Expected Subjective Value Theory (ESVT): A Representation of Decision Under Risk and Certainty

Paul W. Glimcher[*]   Agnieszka A. Tymula[†]

August 21, 2020

Abstract

We present a novel descriptive model of choice that achieves an efficient representation anchored to how the brain represents value. An individual's behavior is fully described by two primitives: an individual's "reward expectation" and a free parameter we call "predisposition". We demonstrate the breadth of the model, showing that it captures many of the behavioral phenomena described by Prospect Theory through a novel reformulation of its primitives. The model accounts for individual heterogeneity in behavior, employs fewer parameters than Prospect Theory, and retains neurobiological plausibility as a causal model of the choice process. Our theory makes a series of novel behavioral predictions amenable to future testing and includes novel explanations of the endowment effect and the observed heterogeneity in the probability weighting functions.

1 Introduction

Decision scientists from various disciplines, including management, economics, psychology, have been working for centuries to create a model that tractably describes and correctly predicts behavior, and that is also easily generalizable and applicable. One principal difficulty faced by this program of research is that the processes that guide the chooser to arrive at

---

[*]Neuroscience Institute, New York University School of Medicine, 435 East 30th St, 11th Floor, New York, NY 10016, USA

[†]School of Economics, University of Sydney, A02, Sydney, NSW 2006, Australia. Phone: +61 2 9351 2840, Fax: +61 2 9351 4341. Tymula acknowledges the support from the Australian Research Council (ARC-DE150101032) while working on this project. Send correspondence to agnieszka.tymula@sydney.edu.au.

a decision are not observable. In the face of this problem, decision scientists have either derived their models from axioms that captured normative features of choice or they looked for mathematical frameworks that best described empirically observed data. Due to the development of widely available technology that now allows us to measure brain function in living organisms, a new approach that grounds the decision models in how the brain performs choice has become possible. By capturing the mechanism that implements decisions in the brain, this approach offers the promise of models that capture both behavior and mechanism. Such models may offer biological foundations for normative theories and shed a novel light on persistent theoretical puzzles. In this paper, we present a model that belongs to this biologically-and-behaviorally motivated class of decision models.

Fundamental to our approach is the recognition by computational biologists that feasible value functions are necessarily bounded and show finite precision due to the biological limitations of the nervous system (Platt and Glimcher, 1999, Glimcher, 2011). Information in the brain is conveyed by the rate with which neurons produce electrochemical impulses, called action potentials. The number of action potentials a neuron can produce is strictly bounded (by the laws of physics) and the number of neurons is demonstrably finite. Therefore, the subjective value function of reward value must necessarily be bounded. Further, the mechanism for generating action potentials is imperfect and hence introduces (largely thermodynamic) noise into any representation. As a result of these constraints, there is a meaningful limit on the information carried in any neural representation by the finite precision and range of the underlying representational system.

Working from this foundation, there is growing evidence that during the last decade neurobiologists have actually succeeded in identifying features of the basic representational form used by the human brain to encode decision variables for causally producing choice. This representation relies on a utility-like construct that is in principle observable, and takes as its natural units the number of *action potentials* produced by a group of neurons per second ("firing rate"). To distinguish it from the traditional mathematical concept of utility, following the literature in neuroeconomics, we refer to it as *subjective value*. Neuroscientists have long argued that the principles that guide this representational form ensure that information encoded by the nervous system is minimally redundant (Wainwright et al., 2002) and hence maximally informative, a notion that has recently gained traction in the decision sciences (Woodford, 2012).

In line with these arguments, Steverson et al. (2019) have shown that a specific class of utility-like functions identified by neuroeconomists allows agents to optimally balance the known biological costs of reducing stochasticity (or equivalently increasing choice precision) against the gains from more accurately representing the utility of options. Unlike other models, these divisive normalization models can thus be shown to be efficient in the sense that they optimally balance biological costs and benefits. Steverson et al. (2019) also demonstrate

1

that this utility-like function imposes uniqueness on the representation, making it possible to derive neurological predictions from choice and vice versa. And perhaps more importantly for economists, recently published papers have shown that this representational form predicts several interesting human and animal choice behaviors that had previously escaped the notice of economists (e.g. Louie et al. (2013), Khaw et al. (2017), Guo and Tymula (2020), Landry and Webb (2018), Webb et al. (2020)).

The standard form of the divisive normalization model has the following key features: 1) it employs a cardinal utility-like construct (aligning well with Steverson et al. (2019)'s observation of full uniqueness); 2) is numerically bounded and has finite precision (aligning with Steverson et al. (2019)'s observation of efficiency); and 3) it adjusts dynamically to changes in the reward expectation, like the recent range normalization models (Kontek and Lewandowski, 2018), in this case via demonstrably efficient divisive functional form.

The model also aligns well with recent work that has highlighted the implications of representational limits of this kind for decision making (Kontek and Lewandowski, 2018, Steverson et al., 2019, Frydman and Jin, 2019, Rayo and Becker, 2007, Netzer, 2009, Glimcher, 2010, Woodford, 2012, Robson and Whitehead, 2016, Webb et al., 2020, Payzan-LeNestour et al., 2020). One principal implication of these kinds of representational limits is that efficient subjective value representations must be dynamic, adjusting to the mean and range (or variance) of encountered choice sets. Kontek and Lewandowski (2018), for example, have recently formulated a range-dependent utility theory that precisely captures an efficient mechanism for utility encoding in the face of such limits. Woodford (2012) provided yet another such analysis. Both of these papers argue that with any bounded and imprecise representation it is optimal for individuals to evaluate options relative to reward expectations and not in absolute terms. The foundational neurobiological model that we adopt in this paper shares this critical property while preserving the normative features of Steverson et al. (2019)'s analysis and a strict adherence to the neurobiological mechanism which produces choice.

Our focus in this presentation is to explore the critical functional properties of a novel-to-decision-science functional form that has begun to receive significant attention in neuroscience. We stress that this functional form is uncontroversial in neuroscience (Heeger, 1992, Schwartz and Simoncelli, 2001, Carandini and Heeger, 2012, Reynolds and Heeger, 2009), is efficient (Steverson et al., 2019, Wainwright et al., 2002), and is grounded in years of study which we review in the next section. After that review, we present the general form of the neurobiological model that accounts for efficiency in value coding imposed by other alternatives in the choice set as well as previously experienced alternatives. For tractability, we restrict our attention to the case of a singleton choice set, the valuation of a single risky or certain alternative.

With this approach, we find that the efficient-neurobiological model of choice (its singleton valuation form), ESVT, reformulates Prospect Theory in a novel way that we believe

2

will be of general interest. Our goal in this paper is: 1) To relate this functional form to Expected Utility Theory, Prospect Theory and some of its closely-related descendants; 2) To examine the specific choice theoretic features of the model from a decision scientist's point of view; and 3) To briefly compare the model's ability to account for behavior with standard Prospect Theory.

Before beginning our formal analysis, we note that ESVT assumes only two primitives: *predisposition* and reward *expectation*. Unlike most economic models of choice, attitudes towards risk are *not* primitives of models in this class. Instead, risk attitudes emerge from and vary with changes in individual expectations and as a function of predisposition. Our primary finding is that this reformulation preserves much of the representational structure of Prospect Theory. It accounts for choice behavior at least as well as Prospect Theory and does so with fewer parameters and in a way that closely aligns with the actual physical mechanism by which choice is empirically produced inside human beings. ESVT also, however, makes novel predictions about behavior that extend beyond the domains Prospect Theory was intended to describe. For example, it makes theoretical predictions about the sources of heterogeneity in individual-level probability weighting functions that have not yet been well predicted by any theory and provides a completely novel explanation of the endowment effect which relies on changes in reward expectation rather than loss aversion.

In the next section, we develop the neurobiological foundations of our functional form for readers interested in the model's early origins and recent developments. Additional information about these foundations can be found in Glimcher (2011). In section 3, we describe the model and give an intuition about the role of each of the two parameters. In section four, we present the results of our mathematical analysis of the model. Here, we set ourselves the task of asking: Given that human choosers represent the values of rewards in their brains using something very close to this class of transformation, and then perform an argmax operation on these *subjective values*, what kinds of preference structures would one expect to see in choice behavior? Would one see the same phenomena Kahneman and Tversky described? Furthermore, could such a representation be a refinement of Prospect Theory, capturing behavior more accurately, under at least some conditions, because it more accurately captures the biological mechanism of choice? In Appendix A we present a preliminary empirical test of the model with a dataset consisting of the risky choices of 2,530 individuals made in a set of gain-only lotteries. There we compare the model fit to that of Prospect Theory and Expected Utility.

2 Neuro Foundations: Normalization Model-Based Representations of Value

The study of how information is represented in the brain largely begins with the study of how sensory systems (the neurobiological mechanisms underlying sight, touch, hearing, taste, and smell) encode and represent properties of the outside world (Ratliff and Hartline, 1959, Barlow, 1961, Attneave, 1954). Nerve cells convey information by "firing rate," the rate at which they produce transient electrochemical impulses called action potentials which they communicate to cells with which they are connected. Three key properties of neural activity will be crucial in understanding the logic and efficiency of the normalization mechanism. First, the firing rate of a neuron is biophysically bounded by a maximum rate at which the neuron can produce action potentials, typically about 100 impulses per second. This, combined with the fact that we have a limited number of neurons in the brain and that a subset of those is devoted to encoding stimuli of different types, means that for any given stimulus type, the absolute maximum brain activity that can be allocated to a representation (measured as the total number of action potentials) is limited. Second, neuronal firing rates are stochastic, they exhibit (thermodynamic) variance which limits the precision with which firing rates can be interpreted. Third, neural activity is expensive in the caloric sense, thus increasing the precision of a representation is metabolically costly. These three properties imply that an efficient representation of stimuli must tradeoff precision against its costs, as we explain below.

2.1 Static framework

To begin to understand the costs and benefits of representation, consider how the brain might represent a static image of a sunset presented on a black-and-white computer monitor's 1,000 x 1,000 grid of 1,000,000 pixels. One obvious possibility (that is nevertheless not the case in reality) might be that a single nerve cell in the brain's visual system would be devoted to conveying value for each of the 1,000,000 pixels, thus requiring 1,000,000 nerve cells to convey the content of the picture. To make this example concrete, we might imagine that each nerve cell would represent that its pixel was black by producing (or firing) one action potential per second and that its pixel was white by firing at a biophysically maximum rate of 100 action potentials per second. Intensities (I) between these two extremes would be represented by a continuous monotone function (F). In such a situation a perfectly white computer screen would cause all 1,000,000 neurons to fire 100 action potentials per second. In this way each pixel is associated with an independent neuron that has an independent firing rate and no information about the pixels in the display is lost by this transformation to firing rates.

4

What neurobiologists realized in the 1960s and formalized in the 1990s, however, is that adjacent pixels in real-world images are non-independent. To put that in more formal terms we can describe each image as a set of 1,000,000 pixels with the intensity of each pixel i labeled as $I_i \in \mathbb{R}_+$. The set of all *possible* images forms set A. What studies of pixel-by-pixel correlations in natural images reveal is that the set of all *observable* images, B, is much smaller than A ($B \subset A$) and is marked by high (approximately 0.7) adjacent pixel correlations (Simoncelli and Olshausen, 2001, Simoncelli, 2003). This non-independence is an incredibly important point. Consider the nine adjacent pixels shown in Figure 1. If one knew in advance that the eight outer pixels were white and the adjacent pixel correlation was 0.7 (a typical estimate), one would know with better than 99.9% certainty that the central pixel was also white. Under these conditions, therefore, the activity of the neuron encoding that central pixel is almost entirely redundant.
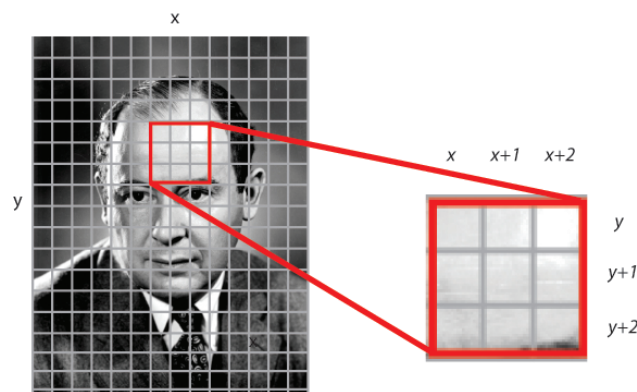

Figure 1   An array of local pixels from a photograph.

