_0)$, and so on. It is useful to note from (12) that a prior decision affects the variance of $\hat{v}_{01}$ in two ways. First, the information embodied in the prior decision directly scales up or down the variance of new information gained in the valuation phase, as reflected in the x^2 multiplying term. Second, a prior decision affects the solution to the optimal effort in the valuation decision, as reflected in the argument of the term $h\left(m/x^2\right)$.

If the individual recalls that she previously rejected the good, then $x^2 = (E[\hat{v}_0|\hat{v}_0 < p_0])^2$ is small (less than 1). We also know from (11b) that the individual thinks less than if no previous decision had taken place. As a result, her final valuation explains less of the variance, is *less accurate,* and tends

[13] This property is common in microeconomic analysis and much work has involved utility functions with constant (CARA) and decreasing (DARA) absolute risk aversion (such as $1-e^{-\alpha c}$ with $\alpha > 0$, $\log(1+\beta c)$, and c^μ with $0 < \mu < 1$).

to rely more heavily on the prior mean. By contrast, when the individual recalls that she previously decided to purchase the good, $\hat{v}_0 \geq p_0$, x^2 is large and she is prompted to think more according to (11a). In this case, her final valuation explains more variance, is *more accurate,* and tends to rely less on the prior mean. Mathematically, we write:

$$\begin{aligned} Var(\hat{v}_{01}|\hat{v}_0 &< p_0) = \Psi(E[\hat{v}_0|\hat{v}_0 < p_0]) < \Psi(1) = Var(\hat{v}_1), \\ Var(\hat{v}_{01}|\hat{v}_0 &\geq p_0) = \Psi(E[\hat{v}_0|\hat{v}_0 \geq p_0]) \geq \Psi(1) = Var(\hat{v}_1). \end{aligned} \tag{13}$$

So far we have shown that a prior decision can have divergent effects depending on the recalled outcome. The more substantive question is whether we can make a general *ex-ante* statement about the final valuation for the good. To put it differently, should we anticipate an increase or decrease in the unconditional variance of the final valuation, namely $Var(\hat{v}_{01})$, if it is preceded by a decision on whether or not to purchase the good? The following proposition addresses this question.

Proposition 3 *Whenever $-V''/V'$ is nonincreasing, we have $Var(\hat{v}_{01}) \geq Var(\hat{v}_1)$.*

The inequality states that the variance of willingness to pay is expected to be higher when information about a previous decision is available. The intuition for this result is as follows. When $-V''/V'$ is non-increasing, the marginal benefit from additional effort does not decrease steeply (i.e., V is not "too concave"). Under this condition, when the individual recalls that $\hat{v}_0 \geq p_0$ she is prompted to expend substantial effort to figure out the true value of v and to scale up the variance of new information gained by $x^2 = E[\hat{v}_0|\hat{v}_0 \geq p_0]^2$, to an extent that overshadows the lower effort she expends when she knows that $\hat{v}_0 < p_0$. Because the total amount of information the individual has about v is greater with recall of a prior decision, there is a higher likelihood that the individual will form a final valuation that is further away from the initial uninformative expectation. Consequently, final valuations that incorporate recall of a prior decision ($\hat{v}_{01}$) are expected to have higher variance.[14]

As explained, Proposition 3 implies that incorporating a past purchase decision is likely to yield final valuations that stray considerably from the ex-ante mean of $E(v) = 1$. Hence, we should expect valuations determined by taking into account prior decisions to be more extreme. Indeed, a canonical example developed in the Technical Appendix shows that recall of prior choices will lead to more subsequent purchases at very high or very low prices. Moreover, in the multiple-good case (Section

[14] If recall of a prior decision had no impact on the amount of effort expended in the valuation phase, the finding in Proposition 3 would not have required any condition on the function V. In such a case, a prior decision would unequivocally add information to that gained from subsequent contemplation and the final valuation would always have greater variance. In our case, this is not true because prior decisions do affect the effort expended (per 11a-11b).

3.6) we show that because incorporating prior decisions increases the variance of final valuations, this will lead to a greater dispersion of monetary valuations between two goods (Corollary 1).

3.4.2 Increasing the Decision Stakes

One might expect the effects of bounded rationality to diminish as the stakes get higher. We examine this issue by multiplying both v and p by some constant $\lambda > 1$, so that the agent's utility is $\lambda(v-p)-c$. This is meaningful if we consider a greater quantity of the same good, or if the type of good is of higher ex-ante expected value. The superscript λ will indicate payoffs that are multiplied by λ.

Proposition 4 $[Var(\hat{v}_{01}^{\lambda}) - Var(\hat{v}_1^{\lambda})]$ *increases with* λ *when (i)* $-V''/V'$ *is nonincreasing, (ii)* $-V'\left(h\left(1/x^2\right)\right)/V''\left(h\left(1/x^2\right)\right)$ *is convex in* x, *and (iii) the condition stated in Proposition 1 holds.*

Hence, the impact of incorporating a previous decision on the variance of the final valuation is in fact more pronounced as the decision stakes increase. To understand the intuition, note that an increase in λ increases $[Var(\hat{v}_{01}^{\lambda}) - Var(\hat{v}_1^{\lambda})]$ in two ways. First, by Proposition 1, as λ increases the individual will have exerted more prior effort to decide whether to purchase at price p_0. This makes the individual treat $\hat{v}_0$ as more informative, and results in more extreme scaling of the new information embodied in $\hat{v}_1$. This phenomenon tends to increase $Var(\hat{v}_{01}^{\lambda})$ from (12), while $Var(\hat{v}_1^{\lambda})$ is unaffected. Second, an increase in λ induces the individual to exert more effort in the valuation stage too. This increases both $Var(\hat{v}_{01}^{\lambda})$ and $Var(\hat{v}_1^{\lambda})$. Under (ii), it impacts $Var(\hat{v}_{01}^{\lambda})$ more heavily.

3.4.3 Impact of a Prior Decision on Subsequent Effort Expended

We have seen that recall of a previous decision affects optimal effort in the valuation stage. According to (11), the individual thinks less if the prior decision deemed the value of the good low ($\hat{v}_0 < p_0$) and vice-versa if the value was deemed high ($\hat{v}_0 \geq p_0$). It is not clear, however, which effect dominates—specifically, whether the unconditional amount of expected effort is higher or lower when a prior decision is taken into account. The following proposition addresses this issue.

Proposition 5 *Expected optimal effort in valuation satisfies* $E\left[\hat{c}_1^I\right] \geq \hat{c}_1$ *if* $-\frac{V''\left(h\left(1/x^2\right)\right)}{V'\left(h\left(1/x^2\right)\right)}x$ *is a non-increasing function of* x.

