Reference-Dependent Preferences*

Ted O'Donoghue[†] Charles Sprenger[‡]

Cornell University UC San Diego

June 22, 2018

Abstract

In this chapter, we present theories and applications of reference-dependent preferences. We provide some historical perspective, but also move quickly to the current research frontier, focusing on developments in reference dependence over the last 20 years. We present a number of worked examples to highlight the broad applicability of reference dependence. While our primary focus is gain-loss utility, we also provide a short treatment of probability weighting and its links to reference dependence.

*We are grateful to the editors, Stefano DellaVigna, Doug Bernheim, and David Laibson, as well as to Nick Barberis, Alex Rees-Jones, Ori Heffetz, Botond Koszegi, and Keith Ericson for comments and valuable conversations. We also thank participants at the Handbook of Behavioral Economics Workshop, held at Stanford University in August 2016.

[†]Ted O'Donoghue, Zubrow Professor of Economics, Department of Economics, Cornell University, 482 Uris Hall, Ithaca, NY 14853. Email: edo1@cornell.edu

[‡]Charles Sprenger, Rady School of Management and Department of Economics, University of California, San Diego, Wells Fargo Hall, Room 4W106, 9500 Gilman Drive #0553, La Jolla, CA 92093-0553. Email: csprenger@ucsd.edu

Contents

1 Introduction

Reference dependence has been a core topic in behavioral economics. The theory of reference-dependent preferences—born out of introspection and observation—captures a central intuition that outcomes are not experienced on an absolute scale, but rather experienced relative to some point of reference. Moreover, losses relative to the reference point are felt more severely than commensurate gains. This notion of "loss aversion" not only rationalizes prominent deviations from the canonical model of expected utility over final wealth, but also has proven remarkably helpful for interpreting a broad swathe of economic behaviors.

This chapter has four principal aims. First, we discuss the intellectual history from decision making under uncertainty that led to the development of models of reference-dependent preferences. Second, we attempt to move readers relatively quickly to the research frontier, devoting substantial text to developments in reference dependence over the last 20 years. Third, we provide a number of worked examples to highlight the potential for the broad applicability of reference dependence—from risk taking, to labor supply, to consumer behavior. Finally, we attempt to evaluate the literature, identifying potential gaps in our collective understanding, and pointing towards valuable new research areas.

Before proceeding, some notes to the reader may be helpful for navigating the text. First, as in the opening paragraph, we interpret "reference-dependent preferences" to mean models with gain-loss utility derived from how realized outcomes compare to some reference point. There exist other models that might also be called "reference-dependent" but which are not articulated in terms of gains and losses—e.g., models of habit formation, inequity aversion, state-dependent preferences, salience, and anticipated regret. This chapter does not cover such models.

Second, reference dependence and loss aversion emerged from the literature on decision making under uncertainty, and thus much of the literature is tied to choices over monetary risk and associated risk preferences. However, reference-dependent preferences have proven to be much more broadly applicable. Many applications come from outside the domain of monetary risk, and indeed often involve completely riskless choice.

Third, whereas Kahneman and Tversky (1979) developed “prospect theory” to include both reference-dependent preferences and nonlinear probability weighting, the subsequent literature has often investigated these two phenomena separately. For the most part, this chapter does so as well—studying the nature and implications of reference-dependent preferences without reference to probability weighting.Recent research, though, has started to identify ways in which reference-dependent preferences can generate predictions that are similar to (and sometimes equivalent to) predictions from probability weighting. Hence, before concluding, we provide a short treatment of probability weighting and discuss these connections.[1]

Fourth, except for a short discussion in the concluding section, this chapter takes a positive approach to reference-dependent preferences. In other words, we focus on the behavioral predictions of various reference-dependent preferences, and the extent to which those behavioral predictions are consistent with observed behavior. In terms of language, then, we use the term “preferences” to refer to models that predict what people do, and not what people would like to do.

Finally, this chapter will hopefully provide benefits to both novice and more advanced readers. Reading from the beginning will provide an overview of the basic models of expected utility and reference dependence in Sections 2 and 3. Readers already versed in these models may want to skim this material and get more quickly to applications in Section 4. Section 5 then provides a treatment of expectations-based reference dependence, and Section 6 describes applications that make use of expectations-based reference dependence. Section 7 provides an overview of reference-dependent “news” utility wherein a person experiences gain-loss utility from a change in beliefs about future consumption. Section 8 provides the above noted treatment of probability weighting, and Section 9 provides concluding thoughts.