The reason this is so important is that action potentials and neurons that generate them are incredibly costly in terms of calorie consumption. Nearly 20% of the calories we consume go to support our brain even though it accounts for only about 3% of our body weight. Brain tissue is on average roughly seven times as expensive as the average tissue in our bodies.

Faced with these high costs, a representational form that uses fully independent neuronal firing rates to encode non-independent pixels intensities is clearly inefficient, in practical terms wasting much more than half of its metabolic expenditures. Recognizing this fact, neurobiologists working in the late 1990s and early 2000s began to ask what forms of representation would maximize Shannon (or Fisher) Information (Shannon and Weaver, 1949) in the aggregate of action potential rates conditional on the correlational structure of the observed sets B (Heeger, 1992, Heeger et al., 1996, Reynolds and Heeger, 2009, Schwartz and Simoncelli, 2001, Wainwright et al., 2002). The general answer to this question is that neurons employ a representation of the general form:

5

$$F_i \propto \frac{I_i^\alpha}{M^\alpha + \sum_j w_{i,j} I_j^\alpha}$$

where F_i is the firing rate of the neuron encoding the intensity of pixel i. I_i is the objective intensity of pixel i and $w_{i,j}$ is a weight which captures the empirically measured ex-ante correlations between the pixel at location i and the pixel at location j. α is a free parameter. M is an expectation over the intensity levels in the image which we return to in section 2.2 when we introduce a dynamic setting. In essence, the denominator is serving to remove from the firing rate of neuron i information already carried by other neurons, thus maximizing information per action potential.

Throughout the preceding discussion, neuronal firing rates have been treated as deterministic. If the normalization equations map intensity to firing rate along a segment of $\mathbb{R}_+$ between 0 and 100, in what sense can we think of the information-bearing capacity of these neurons as limited? Given infinite precision, the restriction that the firing rate of the neuronal set is bounded seems trivial. Firing rates are, however, non-deterministic and roughly Poisson-distributed[1] with mean firing rate and variance highly correlated . This property makes the limited capacity of the system hugely problematic in finite time. If the entire set of neurons encoding an image has only an average aggregate action potential rate of 100 to spend and each element (neuron) has high variance proportional to its mean, then the information carried by the set is quite limited.

For a decision scientist, it is probably useful to think about biological representations of stimuli through the lens of random utility models where each neuron encodes a mean value plus a variance term. More on this approach can be found in Webb (2020).

2.1.1 Evidence in decision-making

For neurobiologists studying how we see, there are tools for assessing the independence of adjacent pixels in huge sets of existing images. Using such an approach to the set of all observed consumer choice sets is problematic. Fortunately, however, another approach is possible. For a neuroeconomist, subjective values are by definition both causally responsible for choice and directly observable (unlike the utilities to which they are related). Using a variety of neurobiological tools, it is now possible to measure directly, or to infer quite precisely, the activity levels of neurons encoding the subjective values of options experimentally presented to choosers. Given that this is possible, rather than asking whether choice sets have correlational structure, one can ask whether the representation of reward value in the brain follows a normalization-based representation.

The first effort to do this was undertaken by Louie et al. (2011) who studied the activity

---

[1]The actual distribution has been very well studied and is an extreme type-II distribution with interesting properties beyond the scope of this discussion.

of single neurons in monkey subjects making choices over two- and three-element choice sets. The structure of those choice sets was systematically varied in a way designed to reveal normalization in value representation if it occurred.[2] Louie et al. (2011) found that the firing rates of the neurons encoding each element in a choice set were non-independent and were extremely well described with normalization-type models. From these findings, they concluded that subjective value representations appear to be normalized in exactly the way one would expect if choice sets they sought to represent efficiently showed deep correlational structure.

Once it was clear that the neural instantiation of subjective value appeared to follow a normalized form of representation, some scholars began to ask whether these normalized representations made any unique behavioral predictions. Speaking broadly, these theoretical investigations have suggested that normalized representations should lead to very specific violations of the axiom of regularity (Becker et al., 1963), a conclusion that has been tested successfully in both monkeys and humans in choice sets of varying size and type (Louie et al. (2013, 2015), Webb et al. (2020), Robinson and Tymula (2019), but see Gluth et al. (2020)). The effects of choice set size and type on stochasticity in choice have also been examined theoretically and empirically in riskless (Louie et al., 2013) and risky choice (Robinson and Tymula, 2019) and once again these studies have found evidence that normalization class representations can go far in explaining many classes of behavior.

To summarize this literature, one can draw both the conclusion that neural representations of subjective value follow the predictions of the normalization-form models and the conclusion that the unique behavioral predictions made about normalization-class representations are observed.

2.2   Dynamic framework

To continue with our analogy to vision, the same way that visual properties of adjacent pixels in real-life images show a high correlation, their properties are also highly correlated across time, meaning that currently experienced stimuli, for example, levels of brightness, carry information about the brightness levels we should expect in the next second. Such temporal dependency of firing rates has long been recognized in neuroscience. For example, Ohzawa et al. (1985) measured firing rates in the cat's visual cortex and found that the same objective levels of contrast lead to a higher firing rate if the cat was previously shown low levels of contrast than when the cat was previously shown high levels of contrast. The observed patterns of firing rates in the study by Ohzawa et al. (1985) are consistent with the

---

[2]They performed these initial measurements on monkeys because it allowed the use of invasive technologies that yield measurements of subjective value virtually without measurement error. Similar studies have subsequently been made in humans (Hunt et al., 2014, Webb et al., 2020) using techniques with much lower signal to noise ratios, and similar results have been obtained.

representational form we introduced in the previous section. There, we briefly defined M to be the expectation. In this section, we describe what is known about its origins in neural activity.

How networks of neurons biophysically generate the transformation specified by the normalization equation has also received significant attention providing important insights into the normalization equations themselves at a positive and at a normative level. Current evidence suggests that each reward in a choice set is represented neurobiologically by what can be schematized (perhaps too) concisely as two dynamically interacting neurons: an *excitatory* and an *inhibitory* neuron. For simplicity, we now focus on the representation of the subjective value of a single reward in a degenerate one-element choice set, a set-up that will be our primary focus in this paper. In the existing literature, e.g. LoFaro et al. (2014), the function that represents action potential production in response to objective input x is presented in the excitatory neuron as the function R and in the inhibitory neuron as the function G. These functions are described by the following first-order differential equations:

$$\tau \frac{dR}{dt} = -R + \frac{x}{1+G}$$

$$\tau \frac{dG}{dt} = -G + R$$

where τ is the time constant of neuronal information integration and x is the objective value of the reward. LoFaro et al. (2014) showed that in networks of this type, for any given x the network always has a unique equilibrium state:

$$R = \frac{x}{1+G}$$

There are two notable aspects of this equation. The first is that the unique equilibrium state of the neuronal computation is essentially equivalent to a standard normalization equation. The second is that the normalization emerges spontaneously at equilibrium from the formal integration of the value inputs over time by the system. Applying some additional simplification (LoFaro et al., 2014) we can represent the action potential rate at time t as:

$$R_t \propto \frac{x_t}{\sum\limits_{k=0}^{t-1} D(k)x_k + x_t}$$

where $D(k)$ is a weighting function such that the term on the left in the denominator $(\sum\limits_{0}^{t-1} D(k)x_k)$ is a time-discounted average of the rewards previously encountered by the system. In the language of modern economics, the reward expectation is based on previously encountered rewards. Some may be willing to call it a reference point. In essence, just as in our exposition of the static framework, the denominator is serving to remove from the firing

rate of a neuron, the information that has been carried by the neurons in the past.

2.2.1 Evidence in decision-making

Empirically, normalization functions have been used to model the time-varying firing rates of subjective value encoding neurons in a way that goes beyond the statics described earlier (Louie et al., 2014, LoFaro et al., 2014). Neurons in the subjective value-representing areas (Louie et al., 2014, Padoa-Schioppa, 2009) are well described by these dynamic equations.

This suggests that these equations might also be used to model the dynamic effects of expectation changes on human choice behavior. Two recent behavioral studies have assessed whether the current valuation of the same snack item (Khaw et al., 2017) or the same monetary gamble (Guo and Tymula, 2020) can be influenced by inconsequential exposure to rewards in the past as predicted by the normalization model. Both of the studies found evidence that is consistent with normalization — an individual's valuation of the same reward is lower after exposure to high-valued rewards than after exposure to low-valued rewards. The effect occurs even though the valuations are elicited using an incentive-compatible mechanism and the comparison is within participants who are providing valuations for the same rewards within the same experimental session.

2.3 Summary

In this literature review, we have provided insight into the origins of the functional form we examine in this paper. Existing theoretical work in economics has shown that this class of functional forms has normative qualities and is equivalent to reducing Shannon entropy in terms of implied behavior (Steverson et al., 2019). Studies of how the brain actually represents subjective value in the circuits causally responsible for choice show that these equations do a better job of predicting steady-state neuronal firing rates than alternatives (Louie et al., 2011, 2015, Yamada et al., 2018). Econometric studies of representations employing these equations suggest a number of unique behavioral patterns that would be the product of representations having these features and many of those properties have now been observed in humans and animals (Louie et al. (2013), Webb et al. (2020), Louie and Glimcher (2012) but see Gluth et al. (2020)). Finally, dynamic versions of these equations have also been developed (LoFaro et al., 2014), used to predict dynamic neuronal firing rates, and used to predict novel choice dynamics then confirmed in the data (Louie et al., 2015, Khaw et al., 2017, Guo and Tymula, 2020).

An interesting feature of this approach is that it raises the possibility that many behaviors that violate Expected Utility may not be failures to maximize as previously thought. Instead, some violations of Expected Utility might, in fact, be formally rational and a result of efficient trading of the marginal benefits against the marginal costs of additional precision in choice.

To date, most of the research on the normalization of subjective value has focused on the static framework, that is on how the elements of the current choice set affect valuation. Far fewer papers have focused on the dynamics of value normalization, on how previously experienced rewards affect current valuation, even though the neuroscientific basis of such temporal adaptation are well understood in sensory systems (Louie et al., 2015). Most of the work has also focused on riskless rather than risky choice. Our goal in this paper is to expand the literature by providing a tractable theoretical framework based on which novel and testable behavioral hypotheses can be derived and then empirically (neurobiologically and behaviorally) tested both in the riskless and risky settings. We also see our analysis as a necessary step towards a comprehensive model that incorporates both types of normalization — normalization by the value of the previously experienced rewards as well as normalization by the value of the rewards in the current choice set.

3   Model

We focus in this presentation on the dynamic aspect of normalization by modelling a value of the current option in a singleton choice set that is normalized by a reward expectation determined by a stream of previously experienced rewards. Following the literature summarized in section 2, we assume that the subjective value (the "decision value"(Kahneman and Tversky, 1984, Hare et al., 2008)) of a reward i at time t, $x_{i,t} \in \mathbb{R}_+$, is given by:

$$S_t(x_{i,t}) = \frac{x_{i,t}^{\alpha}}{x_{i,t}^{\alpha} + M_t^{\alpha}} \tag{1}$$

where M_t is the reward expectation at time t based on the stream of rewards experienced until $t-1$ and $\alpha > 1$ is a free parameter we call *predisposition* to differentiate it from traditional preference parameters.

The model is restricted to the valuation of a reward in a singleton choice set. We make this simplification to gain a clear understanding of how dynamic normalization affects subjective valuations, in the absence of the normalization imposed by the context of a choice set. Furthermore, we restrict our model to nominal gains as there is significant controversy over the neural representation of losses (Tom et al., 2007, Knutson et al., 2008, Samanez-Larkin and Knutson, 2015). Despite these two restrictions, the model still generates novel and testable predictions that shed light on some of the current puzzles in behavioral economics.

The subjective value function takes values between 0 and 1 ($S \in [0,1]$) consistent with the idea that decision makers are bounded in the range of values that they can biophysically assign to the rewards. Unlike traditional utility (but following Steverson et al. (2019)), this subjective value function is cardinal. Its natural unit is *firing rate*, that is the number of action potentials per second generated by neurons representing the subjective value of an

alternative, here projected onto the unit line segment for simplicity.[3] Neuronal firing rates map linearly to purely behavioral estimates of subjective value, as has now been repeatedly observed (Louie and Glimcher, 2010, Kable and Glimcher, 2007, Plassmann et al., 2007, Smith et al., 2014). The subjective value of $x_{i,t}$ is divisively normalized by the expectation and itself. The fact that it is normalized by itself guarantees that the subjective value function always adjusts to the problem at hand. In forms that address larger choice sets, all elements of the choice set are included in the denominator (Glimcher, 2015, Webb et al., 2020), though it is a complication we forego in this presentation for analytic tractability. An interested reader can refer to Robinson and Tymula (2019) for theoretical and experimental treatment of choice-set level normalization and risk attitudes and to Webb (2020), Webb et al. (2020) for econometric analyses of choice set effects in these models.