Proposition 5 states that recall of a prior purchase decision prompts more effort (in expectation) when subsequently determining a valuation for the same good. The condition in the proposition once again relates to the curvature (or local concavity) of V as a function of effort. Here, however, it is closely tied to the notion of *relative risk aversion* due to the multiplication by x. Thus, and somewhat counterintuitively, this result tells us that even though effort has already been expended in the previous purchase decision, we can expect the effort expended in a subsequent valuation decision to be higher.[15]

3.5 When the Individual Recalls Finer Details from a Prior Decision

Consider a partition $P = \{P_1, P_2, \ldots\}$ of $[0, \infty)$, where $P_1, P_2, \ldots$ are disjoint subsets with $[0, \infty) = P_1 \cup P_2 \cup \cdots$. Assume that the individual recalls in which set P_k her previous estimate $\hat{v}_0$ belongs to. In our model, when the individual recalls whether she would purchase the good or not, the relevant partition has two sets $P_1 = [0, p_0)$ and $P_2 = [p_0, \infty)$ and the individual only knows whether $\hat{v}_0$ belongs in P_1 or P_2. A partition $P = \{P_1, P_2, \ldots\}$ is finer than a partition $Q = \{Q_1, Q_2, \ldots\}$ if each Q_k can be written as a union of some of the sets in P. A finer partition means that the individual recalls more.

Proposition 6 *Let P be a finite partition that is finer than Q, and let $\hat{v}_{01}[P]$, $\hat{v}_{01}[Q]$, $\hat{c}_1^I[P]$, and $\hat{c}_1^I[Q]$ denote the final valuations and effort levels under these partitions. We have*

$$\begin{aligned} Var(\hat{v}_{01}[P]) &\geq Var(\hat{v}_{01}[Q]) \ if - V''/V' \ is \ nonincreasing, and \\ E\left[\hat{c}_1^I[P]\right] &\geq E\left[\hat{c}_1^I[Q]\right] \ if - \frac{V''\left(h\left(1/x^2\right)\right)}{V'\left(h\left(1/x^2\right)\right)} x \ is \ a \ nonincreasing \ function \ of \ x. \end{aligned}$$

That is, the effects stated in Propositions 3 and 5 become more pronounced when the individual recalls finer details. Note that we can obtain Q by taking a union of two sets in P one by one. But every time we compare to such a union, $Var(\hat{v}_{01})$ and $\hat{c}_1^I$ increase in expectation by Propositions 3 and 5 respectively. It is this inductive logic that forms the intuition for Proposition 6 and that we formally use in the proof. In sum, our theory applies to the broader case in which the individual recalls more details than whether she decided the object was worth purchasing, and it could also apply to the case in which the previous decision task was to explicitly provide an assessment of the value of the object.

3.6 The Multiple-Good, Multiattribute Case

We extend our analysis to a two-good multiattribute context, where an individual recalls a rank ordering of the goods and then determines her willingness to pay for each.[16] Consider two goods, X

[15] The multiplicative specification plays a role in this result. Under an additive specification we would have $E\left[\hat{c}_1^I\right] = \hat{c}_1$.

[16] The relationship to the single-good case can be gleaned by treating the money the individual would have to pay (= the price) as the second good, and the value (v) of the good as a single attribute which is uncertain (the "money" good

and Y, defined in terms of multiple attributes a_X^j and a_Y^j for $j = 1, 2, 3, \ldots$, which yield utilities $u_X = \sum_j \delta^j w_j \left(a_X^j\right)$ and $u_Y = \sum_j \delta^j w_j \left(a_Y^j\right)$ from consuming each of the goods X and Y respectively. The parameters δ^j are independently distributed random variables. Here, w_j are functions that transform levels of attribute j into utility. The individual knows attribute levels a_X^j and a_Y^j and the functions w_j, but she does not know δ^j. Hence, the individual is uncertain about the relative importance of each attribute to overall utility. The individual can expend effort levels c^j and obtain information about δ^j, to be denoted $\hat{\delta}^j$, such that $\delta^j = \hat{\delta}^j \varepsilon^j$ and all $\left\{\hat{\delta}^j, \varepsilon^j,\ j = 1, 2, \ldots\right\}$ are independently distributed. The total cost of these effort levels to the individual is $\sum_j c^j$; the CDF of each $\hat{\delta}^j$ is $F_j\left(\cdot; c^j\right)$ and its variance is $V_j\left(c^j\right)$, where V_j satisfies the same properties as V defined in Section 3.1.

Ranking. In comparing goods X and Y to determine her preference ranking, the individual gets information $\hat{\delta}_0^j$ by exerting effort $\hat{c}_0^j$. Her valuations for the goods are $\hat{u}_X = \sum_j \hat{\delta}_0^j w_j \left(a_X^j\right)$ and $\hat{u}_Y = \sum_j \hat{\delta}_0^j w_j \left(a_Y^j\right)$, and she prefers good X iff $\hat{u}_X \geq \hat{u}_Y$.

Valuation. In a valuation decision the individual provides valuations $\hat{u}_X$ and $\hat{u}_Y$ so that if confronted with prices p_X and p_Y she will purchase good X iff $\hat{u}_X \geq p_X$ and good Y iff $\hat{u}_Y \geq p_Y$. These prices are independently drawn from a uniform distribution on the interval $[0, m/2]$. This setup is consistent with the BDM mechanism (Becker et. al, 1964) used in our experiments.

We write $\overrightarrow{c}_1 = \left(c_1^1, c_1^2, \ldots\right)$ for the vector of effort levels in a valuation decision and $U_1\left(\overrightarrow{c}_1\right)$ for the expected utility from incurring effort $\overrightarrow{c}_1$ when no prior ranking taking has place; a superscript I indicates that prior ranking is recalled. Because we can apply our previous analysis from the single-good case to each attribute a^j separately, we have $E\left[\hat{c}_1^{j,I}\right] \geq \hat{c}_1^j$ whenever the coefficient of relative risk aversion $\left(-xV_j''\left(h\left(1/x^2\right)\right)/V_j'\left(h\left(1/x^2\right)\right)\right.$ is nonincreasing $\forall j$, as in Proposition 5. Therefore, the sum $\sum_j E\left[\hat{c}_1^{j,I}\right]$ will be at least as high as the total effort without prior ranking, $\sum_j \hat{c}_1^j$. In analogy to (12), we can define a function $\Psi_j(x) = x^2 V_j\left(h_j\left(\frac{1}{A_j x^2}\right)\right), \forall x > 0$. If $-V_j''/V_j'$ is nonincreasing, then Ψ_j is convex, yielding $Var(\hat{\delta}_{01}^j) \geq Var(\hat{\delta}_1^j)\ \forall j$. We write $\hat{u}_X^1$, $\hat{u}_Y^1$, $\hat{u}_X^{01}$, and $\hat{u}_Y^{01}$ for the expectations of u_X and u_Y induced by using $\hat{\delta}_1^j$ and $\hat{\delta}_{01}^j$ respectively ($\hat{\delta}_{01}^j = E[\delta^j | I, \hat{\delta}_1^{j,I}]$).[17] We obtain the following corollary regarding the dispersion of valuations between the two goods:

Corollary 1 *When $-V_j''/V_j'$ is nonincreasing,* $E\left[(\hat{u}_X^{01} - \hat{u}_Y^{01})^2\right] \geq E\left[(\hat{u}_X^1 - \hat{u}_Y^1)^2\right]$.