The expected subjective value of a lottery (x, p) with possible rewards $x_1, x_2, ..., x_n$ that occur with corresponding probabilities $p_1, p_2, ..., p_n$ is calculated by taking a standard expectation:

$$ES_t(x_{i,t}, p) = \sum_{i=1}^{n} p_i S_t(x_{i,t}) \tag{2}$$

3.1 Reward Expectation

We use existing neurobiological data[4] on how expectations are computed in the human brain to select an initial (although admittedly imperfect) functional form for reward expectation, M_t. A large body of neuroscientific evidence has documented the physical existence of a dopaminergic reward prediction error system in the brain (for a review see Glimcher (2011)). Axiomatically (Caplin and Dean, 2008) and empirically (Rutledge et al., 2010), this system has been shown to represent an experience-based expectation of future rewards and punishments. Following this work, our model employs a simple recursive computation that calculates a recency-weighted average of past rewards. Computation in this brain system is typically modeled using discrete time where at each time step t, the brain computes the reward expectation P_t as:

$$P_t = P_{t-1} + \gamma(x_{i,t} - P_{t-1}) \tag{3}$$

---

[3]To transform S into true action potential units, it is typically multiplied by 100, the true maximal action potential rate of most neurons.

[4]In the neuroscientific literature a tremendous amount of directly observed data is available on expectations (Ohzawa et al. (1985), Bayer and Glimcher (2005), Khaw et al. (2017)). Detailed differential equation-based models of the brain mechanism that computes and represents expectations are now available (Wilson and Humanski, 1993, Carandini and Heeger, 2012, Watabe-Uchida et al., 2017) as are models of how the recent history of offers alters choice and valuation in decision making (LoFaro et al., 2014, Khaw et al., 2017).

where $x_{i,t}$ is the reward delivered to the subject at time t and $\gamma \in (0,1)$ sets the forgetting rate. This iterative computation, performed in known brain circuits, is equivalent to the following function (although the iterative form above imposes simpler requirements for data storage in the brain):

$$P_t = (1-\gamma)^t P_0 + \sum_{\tau=1}^{t} \gamma(1-\gamma)^{t-\tau} x_{i,\tau} \tag{4}$$

Following the early reinforcement learning literature, *for simplicity* (Sutton and Barto, 1998) we assume that initial expectations are equal to zero and can thus write the reward expectation (M_t) as:

$$M_t = P_{t-1} = \sum_{\tau=0}^{t-1} \gamma(1-\gamma)^{t-1-\tau} x_{i,\tau} \tag{5}$$

As can be seen, this mechanism computes a time-weighted average of previous reward magnitudes and can be fit with one free parameter γ. The forgetting rate, $\gamma \in (0,1)$, captures the observation that more recent rewards have a bigger impact on the reward expectation than older rewards, distinguishing our model from the models in which rewards are evaluated relative to the current level of wealth instead.

We note critically that this widely used reward expectation is relatively simple (too simple for many environments) because it derives from already well understood (and thus necessarily simple) neurobiological mechanisms. One can say that in these models current and past experiences and no other knowledge (like information about foregone gains or sudden changes in context) impact the reward expectation. Of course, we know that the reward expectation (both behavioral and neurobiological) is much more complicated than this implies (e.g. Tymula and Plassmann (2016), Heffetz and List (2014), Khalil and Wu (2017)) and that other factors influence expectations, but this approach supplies reasonable tractability and predictive power.

3.2 Predisposition (α)

Predisposition is the critical exogeneity that influences choice. When combined with an expectation it specifies behaviorally observed preferences. The value of predisposition is not yet normatively constrained, although there are some biological hints about the relationship between predisposition, the brain during aging (Chung et al. (2017)), and perhaps choice set size (Webb et al. (2020)). More theoretical and empirical work (now ongoing) is, without a doubt, necessary to understand the normative foundations of predisposition and whether we should think about it as an individual-specific or environment-specific parameter. But from a neuroscientific and behavioral point of view, we know that including predisposition

12

as a free parameter is necessary at this stage. The inclusion of this term drastically improves empirical model fits to both behavior and neural data while we work to endogenize this parameter. In neural measurements it is typically estimated to be between 1.0 and 3.5 (Albrecht and Hamilton, 1982, Sclar et al., 1990, Busse et al., 2009, Carandini and Heeger, 2012). It is also a critical source of heterogeneity in the model.

3.3 Model Illustration

Figure 2 plots examples of subjective value functions over gains for a range of predisposition values and expectations. The model can produce a variety of familiar value function shapes: a value function concave for all $x_{i,t}$ that looks like the utility function introduced by Bernoulli, as well as the more modern S-shapes used in Prospect Theory (Kahneman and Tversky, 1979) and other reference-dependent models.
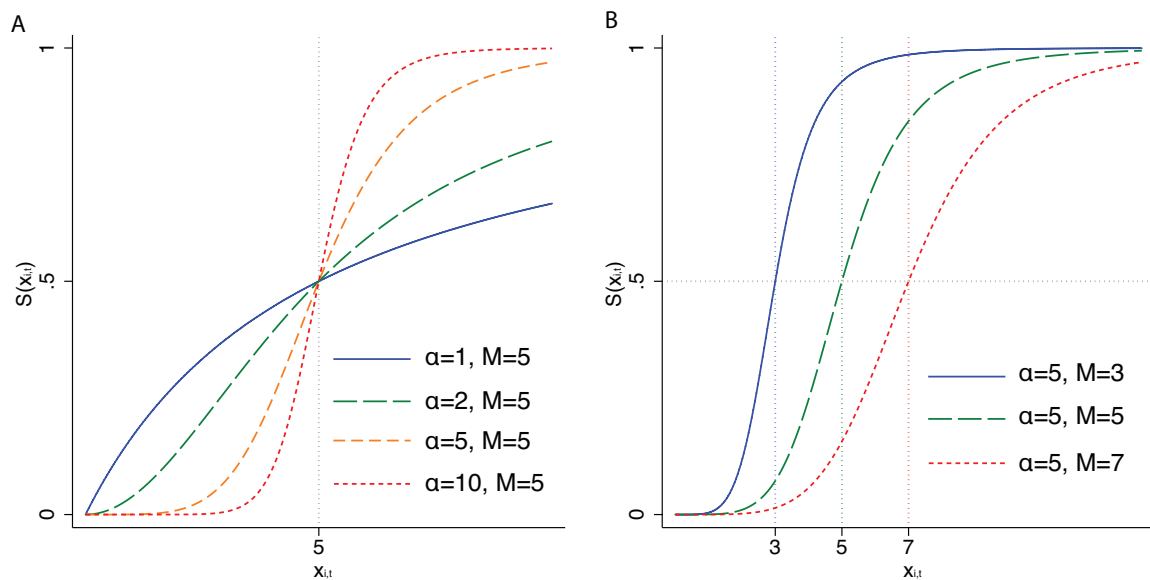

Figure 2 Examples of subjective value functions. In each graph, the vertical dashed lines are drawn at $x_{i,t} = M_t$. A: The effect of changes in the predisposition on the subjective value function. B: The effect of changes in the reward expectation (M_t) on the subjective value function.

To give the reader a quick overview of the role of predisposition and expectation in the subjective value function, in Figure 2 A we manipulate predisposition while keeping the reward expectation constant, and in Figure 2 B we manipulate the reward expectation while keeping the predisposition constant.

In Figure 2 A we see that predisposition affects the curvature of the value function. When predisposition is low ($\alpha = 1$, blue solid line), the subjective value function is concave for all x and thus the decision maker is always risk-averse. Under these conditions, the subjective

value function looks very much like the utility functions that dominated economics before the development of reference-dependent utility theories. As α increases, the subjective value function begins to take on the familiar S-shape from Prospect Theory — it starts as a convex function and then at some point changes to concave.

To illustrate the role of the expectation, Figure 2 B plots three different value functions with different expectations and the same predisposition. As the expectation increases, the subjective value function shifts to the right as in many reference dependent models. Importantly, the shift is such that at the expectation the function always takes the same value, $S_t(M_t) = 0.5$. This can be easily checked by substituting M_t in place of $x_{i,t}$ in Equation 1. For every α, whenever $x_{i,t} = M_t$, then $S_t(x_{i,t}) = 0.5$. This property is consistent with behavioral and neural evidence that people evaluate the rewards and other stimuli as departures from the expectation rather than in absolute terms.

It is also important to note that the subjective value function does not shift in a strictly parallel fashion. As it shifts to the right the overall slope of the function decreases. This has the positive feature of preserving diminishing marginal value (a point we return to later). We take this as a significant feature, distinguishing this class of models from models where the reward expectation shifts the value function rigidly.

4 Model predictions

In this section, we establish the relationship between ESVT's primitives, *predisposition* and *expectation*, and the critical behavioral phenomena captured by the model.

4.1 Risk Taking

To understand the role that expectation and predisposition play in risk-taking, we analyze the curvature of the subjective value function and the Arrow-Pratt index of relative risk aversion (RRA) as a measure of local curvature:

$$RRA_t = -\frac{x_{i,t}S''}{S'} = \frac{(1+\alpha)x_{i,t}^\alpha - (\alpha-1)M_t^\alpha}{x_{i,t}^\alpha + M_t^\alpha} \tag{6}$$

The index can take both positive (local risk aversion) and negative (local risk seeking) values. Given that $S' > 0$ and $x_{i,t} \in \mathbb{R}_+$, the conditions for the RRA to be positive and the subjective value to be concave (and for RRA to be negative and the subjective value to be convex) are equivalent. In traditional reference-dependent models, individuals are assumed to be risk-averse (seeking) whenever the rewards are larger (smaller) than the reference point. The inflection point in our subjective value function, which determines the switch from convexity to concavity as well as the change in the *local* risk attitude, is not only determined by the reward expectation, but also by predisposition.

14

Theorem 1. *The individual with ESVT preferences is locally risk averse and has a concave subjective value function for*

$$x_{i,t} > \sqrt[\alpha]{\frac{\alpha - 1}{\alpha + 1}} M_t \qquad (7)$$

An individual is risk seeking and has a convex subjective value function for $x_{i,t} < \sqrt[\alpha]{\frac{\alpha-1}{\alpha+1}} M_t$.

Proof. Given that $S' > 0$ and $x_{i,t} > 0 \; \forall x_{i,t}$, the proof follows from the definition of Arrow-Pratt relative risk attitude index (Equation 6). ∎

4.1.1 Adaptive Reward Expectation and (In)stability of Preferences

Since the expectation, M_t, varies in ESVT from period to period, the individual risk attitude for a given $x_{i,t}$, as determined by Equation 6 and Theorem 1, can change dramatically over time. This observation reveals that ESVT interacts with a contemporary discussion on the stability of preferences. Economists and psychologists have traditionally taken different views on the issue of preference stability, with psychologists largely arguing against stability (Slovic, 1995) and economists largely wishing for stability (Stigler and Becker, 1977). In this model, the fact that the expectation changes over time as the individual experiences different rewards implies that an individual's willingness to accept a given risky option will also change over time. This change is, however, determined by an individual's reward history and forgetting rate. The model, therefore, walks a line between traditional psychological and more economic approaches to decision-making. It yields a varying risk tolerance due to what could be construed as a rational contextual effect without discontinuities in the overall preference structure.

To derive predictions about how local risk attitude changes due to dynamic normalization, we first establish that the local Arrow-Pratt index of relative risk attitude declines when the reward expectation increases.

Theorem 2. *Under ESVT, the Arrow-Pratt index of relative risk attitude declines when the reward expectation increases.*

Proof. To show: $\frac{\partial RRA_t}{\partial M_t} < 0$

$\frac{\partial RRA_t}{\partial M_t} = -\frac{2\alpha^2 M_t^{\alpha-1} x_{i,t}^\alpha}{(x_{i,t}^\alpha + M_t^\alpha)^2} < 0$ which completes the proof. ∎

Thus, the role of reward history in explaining risk attitudes in ESVT goes beyond the pure wealth effects that economists have modeled for centuries. It is not only the total income earned but also the timing of when each reward was received as well as the individual's forgetting rate, which determines willingness to take risks. This is illustrated in the next two propositions.