This is a testable result that we wish to highlight. The expected squared difference between the

has zero level on this attribute, while the physical good has zero level on the price dimension).

[17] For example, if the individual recalls that she preferred good X to Y, $\hat{\delta}_{01}^j = E[\delta^j | \hat{\delta}_1^{j,I}, \hat{\delta}_0^j \{X \succeq Y\}]\ \forall j$.

valuations of two goods increases when there is information about a previous preference ordering.[18] Intuitively, this happens because recall of a prior ranking will result in final valuations that contain greater variance and that more likely differ from the ex-ante mean. In turn, this will increase the valuation spread. Using Proposition 4, we can show that multiplying each w_j by $\lambda > 1$ increases the impact of a prior decision on final valuation spread—even when the spread is normalized by λ^2:

Corollary 2 *Under the conditions of Proposition 4 for each V_j and F_j, we have that*

$$E\left[(\hat{u}_X^{01,\lambda} - \hat{u}_Y^{01,\lambda})^2\right]/\lambda^2 - E\left[(\hat{u}_X^{1,\lambda} - \hat{u}_Y^{1,\lambda})^2\right]/\lambda^2 \textit{ is increasing in } \lambda. \tag{14}$$

Therefore, all of our results could be generalized to a model with uncertainty about multiple attributes and a prior rank ordering or choice between two goods.

4 Testing Model Implications

We now present a series of experiments designed to examine our model's predictions in actual decision-making situations. We first state the primary hypotheses we wish to test.

With regard to how recall of a prior decision affects the amount of effort individuals expend in a subsequent valuation, we have predicted asymmetric effort levels depending on the outcome of the prior decision (see (11a-11b)). This is captured in the following hypothesis:

H1 *More effort will be expended when recall implies selection of the good than when it implies rejection of the good.*

We have further shown that, unconditional on the outcome, a prior decision will lead to greater effort in expectation. (See Proposition 5.) We therefore state the following hypothesis:

H2 *Recall of a prior decision when determining valuations will result in greater effort.*

Our theory has made predictions about the impact of prior decisions on the variance of valuations and about the effects of increasing the decision stakes. Because one needs multiple observations within the same decision to meaningfully capture such effects, we focus on the implications for the multiple-good case.[19] Based on Corollaries 1 and 2, we can state the following hypotheses:

H3 *Recall of a prior decision will increase the spread between the monetary values for two goods.*

H4 *The impact on valuation spread of recalling a prior decision is increasing in the decision stakes.*

[18] We use the squaring operator to control for the sign of the difference (good X preferred to Y or vice versa).

[19] To test our results on variance with respect to a single good, we would either need each individual's ex-ante mean valuation for the good or multiple observations per subject for the same good.

4.1 Experimental Design and Method

To test the hypotheses, we conducted a series of experiments that presented individuals with decisions about dining at restaurants. The restaurants were described in terms of two or three attributes. The set of possible attributes included location (different areas in Cambridge, Massachusetts), type of food (Asian, Indian, seafood), service level, food quality, and decor. (Descriptions of each restaurant and sample instructions appear in the Technical Appendix. Restaurant names were not given.) Dining at a given restaurant was described to include an appetizer, main course, and dessert. Stimuli were presented via a computer interface, and subjects' responses were recorded via the same interface. The subjects included students (undergraduate and graduate) and nonstudents. Subjects were promised a minimum of $10 and a chance to earn more in cash or prizes, as explained below.

4.2 Single Good

Experiments 1 and 2 were designed to test our hypotheses on effort expended within a single good context. Experiment 1 uses the amount of time that subjects spent to determine their valuations as the measure of cognitive effort. Experiment 2 uses the amount of time that subjects explicitly specified they would need for each subsequent valuation task.

4.2.1 Experiment 1

Subjects were randomly assigned to one of two conditions: "choice and subsequent valuation" (CV) or "valuation only" (VO). The CV condition (N=63) consisted of two stages. In stage I, participants were asked to choose between a dinner-for-one voucher at a specific restaurant or a voucher for any regular-priced CD at a well-known music-store chain. In each decision the restaurant was different, but it was always chosen or rejected against a regular-priced CD. Subjects made five such choices, one at a time. Subsequently, in stage II the same subjects provided their maximum willingness to pay in dollars for a meal at each of the restaurants presented in stage I. *Subjects were reminded each time of their stage-I decision* (the reminder was provided below the restaurant's description).

In the VO condition (N=61), subjects made five irrelevant (filler) decisions in stage I.[20] In stage II, subjects were asked to state their maximum willingness to pay for the same five restaurants employed in the CV condition. Thus, a stage II valuation in the VO condition did not entail recall of a previous decision about the restaurant in question.

[20] The filler task equated the number of decisions across conditions. We thank Al Roth for suggesting this.

Subjects were told the following about how prizes would be awarded at the end of the experiment: The computer would randomly select one of their decisions. If a stage I decision was selected, the subject would receive a dinner-for-one or CD voucher, depending on the choice indicated in that decision. If a stage II decision was selected, the computer would then randomly draw a number. If the dollar value the individual assigned to the restaurant in that decision was greater than the randomly drawn number, the individual would receive a dinner-for-one voucher at that restaurant. Otherwise, the individual would receive the random number in dollars. This procedure induces truth-telling and is consistent with the BDM mechanism. Participants were presented with several examples before stage II to familiarize them with the mechanism. The random number was drawn from the range [0-30], which was pretested with 32 subjects to be the upper bound of their valuations.

Results The computer recorded the amount of time that subjects spent determining their willingness to pay for each restaurant (for average times see Table 1 in the appropriate columns for Experiment 1). To examine H1, we focused on the time subjects spent on valuation in the CV condition. When a restaurant was preferred in stage I, H1 predicts more effort to have been expended in determining valuation for that restaurant in stage II than when a CD was preferred. We estimated the following regression model on stage II data from the CV condition:

$$T_{CV} = \theta_0 + \theta_j d_j + \theta_R REST + e, \tag{15}$$

where T_{CV} is the time that subjects spent determining a restaurant's dollar value in the CV condition, θ_0 is an intercept term, $d_j, j \in [1, 2, ..., 4]$ are dummy variables to control for the specific decision,[21] $REST = 1$ if the subject chose the restaurant in stage I (and 0 if the CD was chosen), and e is a standard normal error term.[22] Regression analysis yields a parameter estimate of θ_R=6.41 ($p = 0.079$). Thus, we find support for the asymmetric prediction in H1. Recalling a prior preference for a restaurant over a CD increased the amount of time spent to determine a restaurant's value by 6.41 seconds.