15

Proposition 1. *Assume two individuals who have the same predisposition (α), the same forgetting rate ($\gamma \in (0,1)$), experienced an identical set of rewards in the past $\sum_t x_{i,t-1}$, and face an identical risky gamble. The individual who received the past rewards in an order such that they increase in value from one period to the next ($x_{i,0} < x_{i,1} < ... < x_{i,t-2} < x_{i,t-1}$) will be more risk tolerant in their valuations than the individual who received the same rewards but in a sequence that decreases in value from one period to the next ($x_{i,0} > x_{i,1} > ... > x_{i,t-2} > x_{i,t-1}$).*

Proof. The proof follows from the fact that the individual who received rewards in an order that is increasing in value has higher reward expectation. ∎

To understand this proposition, imagine two individuals in period t, Oskar and John, who have the same predisposition, α, the same forgetting rate γ, and have earned the same total income until period t. Oskar initially earned a lower income than John, but then his income continued increasing over time. The opposite happened to John. Since their total accumulated wealth is the same in period t, Expected Utility Theory, Prospect Theory, and Range-Dependent Utility Theory would all predict that Oskar and John are going to have the same risk preference. ESVT, however, makes a different prediction. Since the more recent events have a larger impact on M_t, Oskar will have a higher reward expectation. We can now use Theorem 2 to predict that in period t Oskar will be more risk-tolerant.

Another prediction of the model is that the individuals who have the same reward history can differ in their attitude to risk if they have different forgetting rates as established in the next proposition.

Proposition 2. *Two individuals with the same predisposition (α) who have experienced the same stream of rewards in the past will exhibit different risk attitudes if their forgetting rates (γ) differ.*

Proof. The proof follows directly from the fact that the reward expectation depends on the forgetting rate (γ). ∎

Intuitively, an individual with a higher forgetting rate will have risk attitudes that are less affected by the rewards from the more distant past. Individuals with different forgetting rates will also respond differently to positive and negative shocks to the rewards. Imagine now that Oskar and John have the same predisposition and experienced the same stream of rewards until the previous period ($t - 1$) when they both received a reward much smaller than the previous rewards in the stream (a negative shock when benchmarked against their expectations). If we assume that Oskar has a low forgetting rate (a long memory), his expectation reflects many previous rewards that mitigate the influence of the smaller reward in $t - 1$ on his expectation. If John forgets very quickly, in contrast, his current reward expectation could be solely based on the small reward in the previous period. Under these

16

conditions, Oskar would be more risk-tolerant than John because his reward expectation is higher as it incorporates information about the rewards from periods preceding the negative shock. Going forward though, the relationship reverses. If in the current period the reward increases, in the following period, John will immediately "forget" about the negative shock and in the next period will be more risk-tolerant than Oskar who will "remember" about the negative shock for a longer period.

4.1.2 Predisposition and Risk Attitude

Proposition 1 implies that an individual with a very low-valued predisposition ($\alpha = 1$) is risk averse for all $x_{i,t}$, just as in traditional functional forms for utility. At the other extreme, when the individual's predisposition has high values ($\alpha \to \infty$), ESVT predicts risk aversion for gains (SV is concave, $RRA_t > 0$ iff $x_{i,t} > M_t$) and risk seeking for losses (SV is convex, $RRA_t < 0$ iff $x < M_t$) *relative to reward expectation*, a pattern generally in line with the value function of Prospect Theory. The model can thus accommodate the two most popular and competing models of the value function as its extreme cases without much complexity. One might suspect that most decision makers are somewhere in between these two extremes: their subjective value function starts off as convex and as x increases switches to concave. The inflection in S occurs for some $x_{i,t} < M_t$ because $\sqrt[\alpha]{\frac{\alpha-1}{\alpha+1}} \leq 1$.

4.2 Reward Expectation and Diminishing Marginal Value

Many reference-dependent models predict constant marginal utility with respect to the moving reference point; as the reference point shifts upwards, the marginal value of a \$1 gain relative to the reference point remains constant. This is of some concern because the notion of a stable diminishing marginal value with increasing wealth has been central to economic thought since Adam Smith's exposition on the Diamond-Water Paradox (Smith, 1776). To illustrate the point, assume a very simple Prospect Theoretic reference-dependent model where the reward expectation is given by the status quo and the utility is given by $U = u(x - r)$, where x is the actual reward and r is the reward expectation. In such a simple model, the marginal utility from an additional unit (of x) is constant with respect to how many units of the reward individual already owns (see Figure 3A). This, of course, makes the surprising prediction that a single consumption unit offered in excess of the status quo is of equal utility to a pauper as to a billionaire.

ESVT does not make this prediction but rather preserves a form of diminishing marginal value across wealth and expectation levels. In ESVT, the subjective value function *combines* the reward expectation and predisposition; their interaction yields a value function that preserves diminishing marginal value. In ESVT the subjective marginal value that an individual experiences with successive increments in reward always strictly declines as illus-

17

trated in Figure 3B. For tractability and comparison (as with Prospect Theory above), here we assume a situation where an individual's reward expectation is equal to their current status quo. It is clear, from Figure 3B, that the individual values an additional consumption unit less, the more consumption units she already possesses, and that this is true regardless of the expectation level. We establish this formally in the proposition below.
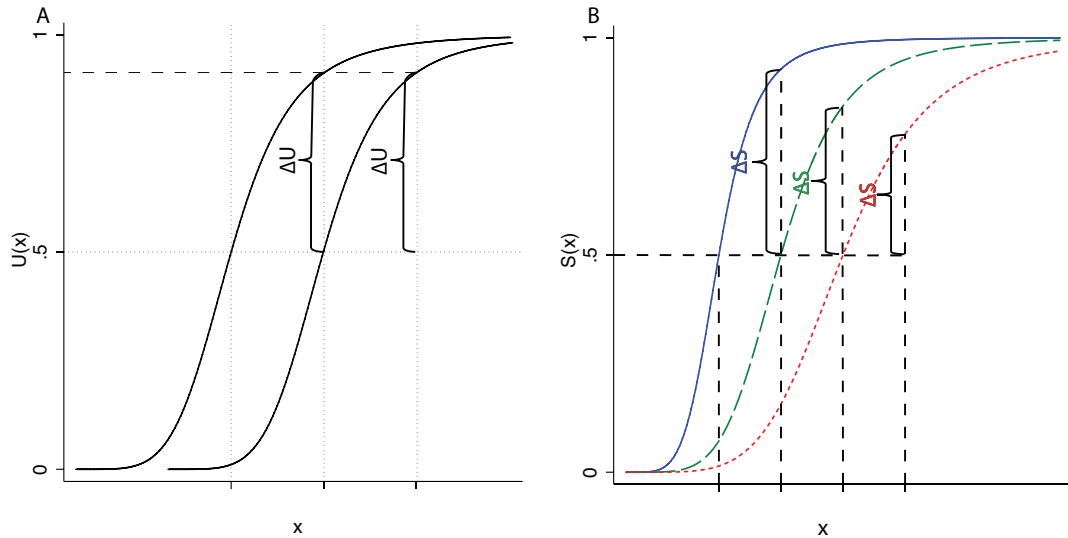

Figure 3   Marginal utility and marginal value in difference-based reward expectation model (A) and in divisive normalization model (B).

Proposition 3. *Assume two individuals with the same predisposition and expectations equal to their status quo. The marginal value of an additional unit of the consumable reward will be lower for the individual who has one more unit of this consumable reward.*

Proof. For an individual with reward expectation $M_{i,t} = x_{i,t}$, an increase in subjective value from another unit of consumable reward (an increase from $x_{i,t}$ to $x_{j,t}$) is given by

$$S_t(x_{j,t}, M_{i,t}) - S_t(x_{i,t}, M_{i,t}) = \frac{x_{j,t}^{\alpha}}{x_{j,t}^{\alpha} + M_{i,t}^{\alpha}} - 0.5$$

Let $x_{1,t} + 2 = x_{2,t} + 1 = x_{x,t}$, $M_{1,t} = x_{1,t}$, and $M_{2,t} = x_{2,t}$.

To show: $S_t(x_{2,t}, M_{1,t}) - S_t(x_{1,t}, M_{1,t}) > S_t(x_{3,t}, M_{2,t}) - S_t(x_{2,t}, M_{2,t})$ (diminishing marginal value condition).

Since $S_t(x_{1,t}, M_{1,t}) = S_t(x_{2,t}, M_{2,t}) = 0.5$, the diminishing marginal value condition boils down to:

$$S_t(x_{2,t}, M_{1,t}) = \frac{x_{2,t}^{\alpha}}{x_{2,t}^{\alpha} + M_{1,t}^{\alpha}} > \frac{x_{3,t}^{\alpha}}{x_{3,t}^{\alpha} + M_{2,t}^{\alpha}} = S_t(x_{3,t}, M_{2,t})$$

Using our assumption that $M_{i,t} = x_{i,t}$ and rearranging, we get that the diminishing marginal value condition holds if and only if

18

$$x_{2,t}^{2\alpha} > x_{1,t}^{\alpha} x_{3,t}^{\alpha} \Leftrightarrow$$
$$x_{2,t}^{2} > x_{1,t} x_{3,t} \Leftrightarrow$$
$$(\tfrac{x_{1,t}+x_{3,t}}{2})^2 > x_{1,t} x_{3,t} \Leftrightarrow$$
$$0.25(x_{1,t}^2 + 2x_{1,t}x_{3,t} + x_{3,t}^2) > x_{1,t} x_{3,t} \Leftrightarrow$$
$$0.25(x_{1,t}^2 - 2x_{1,t}x_{3,t} + x_{3,t}^2) > 0 \Leftrightarrow$$
$$(x_{1,t} - x_{3,t})^2 > 0$$

which always is true and completes the proof.

∎

4.3 Relation to Probability Weighting

One of the most discussed features of human choice behavior is the widely observed over-weighting of low probabilities and the underweighting of high probabilities. For this reason, it is notable that ESVT does not explicitly employ a probability weighting function of the kind envisioned by Savage (1952) and implemented by Prospect Theory. ESVT, however, does achieve a probability weighting-like effect, by embedding a mechanism that captures the overweighting of low probabilities and the underweighting of high probabilities via its 2-parameter value function.[5] When we observe in human choice behavior an overvaluation of low probability offers, we are observing that subjects place higher (certain) monetary values on low probability outcomes than would be predicted from their behavior observed at higher probabilities. Conversely, we see that subjects state lower values, in the form of certainty equivalents, on high probability outcomes than expected. Like prospect theory, ESVT also captures this feature of human choice behavior.

To illustrate this in Figure 4 we plot the certainty equivalent c, the certain monetary reward that the subject would view as subjectively equivalent to a reward x received with probability p. The certainty equivalent formula under ESVT shows curvature that is very similar to a popular probability weighting function, first introduced by Goldstein and Einhorn (1987):

$$c = \left(\frac{p x^{\alpha} M^{\alpha}}{(1-p)x^{\alpha} + M^{\alpha}} \right)^{\frac{1}{\alpha}} \tag{8}$$

Note that just as in Prospect Theory, in ESVT we predict behavior consistent with an "overweighting" of low probabilities and an "underweighting" of larger probabilities (shown here in certainty equivalent decisions). Said more precisely, ESVT individuals make risk-seeking valuations for a gamble at low probabilities but shift to risk averse behaviors as probabilities grow. Our model, even though it does not explicitly include a probability weighting

---

[5]A similar mechanism occurs in range-normalization models as shown previously by Kontek and Lewandowski (2018).
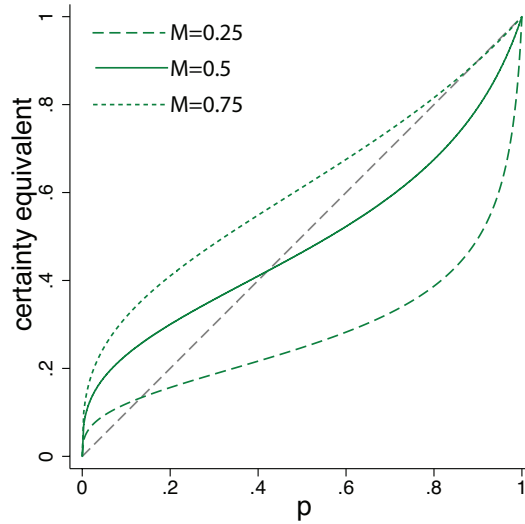
19

Figure 4 Certainty equivalent of a gamble that pays \$1 with probability p drawn as a function of p. The 45-degree line indicates the certainty equivalent for an expected value maximizer. $\alpha = 3$

function, embeds a mechanism that can explain why people are generally risk averse but at the same time are willing to purchase lottery tickets which offer a small probability of large gain and cost more than their expected value.

It is worth noting that empirically observed probability weighting functions estimated at the individual subject level differ between individuals and take shapes that are often far from the canonical inverse-S (Abdellaoui, 2000, Bruhin et al., 2010, Harbaugh et al., 2002, Harrison and Rutström, 2009, Wilcox, 2015, Hsu et al., 2009, Tobler et al., 2008, Fehr-Duda et al., 2011). Existing models do not capture this heterogeneity at the individual level. ESVT, in contrast, generates testable predictions about this heterogeneity in revealed probability distortions. To elaborate on the relationship between expectation, predisposition, and the probability weighting function, we undertook a simple numerical exercise. First, we used ESVT generatively to produce a series of choices in response to a set of lotteries. To do this, we simply implemented our value function, varying both the reward expectation and the predisposition. We then fit that choice data (generated by a synthetic ESVT chooser) with a probability weighting function embedded in traditional Prospect Theory.

We simulated nine individuals with different predispositions ($\alpha_{low} = 1$, $\alpha_{med} = 2$, and $\alpha_{high} = 5$) and different reward expectations ($M_{low} = 40$, $M_{med} = 60$, and $M_{high} = 80$). Given these predispositions and expectations, for each of the individuals we generated their choices in the well-known decision problems used by Gonzalez and Wu (1999). We then fit the ESVT-generated choices of our nine synthetic agents using the standard joint estimation approach with a classic parametrization of Prospect Theory, assuming CRRA utility function and a two-parameter probability weighting function ($w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$). Figure 5 presents
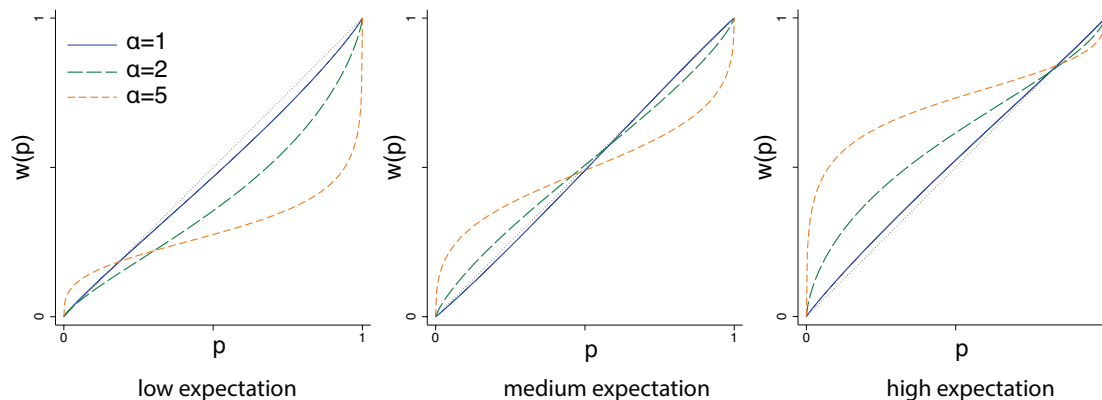
20

the results of this analysis.

Figure 5 Probability weighting functions of ESVT choosers estimated with $w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$ and CRRA utility functions.

What is striking is that our fit of Prospect Theory to the ESVT-generated synthetic choosers gave rise to probability weighting-like behavior when analyzed with prospect theory. If it did not, our estimated probability weighting functions in Figure 5 would all be straight, 45-degree lines. Another interesting feature of our model, illustrated in Figure 5, is that the estimated probability weighting functions systematically differ based on the expectation and predisposition. This raises the possibility that our model can shed some light on the to-date unknown sources of heterogeneity in the estimated probability weighting functions. We see that individuals with the highest-valued predisposition show the most noticeable probability distortions. Individuals with medium expectations show the curvature of the probability weighting function most commonly encountered in the literature. Individuals with low (high) expectations underweight (overweight) probability at almost all levels. Intuitively, the reason for such distortions is that in ESVT, individuals with high reward expectation take more risks. The smallest probability distortions occur for individuals with low-valued predisposition parameters. This is expected because their value function closely resembles CRRA (with which it is fit) thus no significant distortions in probability are inferred.

Probability weighting has been conceived as an explanation for famous failings of Expected Utility such as the Allais Paradox (Allais, 1953). Choosers with ESVT preferences are of interest in this regard because while they have no generative distortions in probability perception, they do still show an overweighting of rewards received with low probabilities and an underweighting of rewards received with high probabilities; an inverse S-shaped probability weighting functions when estimated using Prospect Theory. ESVT does this by tracing differences in the estimated probability weighting functions to the individual's expectations and predispositions. However, a critical shortcoming of the dynamic model that we focus

on here is that it cannot capture the Allais Paradox itself. A similar shortcoming in the range-normalization models has been resolved by assuming loss aversion and setting the reference point to the maxmin (Baucells et al., 2018, 2019). In a similar spirit, an extension of the simplified dynamic normalization model presented here to the more general form which includes choice-set (spatial) normalization (Webb et al., 2020) (see Section 2) would enrich the "reference point" by allowing it to account for other payoffs in the current choice set. This together with below expectation aversion or loss aversion provides an avenue for extensions of ESVT that can capture both the Allais Paradox and the basic probability distortion feature we describe here.

4.4   Below-Expectation Aversion

Loss aversion is a key component of Prospect Theory and many other contemporary models of choice. This concept derives from the observation that in many situations losses loom larger for choosers than equally sized gains. Although our model is limited in scope to positive rewards, it captures a related phenomenon, aversion to outcomes that are below reward expectation, which we call "below-expectation aversion".

Inspired by the traditional definition of loss aversion, we define the related phenomena of below-expectation aversion (λ^e) as the ratio of the subjective value of loss of size k and gain of size k, each relative to the subjective value at the reward expectation M_t (Abdellaoui et al., 2007):

$$\lambda_t^e = \frac{S(M_t) - S(M_t - k)}{S(M_t + k) - S(M_t)} \tag{9}$$

Plugging in the formula for the subjective value function (Equation 1), and simplifying, we get the expression for below-expectation aversion in our model:

$$\lambda_t^e = \frac{(M_t^\alpha - (M_t - k)^\alpha)(M_t^\alpha + (M_t + k)^\alpha)}{((M_t - k)^\alpha + M_t^\alpha)((M_t + k)^\alpha - M_t^\alpha)} \tag{10}$$

Whenever $\lambda^e > 1$, an individual exhibits below-expectation aversion by definition. If $\lambda^e = 1$, an individual treats losses and gains *relative to the expectation* equally.

Theorem 3. *An individual with ESVT preferences always displays below-expectation aversion. The strength of below-expectation aversion depends on stake size and predisposition.*

Proof. To show: For all k, for all M_t such that $0 < k < M_t$ and for all $\alpha > 0$, individual is below-expectation averse ($\lambda^e > 1$).

$\lambda^e > 1 \Leftrightarrow \frac{(M_t^\alpha - (M_t-k)^\alpha)(M_t^\alpha + (M_t+k)^\alpha)}{((M_t-k)^\alpha + M_t^\alpha)((M_t+k)^\alpha - M_t^\alpha)} > 1$

22

Since the denominator is always positive $(((M_t - k)^\alpha + M_t^\alpha)((M_t + k)^\alpha - M_t^\alpha) > 0)$, we can multiply both sides of the inequality by it without changing the sign and rewrite the condition for below-expectation aversion, $\lambda^e > 1$, as:

$(M_t^\alpha - (M_t - k)^\alpha)(M_t^\alpha + (M_t + k)^\alpha) > (M_t^\alpha + (M_t - k)^\alpha)(-M_t^\alpha + (M_t + k)^\alpha)$

Performing the algebra, we get $\lambda^e > 1 \Leftrightarrow 2k^2 > 0$, which holds for all k and completes the proof. ∎

One key feature that distinguishes below-expectation aversion from Prospect Theory's loss aversion is that below-expectation aversion is not a fixed parameter added to the model. Instead, below-expectation aversion is a function of both predisposition and how far the offer under consideration would place you above or below the reward expectation M_t (captured by the stake size k). It is thus clear that both the stake size k and predisposition play important roles in determining the strength of below-expectation aversion. In general, the model predicts that higher predispositions lead to less below-expectation aversion and that below-expectation aversion first increases and then decreases in stake size. These effects are illustrated in Figure 6. While there has been some research suggesting that the stake size affects loss aversion (Pope and Schweitzer, 2011, Blavatskyy and Pogrebna, 2007), more empirical research is needed to understand whether the pattern of behavior is fully consistent with the non-monotonic predictions of our model.
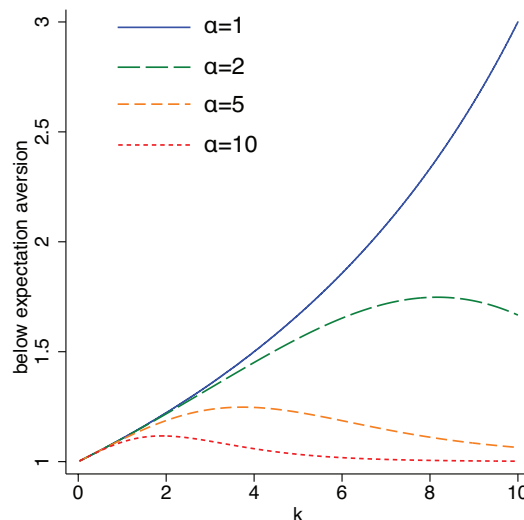

Figure 6   Below-expectation aversion (λ^e) as a function of the gain (loss) size (k) at different values of predisposition (α).

While ESVT in the form presented here does capture something closely akin to loss aversion in the domain of gains, it is important to note that ESVT, as presented here, is a theory of choice over gains and does not include the representation of "true" losses as in

23

Prospect Theory. This raises a number of critical uncertainties that arise when we consider extending ESVT to the domain of "true" losses.

Unfortunately, the neurobiological evidence which might provide some guidance on how to incorporate true losses into the theory is complex and not entirely unambiguous. As noted in section 2 single neurons encode real numbers in their firing rates (in units of action potentials per second), which are bounded between zero and a fixed upper value, typically around 100. These same neurons also have a so-called baseline rate, the rate of action potential generation produced by the neuron in the absence of an input, often observed to be in the range of 10 action potentials per second. This set of observations raises two clear possibilities for how these physical systems encode losses.

One possibility is that firing rates below baseline are effectively employed to encode negative numbers and hence losses. The classic paper of Tom et al. (2007) explicitly sought to test this hypothesis in a loss aversion experiment and concluded that losses were indeed coded as below-baseline firing rates, as observed indirectly with a brain scanner. If, however, the unique zero of neurons is fixed at (for example) 10, then this would imply a lower precision representation for losses than for gains. This would have to be true given our understanding of neuronal variance (Churchland et al., 2011), unless the unique zero point, the baseline firing rate, was dynamically adjusted by the subject's expectation — a possibility that has not yet been tested. If this were the case, then the mechanism we describe for below-expectation aversion could apply to the case of true loss aversion, with only moderate changes.

A second possibility is that separate and parallel neural systems encode $\mathbb{R}_+$, and $\mathbb{R}_-$, with the $\mathbb{R}_-$ encoding system consisting of output neurons that show a negative correlation between firing rate and value. Some evidence exists which supports this interpretation (Knutson et al., 2008, Samanez-Larkin and Knutson, 2015). If this were the case, then true loss aversion might well rest on a unique mechanism that we have not yet modeled. Ongoing neurobiological research should continue to shed light on this issue and we have every expectation that the mechanism by which losses are incorporated into human choice behavior will soon become clearer. It is partly in response to this uncertainty that the current form of the theory remains restricted to the domain of gains.