We estimated the following regression model to examine the unconditional effect of a prior decision

[21] In all the regressions reported in the paper we included dummies to control for the specificities of each decision. The need to do so is justified by the appropriate F-test for pooling across decisions (Greene, 2003). In most cases, though our findings qualitatively hold without controlling for pooling, we find that a null hypothesis of: {all decision dummy coefficients=0} is rejected. Therefore, statistically we need to include the decision dummies. We report the full estimation results with and without dummies, and the F-test for each regression, in the Technical Appendix.

[22] In our regressions we use a variable related to the prior decision as a regressor. If there are unobserved factors (unrelated to our theory) that affect subjects' responses, the measured prior decision variable can be correlated with the model's disturbance (e) due to these factors (see Greene 2003 pages 84-85). In this case, our estimated parameters may be attenuated relative to their true value.

on the effort expended in determining valuations:

$$T = \gamma_0 + \gamma_j d_j + \gamma_{CV} CHOICE + e, \tag{16}$$

where T is the time that subjects spent in stage II determining a valuation for any given restaurant, γ_0 is an intercept term, d_j are as in (15), and $CHOICE$ designates whether a prior decision about the restaurant vs. the CD had taken place (i.e., CV condition). We obtain a parameter estimate of $\gamma_{CV} = 7.42$ ($p = 0.0019$), which confirms H2. Note from Table 1 that recall of a prior choice increased the average amount of time that subjects spent on valuation by over 30 percent.[23]

We also checked the correlation between a good's valuation and the time spent in the VO condition. This correlation was found to be small (0.0533) and nonsignificant ($p = 0.354$). Hence, our results on the impact of prior decisions cannot be explained by positing that inherently better options are allocated more time. When in (16) we included a variable to capture the dollar value given to each restaurant, we still get a significant result for the impact of a prior decision ($\gamma_{CV} = 7.98$, $p = 0.007$).

4.2.2 Experiment 2

In Experiment 1 we used subjects' decision time as our measure of effort. This presumes that the incentives to expend cognitive effort in a given decision are reflected in the amount of time actually taken to make that decision, and is consistent with prior literature (Payne et al. 1993). We wanted to examine our most basic result on the incentives to expend effort (per H1) with a more direct measure. To achieve this, we ran a similar study to the CV condition in Experiment 1 (N=50), except that in stage II there was the following additional step. Before being shown each restaurant, subjects were asked how many seconds they wanted to have available for determining their valuation, with every two seconds costing them one cent.[24] Subjects were told whether the restaurant about to be shown would be one that they chose or rejected (relative to a CD) in stage I. On the next screen subjects saw the details of the restaurant for which they were to provide a dollar value, and a clock placed on the right hand side showed the number of seconds remaining. If the time they had requested expired before a dollar value was entered, the computer randomly selected a value and the next decision appeared.

[23] In one of the five decisions (Decision 2) we get that the average time in the VO condition is slightly higher than in the CV condition. However, this difference is not significant (p=0.67). Furthermore, the median time spent in this decision is actually 3 seconds higher in the CV condition; in accordance with our theory. Hence, a number of outliers in the VO condition are likely skewing the average in this decision.

[24] Various costs were pretested with a set of 38 subjects. The cost used is roughly equal to the money worth of time (subjects expected to participate for one and a half hours and the expected total payment was $25).

Results The average times requested for each decision are presented in Table 2, broken out by whether a Restaurant or CD was preferred. Subjects explicitly requested more time when they knew they were about to face a restaurant that they had previously preferred to the CD than vice versa (by 28% on average). Running regression model (15) on this data yielded $\gamma_{CV} = 1.95$ (p =0.0012). It should be pointed out that making subjects conscious of time issues and tagging a cost (even though small), seems to have reduced their general desire to expend effort (irrespective of what they preferred). Nonetheless, the results support H1.

4.3 Multiple Goods

Experiments 3-4 were designed to test our hypotheses on valuation spread and the decision stakes (H3-H4) and to replicate some of our findings on effort expended in a two-good context.

4.3.1 Experiment 3

In this experiment subjects made five decisions about *pairs* of restaurants. Subjects were randomly assigned either to a "ranking and subsequent valuation" condition (RV) or to a "valuation only" condition (VO). In stage I of the RV condition (N=39), subjects were presented with *two* restaurants in each decision and asked to rank them by designating their preferred restaurant as '1' and their less preferred restaurant as '2'. In stage II of the RV condition, subjects were sequentially shown the same five pairs of restaurants and asked to state their maximum willingness to pay for a dinner-for-one at each; that is, subjects provided two valuations in each decision. *Subjects were reminded of their stage I ranking designations.* In the VO condition (N=37), stage I consisted of five irrelevant (filler) decisions. In stage II of the VO condition, subjects stated their maximum willingness to pay for the same five restaurant pairs used in the RV condition. The mechanism for determining payoffs at the end of this experiment was similar to that in Experiment 1. In this case, if a stage I decision was drawn in the RV condition, a subject would get a voucher for his/her preferred restaurant (designated with a '1'). If a stage II decision was selected, the computer would randomly choose one of the restaurants from that decision and the BDM truth-telling mechanism explained earlier would be employed.

Results As is evident from Table 1, the average time devoted in Experiment 3 to valuation decisions was 50 percent higher in the RV condition than in the VO condition (32.62 vs. 21.72 seconds respectively). A regression analysis similar to (16), with a dummy variable for prior ranking, yields

γ_{RV}=10.9 ($p < 0.01$). Thus, a prior ranking of the alternatives significantly increased the amount of time spent determining willingness to pay, compared to when no such ranking had taken place; reconfirming H2 for the two-good case.[25]

In accordance with Corollary 1, we tested whether a prior preference ranking affected the squared spread of valuations using the following regression model:

$$(\hat{u}^A - \hat{u}^B)^2 = \alpha_0 + \alpha_j d_j + \alpha_{RV} RANK + e, \tag{17}$$

where $\hat{u}^A$ and $\hat{u}^B$ are the valuations that subjects provided for each restaurant in a given pair, α_0 is an intercept term, $d_j, j \in [1, 2, ..., 4]$ are decision dummies, and $RANK$ is a dummy variable denoting whether prior ranking took place. The results yield a parameter estimate of $\alpha_{RV} = 6.25$ ($p = 0.10$), which marginally supports H3 that prior ranking leads to a greater spread. It can be gleaned from Table 3 that the average absolute valuation spread between alternatives (across all pairs) was \$3.26 in the VO condition and \$4.05 in the RV condition (a 24 percent difference). If in regression model (17) we use the absolute difference between valuations instead of the squared difference as the dependent variable, we again find support for H3 (at a greater significance level, $p = 0.01$).

4.3.2 Experiment 4

Our main goal was to examine how an increase in the decision stakes affects valuation spread. The design was similar to that of Experiment 3, except that a different set of subjects now evaluated dinner-for-two vouchers (N=44 in each condition) and the random number drawn by the computer for the payoff mechanism was between 0-55 (pretested with 16 subjects to be an upper bound).

Results We first ran regression (17) on the dinner-for-two data. The results yield a parameter estimate of $\alpha_{RV} = 44.90$ ($p < 0.01$), which strongly supports H3. To test for the effect of increasing the decision stakes, we computed the following variable for each pair of restaurants in a given decision

$$y_{i,j} = \frac{(\hat{u}^A_{i,j} - \hat{u}^B_{i,j})^2}{(\hat{u}^A_{i,j} + \hat{u}^B_{i,j})^2}, \tag{18}$$

where $\hat{u}^A_{i,j}$ and $\hat{u}^B_{i,j}$ are the valuations provided by subject i in decision j for restaurants A and B respectively. We calculated the same variable for each pair of restaurants for each decision and subject in Experiment 3. The normalization by the factor of $(\hat{u}^A_{i,j} + \hat{u}^B_{i,j})^2$ is what allows us to compare the

[25] We measured the time that subjects spent determining valuations for both restaurants on the same screen. Hence, we could not separate out the time spent on the preferred vs. less preferred restaurant to test H1.

data from the two experiments to see if indeed increasing the stakes has a proportionately greater effect on the dispersion of valuations.[26] We estimate the following regression equation:

$$y = \beta_0 + \beta_j d_j + \beta_{RV_1} RV * D1 + \beta_{RV_2} RV * D2 + \beta_{VO_2} V0 * D2 + e, \tag{19}$$

where β_0 is an intercept term, $d_j, j \in [1, 2, ..., 4]$ are decision dummies, RV and VO are dummies for the experimental condition, and $D1, D2$ are dinner-for-one and dinner-for-two indicators respectively. The estimation yields $\hat{\beta}_{RV_2} = 0.15$ ($p < 0.01$) while $\hat{\beta}_{RV_1} = 0.04$ ($p = 0.35$), thus supporting H4. Table 3 shows that, for dinner-for-two decisions, the average absolute valuation spread between alternatives (across all pairs) was \$4.78 in the VO condition and \$7.39 in the RV condition (a 55 percent difference).

4.4 Limitations of the Experiments

Though we have found support for the implications of our model per hypotheses H1-H4, the experiments we conducted have limitations with respect to the link with the model structure. We have demonstrated that the amount of effort expended in a decision is affected by the directional outcome of a prior decision. This is consistent with our model structure for a multiplicative (or more generally supermodular) relationship between information uncovered in separate contemplation occasions (see Section 3.1.1) and would not be consistent with other relationships, for example additive. However, we have not directly tested for this updating process. One could design an experiment where in the first stage, subjects would be asked to write down a valuation for a good based on a specific set of aspects. In a later stage, they would be directed to consider other aspects that might affect their utility and write down a new valuation, and in the last stage, subjects would be reminded of the two separate valuations and asked to form one last valuation. This could allow a direct testing of the relationship between the two separate valuations.

In our model development, we assumed independence of new information gained in a subsequent decision. Though we discuss relaxing this assumption in the Technical Appendix, in our experiments we did not directly manipulate or control for this issue. To check for the incentive to gain new information that is independent rather than correlated with previous information (see also Remark [illegible]), one could conduct an experiment along the following lines. Each subject would act as an agent that is asked to make an optimal purchase decision on behalf of a principal, whose preferences are

[26] Corollary 2 calls for normalizing each individual's squared spread of valuations by $\lambda_{i,j}^2$. Given that we do not observe $\lambda_{i,j}$, we use $(\hat{u}_{i,j}^A + \hat{u}_{i,j}^B)^2$ as a proxy. To see why this works, assume that subject i values a dinner-for two voucher twice as much as a corresponding dinner-for-one voucher. If the impact of prior choice on valuation spread is non-increasing in the decision stakes then $y_{i,j}(D1) = \frac{(\hat{u}_{i,j}^A - \hat{u}_{i,j}^B)^2}{(\hat{u}_{i,j}^A + \hat{u}_{i,j}^B)^2} = \frac{(2\hat{u}_{i,j}^A - 2\hat{u}_{i,j}^B)^2}{(2\hat{u}_{i,j}^A + 2\hat{u}_{i,j}^B)^2} = y_{i,j}(D2)$. To confirm H4 we expect $y_{i,j}(D1) < y_{i,j}(D2)$.

unknown at the outset. In stage I, the agent would be given the option to uncover costly information regarding the principal's utility function (e.g., about attribute importance weights) by selecting specific items to learn about. In stage II, the agent would need to enter a bid for a good on behalf of the principal and could choose to learn more costly items; subjects would be given examples of how related items may reveal correlated information. A subject's final payment is based on the utility the principal derives from the outcome minus the cost of information uncovered. One could then examine to what extent the individual chooses to learn new information that is correlated with previous information.

We also acknowledge several limitations of our results in support of H1 and H2. First, if knowing that a restaurant was previously preferred serves as a signal that spending time reveals useful information (which is equivalent to the subject perceiving that the cost of effort is itself lower if the restaurant was preferred), this might also explain the pattern in Table 2. Second, each time subjects in the CV condition made a stage II valuation decision they were reminded of their previous decision. This raises a concern that subjects spent more time in the CV condition because they had to process additional information. To rule out this alternative explanation, we ran a variant of the VO condition of Experiment 1 in which subjects were reminded in the valuation stage of irrelevant prior decisions they made regarding office supplies (e.g., a choice between a blue and a black pen). Running regression model (16) with this new data yields $\gamma_{CV} = 16.71$ ($p < 0.001$). Lastly, we note from Table 1 that subjects spent less time on each successive decision. This is true in *both* conditions and may be expected given subjects' increasing familiarity (or boredom) with the task as they made more decisions, or their growing desire to complete the study as they approached the final decision (and is consistent with prior research: Lenk et al. 1996; Slovic et al. 1965). Note though that our inclusion of decision dummies should partially tease out this effect and that the support for H2 was found in both Experiments 1 and 3. It is possible that had we used fewer decisions (say only three) and more subjects to compensate for less data points this issue would be less acute.