The second set of issues that serve to constrain the theory to the domain of gains revolves around the nature of the ESVT value function. It is important to note that the ESVT value function does not include a discontinuity or sudden change in slope at the expectation point. This strongly differentiates it from Prospect Theory in which the value function shows a discrete "kink" at an x-value of zero. While many behavioral models and studies tend to rely on such a value function, neurobiological data do not appear to support such representation. There is to our knowledge, no evidence to suggest that neural firing rate functions show a discontinuity at the baseline firing rate, although the curvature of the firing rate function may well change continuously at that point. Still, the kink has served a significant role in

24

broadening our understanding (e.g. Köszegi and Rabin (2006), Baucells and Hwang (2017)). Today, ESVT walks a middle ground in addressing these issues by restricting itself to a more limited domain (gains only) within which it can reproduce a below-expectation aversion without invoking a sharp change in the slope of the value function, a restriction that keeps ESVT in line with neurobiological evidence in Tom et al. (2007).

4.5 The Endowment Effect

The traditional explanation for the endowment effect is *loss aversion*: People endowed with a mug demand more money to sell it than people who do not have the mug are willing to pay for it because losses loom larger than gains. This explanation has been challenged in the literature and revisions have been suggested which center on an interaction between changes in expectations and loss aversion (Köszegi and Rabin, 2006). This newer model of the gap between willingness to pay (WTP) and willingness to accept (WTA) predicts that the more unexpected a trade is, the larger should be the endowment effect (the WTP-WTA gap) for a given level of loss aversion. Empirical tests of these newer models have, however, produced mixed results (Heffetz and List, 2014, Ericson and Fuster, 2011, Khalil and Wu, 2017).

ESVT captures the endowment effect in a somewhat novel way that we believe remains consistent with empirical findings and which may merit examination:

Proposition 4. *Let $x_{i,1}$ be the individual's initial ownership level of the good under consideration. In period 2, the individual is endowed with an additional amount of the good such that, in $t \geq 2$, $x_{j,t} > x_{i,1}$ is the ownership level after endowment. Let M_1 be the reward expectation before the endowment and M_t be the reward expectation in periods $t \geq 2$ after the endowment. The endowment effect occurs if $M_t < \frac{x_{i,1}x_{j,t}}{M_1}$. The endowment effect does not occur if $(M_1^\alpha - M_t^\alpha)(M_1^\alpha M_t^\alpha - x_{i,1}^\alpha x_{j,t}^\alpha) = 0$.*

Proof. The endowment effect occurs whenever:

$WTA = \frac{x_{j,t}^\alpha}{(M_t^\alpha + x_{j,t}^\alpha)} - \frac{x_{i,1}^\alpha}{(M_t^\alpha + x_{i,1}^\alpha)} > \frac{x_{j,t}^\alpha}{(M_1^\alpha + x_{j,t}^\alpha)} - \frac{x_{i,1}^\alpha}{(M_1^\alpha + x_{i,1}^\alpha)} = WTP$

Which is equivalent to:

$WTA = \frac{(x_{j,t}^\alpha - x_{i,1}^\alpha)M_t^\alpha}{(M_t^\alpha + x_{j,t}^\alpha)(M_t^\alpha + x_{i,1}^\alpha)} > \frac{(x_{j,t}^\alpha - x_{i,1}^\alpha)M_1^\alpha}{(M_1^\alpha + x_{j,t}^\alpha)(M_1^\alpha + x_{i,1}^\alpha)} = WTP$

After simplifying, we get that $WTA > WTP$ if and only if

$(M_1^\alpha - M_t^\alpha)(M_1^\alpha M_t^\alpha - x_{i,1}^\alpha x_{j,t}^\alpha) > 0$

Since by assumption, $M_1^\alpha - M_t^\alpha < 0$, $WTA > WTP$ if and only if

$M_1^\alpha M_t^\alpha - x_{i,1}^\alpha x_{j,t}^\alpha < 0 \Leftrightarrow M_t M_1 < x_{i,1} x_{j,t}$.

The endowment effect does not occur whenever

$WTA = WTP \Leftrightarrow (M_1^\alpha - M_t^\alpha)(M_1^\alpha M_t^\alpha - x_{i,1}^\alpha x_{j,t}^\alpha) = 0$

∎

This proposition implies that there are two particular situations when the endowment effect does not occur which we can understand with the help of Figure 7A. In Figure 7A,

in blue we plot a subjective value function for an ESVT chooser who has three mugs, and is expecting to have three mugs, and indicates the corresponding willingness to pay (WTP) for a fourth mug. The first situation under which the endowment effect does not occur is when there is no change in the reward expectation after an endowment $M_1 = M_t$ and therefore post-endowment and pre-endowment subjective value functions are the same blue curve and trivially the valuation of the fourth mug stays the same. Empirically such cases have been demonstrated in List (2004), Heffetz and List (2014), Khalil and Wu (2017). This implies that an endowment that actively changes the reward expectation is necessary for the endowment effect to occur. The second case when the endowment effect does not occur is when the reward expectation changes after the endowment and is equal to the new level of ownership ($M_1 = x_{i,1}$ and $M_t = x_{j,t}$). Such a post-endowment subjective value function is given by the red curve in Figure 7A. As proven in Proposition 4, in such a case there is no endowment effect. This is likely to occur when a sufficiently long time (or perhaps a sufficient number of offers) has passed between the endowment and the elicitation of willingness to accept; a time sufficient for the individual to fully adjust her reward expectation to the new ownership level relative to her forgetting rate. ESVT's model of the endowment effect thus rests on the movement of the overall expectation and curvature of the value function, in contrast to Köszegi and Rabin (2006) in which loss aversion and the extent to which a future trading event is expected controls the level of observed endowment effect.
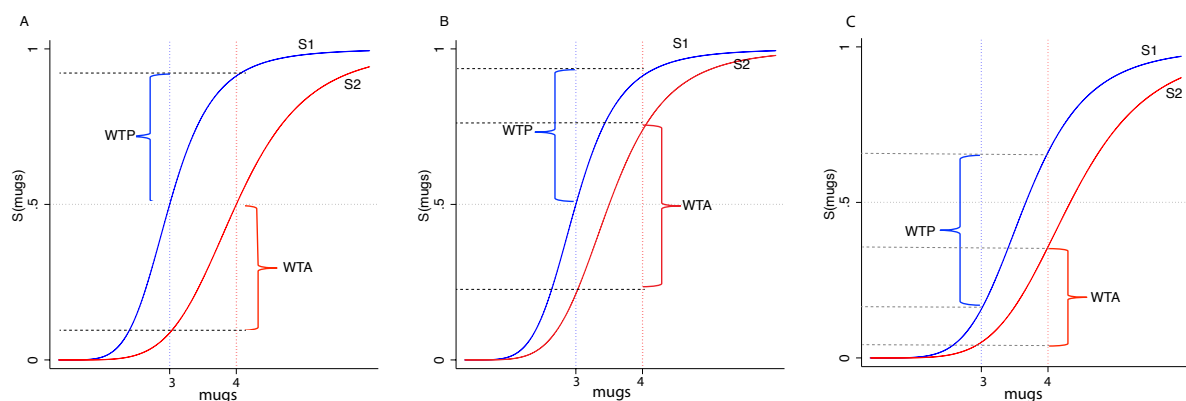

Figure 7 Endowment effect and reward expectation. S_1 in blue (S_2 in red) represent subjective value function before (after) endowment. A: No endowment effect with no or full adaptation of the reward expectation to current ownership. B: Endowment effect occurs because of incomplete adjustment of the reward expectation to current ownership. C: Reverse endowment effect occurs because initial reward expectation is inflated above the ownership level.

To better understand how the endowment effect occurs in ESVT, let us concentrate on one simple case: let us assume that before the endowment occurs, the individual's reward expectation coincides with the status quo ownership level $M_1 = x_{i,1} = 3$ (as illustrated

26

by the blue curve in Figure 7B). The endowment effect will be observed as long as the reward expectation after the endowment has not fully updated to the new ownership level ($M_t < x_{j,t} = 4$), for example, because the forgetting rate is sufficiently low or an insufficient amount of time has passed since the endowment. This is illustrated by an example red curve in Figure 7B. In ESVT, after the endowment, the individual's reward expectation increases, and hence the subjective value function shifts to the right, but due to the slow forgetting rate does not change all the way to the new ownership level immediately. Instead, the new reward expectation is a time-weighted average somewhere between the new and old ownership levels. In ESVT, as long as the reward expectation remains between the new and old ownership levels, the endowment effect occurs. As established in the next proposition, the endowment effect first gets stronger as time since endowment passes (consistent with the evidence in Strahilevitz and Loewenstein (1998)), and then it begins to weaken.

Proposition 5. *Assume that the individual's reward expectation is at the status quo ownership level. After an endowment of an additional unit of the good, as time goes by the endowment effect first increases and then decreases.*

Proof. Let $x_{i,1}$ be the initial level of ownership. In $t = 2$ the individual receives an additional amount of the good such that in $t \geq 2$ $x_{j,t} - x_{i,1} > 0$ is the additional endowment received in period 2. Assume that after period 2, there are no further endowments and that the ownership remains at $x_{j,2}$.

To show: For $t > 2$, $\frac{\partial WTA_t}{\partial M_t} > 0$ for $M_t \in (x_{i,1}, z)$ and $\frac{\partial WTA_t}{\partial M_t} < 0$ for $M_t \in (z, x_{j,2})$ where $z \in (x_{i,1}, x_{j,2})$.

$WTA_t = \frac{x_{j,2}^\alpha}{(M_t^\alpha + x_{j,2}^\alpha)} - \frac{x_{i,1}^\alpha}{(M_t^\alpha + x_{i,1}^\alpha)}$

$\frac{\partial WTA_t}{\partial M_t} = \frac{\alpha M_t^{\alpha-1}}{(M_t^\alpha + x_{j,2}^\alpha)^2 (M_t^\alpha + x_{i,1}^\alpha)^2} [x_{i,1}^\alpha (M_t^\alpha + x_{j,2}^\alpha)^2 - x_{j,2}^\alpha (M_t^\alpha + x_{i,1}^\alpha)^2]$

Since $\frac{\alpha M_t^{\alpha-1}}{(M_t^\alpha + x_{j,2}^\alpha)^2 (M_t^\alpha + x_{i,1}^\alpha)^2} > 0$ always holds, the sign of $\frac{\partial WTA_t}{\partial M_t}$ is determined by $x_{i,1}^\alpha (M_t^\alpha + x_{j,2}^\alpha)^2 - x_{j,2}^\alpha (M_t^\alpha + x_{i,1}^\alpha)^2$.

$\frac{\partial WTA_t}{\partial M_t} > 0 \Leftrightarrow$

$x_{i,1}^\alpha (M_t^\alpha + x_{j,2}^\alpha)^2 - x_{j,2}^\alpha (M_t^\alpha + x_{i,1}^\alpha)^2 > 0 \Leftrightarrow$

$(x_{j,2}^\alpha - x_{i,1}^\alpha)(x_{i,1}^\alpha x_{j,2}^\alpha - M_t^{2\alpha}) > 0 \Leftrightarrow$

Since $(x_{j,2}^\alpha - x_{i,1}^\alpha) > 0$ by assumption, the above inequality holds when $x_{i,1}^\alpha x_{j,2}^\alpha - M_t^{2\alpha} > 0$ $\Leftrightarrow M_t < \sqrt{x_{i,1} x_{j,2}}$.

We can now conclude that $\frac{\partial WTA_t}{\partial M_t} > 0$ for $M_t \in (x_{i,1}, \sqrt{x_{i,1} x_{j,2}})$ and $\frac{\partial WTA_t}{\partial M_t} < 0$ for $M_t \in (\sqrt{x_{i,1} x_{j,2}}, x_{j,2})$ which completes the proof.

∎

Our framework can be used to derive more cases when the endowment effect will or will not occur or even when it, paradoxically might be observed to reverse. Figure 7C presents a graphical example of a reverse endowment effect predicted by our model. The intuition

for this case is that the reverse endowment effect can occur if initially, the participant has reward expectations that are above their ownership level. These are illustrated in Figure 7 via the blue curve that has a reward expectation above 3 (even though the actual ownership is 3). In our framework, such a situation can occur if an individual has a low forgetting rate and has experienced a loss shortly before the endowment. Suppose that such an individual is endowed with an additional mug. For a range of initial inflated reward expectations and forgetting rates, the reward expectation after the endowment may increase, leading to a smaller willingness to accept than willingness to pay.