5 Conclusion

Many decision contexts surface relevant past decisions made by the same individual. When the decision at hand calls for costly effort to resolve uncertain aspects of utility, these past related decisions are likely to be invoked. In this paper, we have examined how prior decisions about single or multiple goods affect subsequent valuations using a simple model that incorporates features of both rational optimizing behavior and bounded rationality.

It may also be plausible to model individuals' behavior when making sequential decisions using some anomalous preference relation, such as an inherent preference for consistency. However, our theory predicts certain patterns of behavior that could not possibly be accounted for by such anomalous preference relations. For example, readily available information about a past choice increases the expected valuation spread between two goods. Furthermore, in expectation, an individual expends more effort on a decision when informed about a past decision. Subjects in our controlled lab experiments did, in fact, exhibit such behavior. Moreover, as the stakes increase, certain patterns of behavior generated by bounded rationality become even more pronounced both in theory and in our experiments. These findings suggest that models that take into account contemplation or thinking costs and recall of prior decisions may yield valuable insights about consumer behavior.

We believe our findings to have important implications for many business and managerial contexts. Salespeople often use a two-stage process that potentially affects final payment. For example, car dealers often refuse to discuss the final price of a vehicle before they have taken the consumer around the lot and prompted her to rank the set of relevant cars. Consistent with our findings, encouraging a ranking of alternatives (mainly on nonprice attributes) may work to the dealer's advantage in subsequent price negotiations on the buyer's most preferred car. The growing pervasiveness of Internet commerce has made the relevance of prior decisions all the more pronounced. In electronic auctions, for example, a consumer is typically shown a listing of relevant products. The consumer must decide which item(s) to bid on and in which order. She must then decide on her valuation for a given item when entering a bid. Our results imply that the bid entered for any item will be affected by integrating information about that item's preference ranking relative to the remaining alternatives in the set. Consider also a typical management approach when formulating a product development roadmap. It is common practice to begin by rank-ordering the various R&D projects to be funded, and only then to determine the exact allocation of resources to each (Parnell 2001). Our findings suggest that such a process would result in more time spent on determining the funding of the top ranked project and more variance in the resource allocations to all projects than if prior ranking had not taken place.

Appendix

A Parametric Specification for Information Acquisition. To see how all of our assumptions fit together, let the set of all aspects be $\mathbb{N}$, the set of natural numbers. Take any strictly increasing, continuous function g with $g(0) = 0$. Let the cost of learning the set $N \subset \mathbb{N}$ of aspects be $g\left(\sum_{n \in N} 1/2^n\right)$. Then, for any $c \in [0, g(2)]$, there will exist a set $N(c) \subseteq \mathbb{N}$ of aspects that costs exactly c to learn. (Note that $N(c)$ corresponds to the digits with a "1" in the binary expansion of $g^{-1}(c)$.) Hence, effort level can be chosen from a continuum, as in our model. To find a specific V that satisfies all our assumptions, including Assumption 1, take $g(x) = C^x - 1$ and $E\left[y_n^2\right] = D^{1/2^n}$ for some $C > D > 1$. Then, $V(c) = Var\left(\prod_{n \in N(c)} y_n\right) = D^{\sum_{n \in N(c)} 1/2^n} - 1 = (1+c)^\alpha - 1$, where $\alpha = \log(D) / \log(C) \in (0,1)$, for $c = C^{\sum_{n \in N(c)} 1/2^n} - 1$. Assumption 2 on F is satisfied when there exists a sequence (b_n) such that $|y_n - 1| < b_n$ for all n and $\prod_{n=0}^{\infty}(1 + b_n) < \infty$, which is consistent with the above specification. Finally, given any two sets N and N' of aspects, $Var\left(\prod_{n \in N} y_n\right) = Var\left(\prod_{n \in N'} y_n\right)$ whenever it costs the same to learn the aspects in either of the sets. Hence, as explained in Remark 3, under most distributions, the individual will optimally choose to learn about a new set of aspects that she has not thought about before, as in our model. ■

Proof of Proposition 1. To prove the first statement, we compute that

$$\frac{\partial}{\partial p_0} U_0(c_0; p_0) = \frac{\partial}{\partial p_0}\left[\int_{p_0}^{\bar{v}} (\hat{v}_0 - p_0) f(\hat{v}_0; c_0)\, d\hat{v}_0 - c_0\right] = -\int_{p_0}^{\bar{v}} f(\hat{v}_0; c_0)\, d\hat{v}_0 = F(p_0; c_0) - 1,$$

where f is the PDF of $F(\cdot; c_0)$. Hence, $\frac{\partial^2 U_0(c_0;p_0)}{\partial c_0 \partial p_0} = \frac{\partial F(p_0;c_0)}{\partial c_0}$. Then, by the condition stated in the proposition, for any $p_0 > 1$, we have $\frac{\partial^2 U_0(c_0;p_0)}{\partial c_0 \partial p_0} = \frac{\partial F(p_0;c_0)}{\partial c_0} < 0$. Therefore, $\arg\max_{c_0 \in [0,\bar{c}]} U_0(c_0; p_0)$ is decreasing with p_0, when $p_0 > 1$. Similarly, for any $p_0 < 1$, $\frac{\partial^2 U_0(c_0;p_0)}{\partial c_0 \partial p_0} > 0$, and $\arg\max_{c_0 \in [0,\bar{c}]} U_0(c_0; p_0)$ is increasing with p_0 in this region. To prove the second statement note that when v and p_0 are multiplied by λ, the expected payoff is proportional to

$$U_0(c_0; p_0)/\lambda = E[v - p_0 | \hat{v}_0 \geq p_0] \Pr(\hat{v}_0 \geq p_0) - c_0/\lambda.$$

Since $\partial^2 \left[U_0(c_0; p_0)/\lambda\right] / \partial\lambda \partial c_0 = 1/\lambda^2 > 0$, the optimal effort is increasing in λ. ■

Proof of Proposition 2. According to the mechanism for determining payoffs

$$U_1(c_1) + c_1 = E\left[\frac{2}{m}\int_0^{\hat{v}_1} (v - p)\, dp\right] = E\left[\left(2\hat{v}_1 v - \hat{v}_1^2\right)/m\right] = E\left[\hat{v}_1^2\right] E[2\varepsilon - 1]/m = E\left[\hat{v}_1^2\right]/m,$$

where the penultimate equality is due to the facts that $v = \hat{v}_1 \varepsilon$ and $\hat{v}_1$ and ε are independent. Since $V(c_1) = Var(\hat{v}_1) = E\left[\hat{v}_1^2\right] - 1$, this yields (5). ■

Proof of Proposition 3. In this proof we will omit the arguments of the functions V, V', V'', and V''', which will always be $h\left(m/x^2\right)$. Since $V'' < 0$ by assumption,

$$\Psi'(x) = 2xV + x^2 \cdot (V') \cdot \left(h\left(m/x^2\right)\right)' = 2xV - \frac{2m}{x^3 V''} > 0.$$

Thus Ψ is increasing in x. Toward showing that Ψ is convex in x, we can further compute that

$$\Psi''(x) = 2V + 2x \cdot (V') \cdot \left(h\left(m/x^2\right)\right)' + m \cdot \left(h\left(m/x^2\right)\right)'' = 2\left[V + \frac{(V')^2}{V''}\left(1 - \frac{2V''' \cdot V'}{(V'')^2}\right)\right]. \tag{20}$$

Since $V \geq 0$ and $V'' < 0$, this implies that $\Psi'' > 0$ whenever $V''' \geq \frac{(V'')^2}{2V'}$. But when $-V''/V'$ is nonincreasing, $V''' \geq \frac{(V'')^2}{V'} \geq \frac{(V'')^2}{2V'}$, that is, Ψ is convex. Write I for the set of all information the individual recalls, minimally including the events $[\hat{v}_0 \geq p_0]$ and $[\hat{v}_0 < p_0]$. If Ψ is convex, then

$$\begin{aligned} Var(\hat{v}_{01}) &= \int_{I\in\mathcal{I}} Var(\hat{v}_{01}|I) d\Pr(I) = \int_{I\in\mathcal{I}} \Psi(E[\hat{v}_0|I]) d\Pr(I) \\ &\geq \Psi\left(\int_{I\in\mathcal{I}} E[\hat{v}_0|I] d\Pr(I)\right) = \Psi\left(E\left[\hat{v}_0\right]\right) = \Psi(1) = Var(\hat{v}_1), \end{aligned}$$

where the inequality is due to the convexity of Ψ and Jensen's inequality. ■

Proof of Proposition 4. Define $\phi(\lambda; m) = E\left[\Psi(\hat{v}_0^\lambda; m)\right] - \Psi(1; m)$, so that $\phi(\lambda; m/\lambda) = Var(\hat{v}_{01}^\lambda) - Var(\hat{v}_1^\lambda)$. First note that

$$E\left[\Psi(\hat{v}_0^\lambda; m)\right] = \Pr\left(\hat{v}_0^\lambda < p_0\right)\Psi\left(E\left[\hat{v}_0^\lambda|\hat{v}_0^\lambda < p_0\right]; m\right) + \Pr\left(\hat{v}_0^\lambda \geq p_0\right)\Psi\left(E\left[\hat{v}_0^\lambda|\hat{v}_0^\lambda \geq p_0\right]; m\right), \text{ and}$$

$$\Pr\left(\hat{v}_0^\lambda < p_0\right) E\left[\hat{v}_0^\lambda|\hat{v}_0^\lambda < p_0\right] + \Pr\left(\hat{v}_0^\lambda \geq p_0\right) E\left[\hat{v}_0^\lambda|\hat{v}_0^\lambda \geq p_0\right] = 1. \tag{21}$$

When λ increases, identity (21) remains intact and $\hat{c}_0$ increases (by Proposition 1). This results in $E\left[\hat{v}_0^\lambda|\hat{v}_0^\lambda < p_0\right]$ decreasing and $E\left[\hat{v}_0^\lambda|\hat{v}_0^\lambda \geq p_0\right]$ increasing in λ. Since Ψ is convex, $E\left[\Psi(\hat{v}_0^\lambda; m)\right]$ increases and therefore $\phi(\lambda; m)$ is increasing in λ. Second, ϕ is decreasing in m. To see this, compute that $\frac{\partial\Psi(x;m)}{\partial m} = \frac{V'(h(m/x^2))}{V''(h(m/x^2))}$, and hence $\frac{\partial\phi}{\partial m} = E\left[V'\left(h\left(m/\left(\hat{v}_0^\lambda\right)^2\right)\right)/V''\left(h\left(m/\left(\hat{v}_0^\lambda\right)^2\right)\right)\right] - \frac{V'(h(m))}{V''(h(m))}$. Since $\frac{V'(h(1/z^2))}{V''(h(1/z^2))}$ is concave in z, by Jensen's inequality, we have $\partial\phi/\partial m \leq 0$. Now, given any λ and λ' with $\lambda \leq \lambda'$, we have

$Var(\hat{v}_{01}^{\lambda'}) - Var(\hat{v}_1^{\lambda'}) = \phi\left(\lambda'; m/\lambda'\right) \geq \phi\left(\lambda; m/\lambda'\right) \geq \phi(\lambda; m/\lambda) = Var(\hat{v}_{01}^\lambda) - Var(\hat{v}_1^\lambda)$. ■

Proof of Proposition 5. The unconditional expected effort in the valuation stage with information from a prior decision is $E\left[\hat{c}_1^I\right] = \int_{I\in\mathcal{I}} h(m/(E[v|I])^2) d\Pr(I)$, where I is the set of all information recalled. On the other hand, $\hat{c}_1 = h(m) = h\left(m/E[v]^2\right)$. It is straightforward to establish that when $-\frac{V''(h(1/x^2))}{V'(h(1/x^2))}x$ is nonincreasing in x, then $h\left(1/x^2\right)$ is convex in x. This proves the proposition by Jensen's inequality. ■

Proof of Proposition 6. We will use induction on $n = |P| - |Q|$ to show that $Var(\hat{v}_{01}[P]|Q_k) \geq Var(\hat{v}_{01}[Q]|Q_k)$ and $E\left[\hat{c}_1^I[P]|Q_k\right] \geq E\left[\hat{c}_1^I[Q]|Q_k\right]$ for each k, where the expectations are taken as conditional on $\hat{v}_0 \in Q_k$. For $n = 0$, we have $P = Q$, and the statements are trivially true. Assume that the statements are true for $n-1$. Let P' be a partition that is finer than Q but coarser than P, and that satisfies $|P| - |P'| = 1$, so that $|P'| - |Q| = n - 1$. Then, by the induction hypothesis $Var(\hat{v}_{01}[P']|Q_k) \geq Var(\hat{v}_{01}[Q]|Q_k)$ and $E\left[\hat{c}_1^I[P']|Q_k\right] \geq E\left[\hat{c}_1^I[Q]|Q_k\right]$ for each k. Since $|P| - |P'| = 1$, there exist k^* with $P_{k^*} = P'_{k'} \cup P'_l$ and $P_k = P'_k$ for each $k \neq k^*$. For any $k \neq k^*$, $Var(\hat{v}_{01}[P]|Q_k) = Var(\hat{v}_{01}[P']|Q_k) \geq Var(\hat{v}_{01}[Q]|Q_k)$ and $E\left[\hat{c}_1^I[P]|Q_k\right] = E\left[\hat{c}_1^I[P']|Q_k\right] \geq E\left[\hat{c}_1^I[Q]|Q_k\right]$. For k^*, by Proposition 3, we have $Var(\hat{v}_{01}[P]|Q_{k^*}) \geq Var(\hat{v}_{01}[P']|Q_{k^*}) \geq Var(\hat{v}_{01}[Q]|Q_{k^*})$. Similarly, by Proposition 5, $E\left[\hat{c}_1^I[P]|Q_{k^*}\right] \geq E\left[\hat{c}_1^I[Q]|Q_{k^*}\right]$. ■