Here we have offered a description of a subset of ESVT's key testable predictions on when the endowment effect, reverse endowment effect, or no endowment effect will occur for relatively simple circumstances like those that have been examined experimentally. The key force driving the endowment effect under these conditions is whether and how fast the reward expectation updates after the endowment. In presenting our analysis of the endowment effect in ESVT we have, of course, examined only quite simple and generic situations. For example, we have assumed that the consumption/experienced value of the good does not change with time which is not true for all types of goods. Goods such as one's mother's jewelry might well grow in experienced value as years pass, a feature not captured by this model. To take another example, the market value of wine or property usually appreciates over time. All of these elements of consumption or experienced value vary from good to good exogenously to our model (and perhaps to the chooser). Despite these simplifications, our model does make a number of unexpected and testable predictions. We predict that under some conditions a reverse endowment effect can arise. We also predict that if sufficient time has passed after the endowment is made (and no changes in the remembered value of the good are imposed by emotional events), the endowment effect associated with the endowed good diminishes.

5   Discussion

In this paper we have presented an analysis of the preference structures produced by Expected Subjective Value Theory, attempting to assess its strengths and weaknesses as a descriptive model of human choice behavior. Expected Subjective Value Theory rests on the optimal coding normalization models from neuroscience (Carandini and Heeger, 2012) which have been gaining popularity in neuroeconomics (Louie et al., 2013) and which have recently been shown to be normatively efficient (Steverson et al., 2019). The analyses presented in this paper suggest that many of the key behavioral phenomena that Prospect Theory was designed to capture can also be explained by Expected Subjective Value Theory while preserving the notion of diminishing marginal value which was abandoned under Prospect Theory (Kahneman and Tversky, 1979). We also find that Expected Subjective Value Theory makes a series of novel predictions that distinguish it from Prospect Theory behaviorally, predictions

28

that are amenable to future testing. Specifically, ESVT provides a new explanation for the endowment effect that is unrelated to loss aversion (or its ESVT analogue below-expectation aversion) and offers a theoretical explanation for the heterogeneity in estimated probability weighting functions. It is, at least to us, surprising that this simple model, created to describe optimal representational frameworks in the brain, can capture such a broad variety of human choice behaviors — phenomena that it was not originally designed to capture.

We take it as a given that divisive normalization, the parent theory for ESVT, outperforms other models with regard to neurobiological observables in context-driven situations (Glimcher, 2015, Louie et al., 2014, LoFaro et al., 2014, Louie et al., 2013). Models like Prospect Theory cannot, by design, capture the effect of temporal context effects or choice-set context effects which have now been observed neurobiologically. Here we have focused on the temporal context effects induced by the expectation, neglecting choice-set context effects for tractability, but we acknowledge that this ability to capture neurobiological context effects will be only marginally interesting to pure economists. As a result, the focus of this paper has been to examine how and where one can compare the performance of Prospect Theory and ESVT in behavioral data under quite simple conditions. We believe our analysis suggests that ESVT compares well with Prospect Theory in many regards at the behavioral level.

5.1   Primitives of the Theory

It is important to note that ESVT differs fundamentally from nearly all other descriptive choice theories in the basic nature of its primitives, although in a way driven by recent insights into the neurobiology of decision-making. Expected Utility, at least when typically used as a predictive tool, relies on the specification of a functional form for utility to define a chooser's preferences. And of course, this means that when one fits Expected Utility to observed behavior to formulate a quantitative prediction about future choices, one parameterizes the utility function one has selected, typically with a curvature specification and a wealth level. To restate this in a more philosophical vein, Expected Utility represents choosers as being exogenously provided with a utility function, and as having an endogenous wealth level, nothing more.

Prospect Theory broadened the parameter set for the utility function by introducing the value function and the probability weighting function (Kahneman and Tversky, 1979). This was accomplished by introducing two curvature terms (one for gains and one for losses), a loss aversion term, a reference point, and subjective distortions in probability perception (via one or two additional parameters). Effectively, Kahneman and Tversky broadened the primitives of their theory (relative to Expected Utility) to include not just the notion of utility function curvature but also the notions of a reference point, loss aversion, and

29

subjective probabilities. These four ideas (utility curvature, a reference point, loss aversion, and probability weighting) form the core primitives of value and choice in their theory and are the four (or six) exogenous elements with which choosers are provisioned by the theory, although the reference point is sometimes viewed as endogenous.

In contrast, ESVT's value function *emerges* from an interaction between the two core primitives of its theory: *predisposition* and *expectation*. When one says that a chooser employing Prospect Theory shows overweighting of low probabilities, we mean that overweighting of low probabilities is both a logical primitive and an endowed property of the chooser. In some sense, we mean that choosers behave *as if* they had a specific mechanism inside their brains which differentially weighted small and large probabilities. In contrast, when we say that a chooser employing ESVT overweights small probabilities, we mean that were one to use standard econometric tools to measure a probability weighting function, one would observe values consistent with overweighting of small probabilities.

5.2 Related Models of Choice and Future Directions

The core idea underlying divisive normalization models like ESVT is that precision in choice is costly. These costs have been absent from much but not all (for example Simon (1976), Robson (2001), Rayo and Becker (2007), Netzer (2009), Glimcher (2010), Woodford (2012), Bordalo et al. (2012), Robson and Whitehead (2016), Steverson et al. (2019)) decision-scientific thinking. Neurobiological evidence available today clearly indicates that increasing the precision of any neural representation is without a doubt costly. The model we analyze here is not the first one in economics or decision science to study how limited neural resources affect an individual's attitude towards risk. Woodford (2012), Robson and Whitehead (2016), Rayo and Becker (2007) all come to the conclusion that the efficient use of finite processing resources would result in an S-shaped value function. In these models, known statistics about the reward structure in the environment determine the exact shape of the efficient value function. Such adaptive coding is also present in the neurobiological model of Padoa-Schioppa and Rustichini (2014). The divisive normalization models of which ESVT is an exemplar can also be thought of as an implementation of the ideas conveyed in these earlier theoretical models, related to the idea of limited neural resources that drive modern economic models of attention.[6]

Finally, we note that it is quite unsatisfying to us that the central free parameter of the model, predisposition, is not yet specified in some deep philosophical form ultimately amenable to endogenization within the model (or at least within the chooser). We acknowledge that neurobiologists have only made cursory attempts to explain what predisposition means. Heeger (1992) argues that it reflects the nonlinear response to inputs. Webb et al.

---

[6]For a review of models of attention see Caplin (2016).

(2020) relates it to the capacity constraints of the nervous system. Chung et al. (2017) hints that it may be related to age and the size of certain brain areas. Given that increases in the predisposition steepen the subjective value function, it is also possible that the parameter tracks historical (or chosen item) reward magnitude statistics such as variance. We find these suggestions deeply interesting. More theoretical and experimental research is needed, however, to define what predisposition is before we render it a directly observable property of choosers.

References

Abdellaoui, M. (2000). Parameter-Free Elicitation of Utility and Probability Weighting Functions. *Management Science 46*(11), 1497–1512.

Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007, oct). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science 53*(10), 1659–1674.

Albrecht, D. G. and D. B. Hamilton (1982, jul). Striate cortex of monkey and cat: contrast response function. *Journal of neurophysiology 48*(1), 217–37.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica 21*(4), 503–546.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review 61*, 183–193.

Barlow, H. A. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication*. Cambridge, MA: MIT Press.

Baucells, M. and W. Hwang (2017). A model of mental accounting and reference price adaptation. *Management Science 63*(12).

Baucells, M., K. Kontek, and M. Lewandowski (2018). Range and Sign Dependent Utility for Risk and Time.

Baucells, M., M. Lewandowski, and K. Kontek (2019). Range Utility Theory for Uncertain Cash-flows.

Bayer, H. M. and P. W. Glimcher (2005, jul). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron 47*(1), 129–41.

Becker, G. M., M. H. DeGroot, and J. Marschak (1963). Stochastic models of choice behavior. *Behavioral Science 8*(1), 41–55.

Blavatskyy, P. and G. Pogrebna (2007, jun). Loss aversion? Not with half-a-million on the table! *SSRN Electronic Journal*.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012, aug). Salience Theory of Choice Under Risk. *The Quarterly Journal of Economics 127*(3), 1243–1285.

Bruhin, A., H. Fehr-Duda, and T. Epper (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica 78*(4), 1375–1412.

Busse, L., A. R. Wade, and M. Carandini (2009, dec). Representation of concurrent stimuli by population activity in visual cortex. *Neuron 64*(6), 931–42.

Caplin, A. (2016, oct). Measuring and Modeling Attention. *Annual Review of Economics 8*(1), 379–403.

Caplin, A. and M. Dean (2008, may). Dopamine, Reward Prediction Error, and Economics. *Quarterly Journal of Economics 123*(2), 663–701.

Carandini, M. and D. J. Heeger (2012, jan). Normalization as a canonical neural computation. *Nature reviews. Neuroscience 13*(1), 51–62.

Chung, H.-K., A. Tymula, and P. Glimcher (2017, dec). The Reduction of Ventrolateral Prefrontal Cortex Gray Matter Volume Correlates with Loss of Economic Rationality in Aging. *The Journal of neuroscience : the official journal of the Society for Neuroscience 37*(49), 12068–12077.

Churchland, A. K., R. Kiani, R. Chaudhuri, X. J. Wang, A. Pouget, and M. N. Shadlen (2011). Variance as a Signature of Neural Computations during Decision Making. *Neuron 69*(4), 818–831.

Ericson, M. and A. Fuster (2011, oct). Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments. *The Quarterly Journal of Economics 126*(4), 1879–1907.

Fehr-Duda, H., T. Epper, A. Bruhin, and R. Schubert (2011). Risk and rationality: The effects of mood and decision rules on probability weighting. *Journal of Economic Behavior and Organization 78*(1-2), 14–24.

Frydman, C. and L. J. Jin (2019). Efficient coding and risky choice.

Glimcher, P. (2010). *Foundations of Neuroeconomic Analysis*. Oxford University Press.

Glimcher, P. (2015, jan). Understanding the Hows and Whys of Decision-Making: From Expected Utility to Divisive Normalization. *Cold Spring Harbor symposia on quantitative biology 79*, 169–76.

Glimcher, P. and A. Tymula (2017). Let the sunshine in? The effects of luminance on economic preferences, choice consistency and dominance violations. *PLoS ONE 12*(8).

Glimcher, P. W. (2011, sep). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America 108 Suppl*(Supplement_3), 15647–54.

Gluth, S., N. Kern, M. Kortmann, and C. L. Vitali (2020). Value-based attention but not divisive normalization influences decisions with multiple alternatives. *Nature Human Behaviour 4*(6), 634–645.

Goldstein, W. M. and H. J. Einhorn (1987). Expression Theory and the Preference Reversal Phenomena. *Psychological Review*.

Gonzalez, R. and G. Wu (1999, feb). On the shape of the probability weighting function. *Cognitive psychology 38*(1), 129–66.

Guo, J. and A. Tymula (2020). Waterfall illusion in risky choice.

Harbaugh, W., K. Krause, and L. Vesterlund (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics 5*, 53–84.

Hare, T. A., J. O'Doherty, C. F. Camerer, W. Schultz, and A. Rangel (2008, may). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of neuroscience : the official journal of the Society for Neuroscience 28*(22), 5623–30.

Harrison, G. W. (2008). Maximum likelihood estimation of utility functions using Stata.

Harrison, G. W. and E. E. Rutström (2009, jun). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics 12*(2), 133–158.

Heeger, D. J. (1992, aug). Normalization of cell responses in cat striate cortex. *Visual neuroscience 9*(2), 181–197.

Heeger, D. J., E. P. Simoncelli, and J. A. Movshon (1996, jan). Computational models of cortical visual processing. *Proceedings of the National Academy of Sciences of the United States of America 93*(2), 623–7.

Heffetz, O. and J. A. List (2014, oct). Is the endowment effect an expectations effect? *Journal of the European Economic Association 12*(5), 1396–1422.

Hsu, M., I. Krajbich, C. Zhao, and C. F. Camerer (2009, feb). Neural Response to Reward Anticipation under Risk Is Nonlinear in Probabilities. *Journal of Neuroscience 29*(7), 2231–2237.

Hunt, L. T., R. J. Dolan, and T. E. J. Behrens (2014, nov). Hierarchical competitions subserving multi-attribute choice. *Nature neuroscience 17*(11), 1613–22.

Kable, J. W. and P. W. Glimcher (2007, nov). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience 10*(12), 1625–1633.

Kahneman, D. and A. Tversky (1979). Prospect Theory. *Econometrica 47*(2), 311.

Kahneman, D. and A. Tversky (1984). Choices, values, and frames. *American Psychologist 39*(4), 341–350.

Khalil, E. and K. Wu (2017, may). Explicit vs implicit proprietorship: Can endowment effect theory explain exchange asymmetry? *Economics Letters 154*, 117–119.

Khaw, M. W., P. W. Glimcher, and K. Louie (2017, nov). Normalized value coding explains dynamic adaptation in the human valuation process. *Proceedings of the National Academy of Sciences 114*(48), 12696–12701.

Knutson, B., G. E. Wimmer, S. Rick, N. G. Hollon, D. Prelec, and G. Loewenstein (2008, jun). Neural Antecedents of the Endowment Effect. *Neuron 58*(5), 814–822.

Kontek, K. and M. Lewandowski (2018). Range-Dependent Utility. *Management Science 64*(6), 2473–2972.

Köszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics 121*(4), 1133–1165.

Landry, P. and R. Webb (2018). Pairwise Normalization: A Theory of Multi-Attribute Choice.

Lattimore, P. K., J. R. Baker, and A. D. Witte (1992). The influence of probability on risky choice: A parametric examination. *Journal of Economic Behavior & Organization 17*(3), 377–400.

Levy, I., J. Snell, A. J. Nelson, A. Rustichini, and P. W. Glimcher (2010, feb). Neural representation of subjective value under risk and ambiguity. *Journal of Neurophysiology 103*(2), 1036–47.

List, J. A. (2004). Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace.

LoFaro, T., K. Louie, R. Webb, and P. W. Glimcher (2014, mar). The Temporal Dynamics of Cortical Normalization Models of Decision-making. *Letters in Biomathematics 1*(2), 209–220.

Louie, K. and P. W. Glimcher (2010, apr). Separating value from choice: delay discounting activity in the lateral intraparietal area. *The Journal of neuroscience 30*(16), 5498–507.

Louie, K. and P. W. Glimcher (2012). Set-size effects and the neural representation of value. In R. J. Dolan and T. Sharot (Eds.), *Neuroscience of Preference and choice: Cognitive and neural mechanisms*, pp. 143–169. London: Academic Press.

Louie, K., P. W. Glimcher, and R. Webb (2015, oct). Adaptive neural coding: From biological to behavioral decision-making. *Current opinion in behavioral sciences 5*, 91–99.

Louie, K., L. E. Grattan, and P. W. Glimcher (2011, jul). Reward value-based gain control: Divisive normalization in parietal cortex. *The Journal of Neuroscience 31*(29), 10627–10639.

Louie, K., M. W. Khaw, and P. W. Glimcher (2013, apr). Normalization is a general neural mechanism for context-dependent decision making. *Proceedings of the National Academy of Sciences of the United States of America 110*(15), 6139–6144.

Louie, K., T. LoFaro, R. Webb, and P. W. Glimcher (2014, nov). Dynamic divisive normalization predicts time-varying value coding in decision-related circuits. *The Journal of Neuroscience 34*(48), 16046–57.

Netzer, N. (2009, may). Evolution of Time Preferences and Attitudes toward Risk. *American Economic Review 99*(3), 937–955.

Ohzawa, I., G. Sclar, and R. D. Freeman (1985). Contrast gain control in the cat's visual system. *Journal of neurophysiology 54*(3), 651–667.

Padoa-Schioppa, C. (2009, nov). Range-adapting representation of economic value in the orbitofrontal cortex. *The Journal of Neuroscience 29*(44), 14004–14.

Padoa-Schioppa, C. and A. Rustichini (2014, may). Rational Attention and Adaptive Coding: A Puzzle and a Solution. *American Economic Review 104*(5), 507–513.

Payzan-LeNestour, E., L. Pradier, and T. Putnins (2020). The 'Waterfall Illusion' in Financial Markets: How Risk Perception Is Distorted After Exposure to Extreme Risk.

Plassmann, H., J. O'Doherty, and A. Rangel (2007, sep). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience 27*(37), 9984–9988.

Platt, M. L. and P. W. Glimcher (1999, jul). Neural correlates of decision variables in parietal cortex. *Nature 400*(6741), 233–8.

Pope, D. G. and M. E. Schweitzer (2011, feb). Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes. *American Economic Review 101*(1), 129–157.

Ratliff, F. and H. K. Hartline (1959, jul). The responses of Limulus optic nerve fibers to patterns of illumination on the receptor mosaic. *The Journal of General Physiology 42*(6), 1241–55.

Rayo, L. and G. S. Becker (2007). Evolutionary Efficiency and Happiness. *Journal of Political Economy 115*(2), 302–337.

Reynolds, J. H. and D. J. Heeger (2009, jan). The normalization model of attention. *Neuron 61*(2), 168–85.

Robinson, T. and A. Tymula (2019). Divisive normalisation of value explains choice-reversals in decision-making under risk.

Robson, A. and L. A. Whitehead (2016). Rapidly adaptive hedonic utility.

Robson, A. J. (2001). The Biological Basis of Economic Behavior. *Journal of Economic Literature 39*(1), 11–33.

Rutledge, R. B., M. Dean, A. Caplin, and P. W. Glimcher (2010). Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *Journal of Neuroscience 30*(40).

Samanez-Larkin, G. R. and B. Knutson (2015). Decision making in the ageing brain: Changes in affective and motivational circuits.

Schwartz, O. and E. P. Simoncelli (2001, aug). Natural signal statistics and sensory gain control. *Nature neuroscience 4*(8), 819–25.

Sclar, G., J. H. Maunsell, and P. Lennie (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision research 30*(1), 1–10.

Shannon, C. E. and W. Weaver (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Simon, H. (1976). *Administrative Behavior*. New York: The Free Press.

Simoncelli, E. P. (2003, apr). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology 13*(2), 144–149.

Simoncelli, E. P. and B. A. Olshausen (2001, jan). Natural image statistics and neural representation. *Annual review of neuroscience 24*, 1193–1216.

Slovic, P. (1995). The construction of preference. *American Psychologist 50*(5), 346–371.

Smith, A. (1776). *The Wealth of Nations*. London: W. Strahan and T. Cadell.

Smith, A., B. D. Bernheim, C. Camerer, and A. Rangel (2014, may). Neural Activity Reveals Preferences Without Choices. *American economic journal. Microeconomics 6*(2), 1–36.

Steverson, K., A. Brandenburger, and P. Glimcher (2019). Rational imprecision: information processing, neural, and choice-rule perspectives. *Journal of Economic Behavior & Organization*.

Stigler, G. J. and G. S. Becker (1977). De Gustibus Non Est Disputandum. *The American Economic Review 67*(2), 76–90.

Strahilevitz, M. A. and G. Loewenstein (1998). The Effect of Ownership History on the Valuation of Objects. *Journal of Consumer Research 25*(3), 276–289.

Sutton, R. S. and A. G. Barto (1998). *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.

Tobler, P. N., G. I. Christopoulos, J. P. O'Doherty, R. J. Dolan, and W. Schultz (2008). Neuronal distortions of reward probability without choice. *The Journal of Neuroscience 28*(45), 11703–11.

Tom, S. M., C. R. Fox, C. Trepel, and R. A. Poldrack (2007, jan). The neural basis of loss aversion in decision-making under risk. *Science 315*(5811), 515–518.

Tymula, A. and H. Plassmann (2016). Context-dependency in valuation. *Current Opinion in Neurobiology 40*, 59–65.

Wainwright, M. J., O. Schwartz, and E. P. Simoncelli (2002). Natural image statistics and divisive normalization: Modeling nonlinearity and adaptation in cortical neurons. In R. P. Rao, B. A. Olshausen, and M. S. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function*, pp. 203–222. Cambridge, MA: MIT Press.

Watabe-Uchida, M., N. Eshel, and N. Uchida (2017). Neural Circuitry of Reward Prediction Error. *Annual Review of Neuroscience 40*, 373–394.

Webb, R. (2020). The (neural) dynamics of stochastic choice. *Management Science 65*(1).

Webb, R., P. W. Glimcher, and K. Louie (2020, jul). The Normalization of Consumer Valuations: Context-Dependent Preferences from Neurobiological Constraints. *Management Science*.

Wilcox, N. (2015). Unusual Estimates of Probability Weighting Functions.

Wilson, H. R. and R. Humanski (1993). Spatial frequency adaptation and contrast gain control.

Woodford, M. (2012, may). Prospect Theory as efficient perceptual distortion. *American Economic Review 102*(3), 41–46.

Yamada, H., K. Louie, A. Tymula, and P. W. Glimcher (2018). Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nature Communications*.

Appendix A Empirical Fit of the Model

As a demonstration, we fit the model to a dataset that includes the risky decisions of 2,530 individuals. These individuals visited the National Academy of Sciences Museum in Washington, DC where, as a part of the LifeLab exhibit, they had an opportunity to assess their own risk attitudes. Risk attitudes were measured using a standard economic task in which participants are asked to choose between a certain option ($5 for sure) and a lottery with changing reward ($5, $8, $20, $50, and $125) and a changing probability of winning this reward (13%, 25%, 38%, 50%, and 75%) (Levy et al., 2010). The task involved a total of 40 decisions per individual — 25 in the risky condition and 15 in the ambiguous condition. Here we analyze data from the risky condition only. One randomly selected participant per month was paid based on one of his/her own choices that was randomly selected as well. Payments were delivered as gift cards. For more details on the procedure, see Glimcher and Tymula (2017).

In this dataset, we treat expectations as a free parameter rather than estimate it from the history of rewards because of the presence of ambiguous rewards that we do not include in the analysis. To begin, we estimate ESVT parameters and compare our model fit to Prospect Theory and Expected Utility. In our estimation procedures, we follow Harrison (2008). For ESVT, we assume the value function as specified in our model in Equation 1. For Prospect Theory, we assume a CRRA utility function (since we did not examine losses) ($u(x) = x^r$) and a probability weighting function as in Lattimore et al. (1992) ($w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$). For Expected Utility, we assume CRRA utility function ($u(x) = x^r$) and no distortion in probability weighting.

For all the models, we then created an index that says that an individual selects the risky option whenever the expected value of the risky option (v_r) is larger than the value of the safe option (v_s) subject to some additive stochastic error $\epsilon \sim N(0, \sigma^2)$:

$$v_r - v_s + \epsilon > 0$$

We related this latent index to our choice data using a logistic choice function where the probability of choosing the risky lottery is given by:

$$P(x, p) = \frac{1}{1 + exp(-(v_r - v_s)/\sigma)}$$

We fit the data using a maximum likelihood procedure, clustering standard errors on the subject level. Table C1 presents the results. Out of all the models, ESVT performs the best according to Akaike and Bayesian information criteria. Prospect Theory comes second.

Thus in this gains-only dataset, ESVT performs well as an alternative to Prospect Theory. Of course, much more detailed comparisons between Prospect Theory and ESVT will need to be made.

Table A1 *Paramter estimates of ESVT, Prospect Theory, and Expected Utility. Standard errors clustered on participant are in parenthesis. Akaike and Bayesian information criterion is in the last two rows.*

A. ESVT			B. Prospect Theory			C. Expected Utility	
α	1.153		r	0.185		r	0.456
	(0.019)			(0.014)			(0.004)
M	12.248		δ	3.332		*noise*	0.739
	(0.351)			(0.300)			(0.014)
noise	0.084		γ	0.661		*AIC*	57632
	(0.002)			(0.019)		*BIC*	576511
AIC	55427		*noise*	0.198			
BIC	55456			(0.020)			
			AIC	56813			
			BIC	56849			

Interestingly, in our sample, we do not see any effect of age on risk attitudes using the standard CRRA utility function. One could infer from this that the utility function does not change as people age or vary with age cohorts. Fitting our participants' choices with ESVT, however, we find evidence that suggests that such a conclusion might be premature. Consistent with lifetime wealth effects, we find that expectations increase with age across our population. Predisposition, on the other hand, decreases. These two changes push risk attitudes in opposite directions. The decline in predisposition is pushing older people toward more risk averse behavior while an increase in expectations is pushing toward more risk tolerant behavior. As these effects counteract one another behavior remains the same, but the subjective value function underlying these choices changes.