Proof of Corollary 1. Let $\Delta_j = w_j\left(a_X^j\right) - w_j\left(a_Y^j\right)$. We obtain $E\left[(\hat{u}_X^{01} - \hat{u}_Y^{01})^2\right] = E\left[\left(\sum_j \Delta_j \hat{\delta}_{01}^j\right)^2\right]$ $= \sum_j \Delta_j^2 E\left[\left(\hat{\delta}_{01}^j\right)^2\right] + K = \sum_j \Delta_j^2\left(Var\left(\hat{\delta}_{01}^j\right) + 1\right) + K$, where $K = \sum_{i\neq j} \Delta_i \Delta_j$. When $-V_j''/V_j'$ is nonincreasing, since $Var(\hat{\delta}_{01}^j) \geq Var(\hat{\delta}_1^j)$ for each j the corollary immediately holds. The proof of Corollary 2 follows by dividing each squared valuation spread by λ^2 and taking the derivative with respect to λ. ■

References

Banerjee, A. V. 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107(3) 797-817.

Becker, G. M., M. H. DeGroot, J. Marschak. 1964. Measuring utility by a single-response sequential method. *Behavioral Science* 9(3) 226-232.

Bikhchandani, S., D. Hirshleifer, I. Welch. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy 100*(5) 992-1026.

Chen Y., G. Iyer, A. Pazgal. 2005. Limited memory and market competition. Working paper.

Conlisk, J. 1996. Why Bounded Rationality? *Journal of Economic Literature* 34 (June) 669-700.

Dow, J. 1991. Search decisions with limited memory. *Review of Economic Studies* 58 1-14.

Eliashberg, J. and J. R. Hauser. 1985. A measurement error approach for modeling consumer risk preference. *Management Science* 31(1) 1-25.

Engelkamp, J. 1998. Memory for actions. Hove, England: Psychology Press.

Ergin, H. 2003. Contemplation costs. mimeo, MIT.

Fischer, G. W., M. F. Luce, J. Jia. 2000a. Attribute conflict and preference uncertainty: Effects on judgment time and error. *Management Science* 46(1) 88-103.

Fischer, G. W., J. Jia, M. F. Luce. 2000b. Attribute conflict and preference uncertainty: The RandMAU model. *Management Science* 46(5) 669-684.

Gale, D. 1996. What have we learned from social learning? *European Economic Review* 40 617-628.

Greene, W. H. 2003. Econometric analysis. Prentice Hall, Upper-Saddle River, NJ 5th ed.

Gul, F. and R. Lundholm. 1995. Endogenous timing and the clustering of agents' decisions. *Journal of Political Economy* 103(5) 1039-1066.

Hirshleifer, D. and I. Welch. 2002. An economic approach to the psychology of change: Amnesia, inertia, and impulsiveness. *Journal of Economics and Management Strategy* 11(3) 379-421.

Ho, T-H. and K. Weigelt. 1996. Task complexity, equilibrium selection, and learning: An experimental study. *Management Science* 42(5) 659-679.

Hsee, C. K. 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.

Hsee, C. K., S. Blount, G. F. Loewenstein, M. H. Bazerman. 1999. Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin* 125(5) 576-590.

Khan, B. E. and R. J. Meyer. 1991. Consumer multiattribute judgments under attribute-weight uncertainty. *Journal of Consumer Research* 17(4) 508-522.

Lingle, J. and T. Ostrom. 1979. Retrieval selectivity in memory-based impression judgements, *Journal of Personality and Social Psychology,* 37 180-194.

Lipman, B. 1991. How to decide how to decide how to...: Modeling limited rationality. *Econometrica* 59(4) 1105-1125.

March, J. G. 1978. Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics* 9 587-608.

McCardle, K. F. 1985. Information acquisition and the adoption of new technology. *Management Science* 31(11) 1372-1389.

Mehta, N., S. Rajiv, K. Srinivasan. 2004. Role of forgetting in memory-based choice decisions: A structural model. *Quantitative Marketing and Economics* 2(2) 107-140.

Parnell, G. S. 2001. Work-package-ranking system for the Department of Energy's Office of Science and Technology. *Interfaces* 31(4) 109-111.

Payne, J., J. Bettman, E. Johnson. 1993. The Adaptive Decision Maker. Cambridge University Press.

Rabin, M. and J. L. Schrag. 1999. First impressions matter: A model of confirmatory bias. *Quarterly Journal of Economics* (February) 37-82.

Rubinstein, A. 1998. Modeling bounded rationality, MIT Press, Cambridge, MA.

Shugan, S. M. 1980. The cost of thinking. *Journal of Consumer Research* 7(September) 99-111.

Zwick, R., A. Rapoport, A. K. Chung Lo, A. V. Muthukrishnan. 2003. Consumer sequential search: Not enough or too much? *Marketing Science* 22(4) 503-519.

Table 1 Time Spent on Valuation

Decision	Experiment 1 (single good) VO (N=61)	Experiment 1 (single good) CV (N=63)	Experiment 3 (two goods) VO (N=37)	Experiment 3 (two goods) RV (N=39)
1	31.77	54.86	27.41	53.46
2	29.38	27.17	17.78	35.13
3	16.16	25.24	20.78	23.33
4	21.18	26.65	20.89	31.46
5	23.33	24.98	21.73	19.70
Average	24.36	31.78	21.72	32.62

Table 2 Time Requested for Valuation in Experiment 2 (N=50)

Decision	Subject's Choice in Stage I Restaurant	CD
1	9.62	7.68
2	10.64	6.39
3	8.43	7.07
4	8.36	6.92
5	7.48	6.84
Average	8.91	6.98

Table 3 Mean Absolute Difference in Valuations (in Dollars)

Decision Pair	Experiment 3 (dinner for one) VO Condition (N=37)	RV Condition (N=39)	Experiment 4 (dinner for two) VO Condition (N=44)	RV Condition (N=44)
1	2.97	4.10	4.02	7.75
2	2.97	4.00	4.59	6.91
3	3.23	4.31	4.84	7.59
4	3.40	3.54	4.36	6.23
5	3.75	4.31	6.11	8.48
Average	\$3.26	\$4.05	\$ 4.78	\$ 7.39

Figure 1: Effort-Accuracy Tradeoff Depending on Information from a Prior Decision