[1] More complete treatments of probability weighting include Harless and Camerer (1994); Camerer (2000); Fehr-Duda and Epper (2012); and Barberis (2013).

2 Expected Utility

2.1 Overview and Some History

The foundations of modern theory of decision making under uncertainty begin in the eighteenth century. The "St. Petersburg paradox," discussed initially in correspondence by the Bernoulli cousins, Daniel and Nicolas, and the mathematicians Pierre Montmort and Gabriel Cramer (Bernouilli 1738, 1954), contradicted the idea that people ought to care only about expected value. The paradox considers Peter and Paul. Peter proposes to toss a coin until it lands on heads and give Paul 2^n ducats if heads is first observed on the n-th toss. Though such a prospect has infinite expected value, Paul is presumed unlikely to require an infinite sum to sell this prospect (contradicting expected-value maximization).

Daniel Bernoulli's proposed resolution (Bernouilli 1954) was that, rather than care about the expected value of the monetary outcome, people instead care about the expected utility associated with the monetary outcome. Moreover, Bernoulli identified the tight connection between diminishing marginal utility for money and risk preference. Two centuries later the expected-utility (EU) model took a central place in economics due to the work of Savage and Samuelson (Savage 1953; Samuelson 1952, 1953). The resulting expected-utility theorem is among the most elegant results in all of economics, providing a set of axioms for preferences over prospects, the satisfaction of which ensures that a person's behavior is consistent with the EU model.

In this section, we review the EU model and some important issues that are relevant for the literature on reference-dependent preferences. For more complete textbook treatments, see Mas-Colell et al. (1995); Varian (1992); and Kreps (1990).

2.2 The EU Model and Risk Aversion

We consider prospects of the form $L \equiv (x_1, p_1; x_2, p_2; ...; x_N, p_N)$, where outcome x_n occurs with probability p_n and $\sum_{n=1}^{N} p_n = 1$. According to the EU model, when facing a choice set $\mathcal{L}$, a person should choose the $L \in \mathcal{L}$ that yields the largest expected utility—that is, the $L \in \mathcal{L}$ that maximizes

$$U(L) \equiv \sum_{n=1}^{N} p_n u(x_n) \tag{1}$$

where $u(\cdot)$ is a utility function defined over outcomes, unique up to an affine transformation.

Work by Savage and Samuelson (Savage 1953; Samuelson 1952, 1953) identified a simple set of axioms that imply the EU model. Specifically, as long as preferences over prospects are complete, continuous, transitive, and satisfy the independence axiom, then those preferences can be represented by the EU model. These axioms represent a set of testable properties of the EU model. The axiom that has received the most attention is the independence axiom.

Independence Axiom: The preference relation $\succsim$ satisfies the independence axiom if for all L, L', L'' and $\alpha \in (0, 1)$

$$L \succsim L' \iff \alpha L + (1-\alpha)L'' \succsim \alpha L' + (1-\alpha)L''.$$

As in standard utility theory, completeness, continuity, and transitivity serve to deliver a well-behaved function $U(\cdot)$. The independence axiom then implies that $U(\cdot)$ takes the EU form, and in particular that $U(\cdot)$ is linear in probabilities.[2]

The EU model does not place any restriction on $u(\cdot)$. It turns out that the shape of $u(\cdot)$ is tightly linked to one's risk preferences. We use the following definitions of risk preferences:

[2] Expected utility became known as von Neumann-Morgenstern (vNM) preferences after the publication of von Neumann and Morgenstern (1944). Independence, however, was not among the discussed axioms, but rather implicitly assumed. Samuelson (1952, 1953) discusses the resulting confusion and his suspicion of an implicit assumption of independence in the vNM treatment. Samuelson's suspicion was then confirmed in a note by Malinvaud (1952). For an excellent discussion of the history of the independence axiom, see Fishburn and Wakker (1995).

Definition: A person is *globally risk-averse* if, for any lottery L, she prefers a certain payment equal to the expected value of L over the lottery L itself. A person is *locally risk-averse over range* $[x', x'']$ if, for any lottery L with support a subset of $[x', x'']$, she prefers a certain payment equal to the expected value of L over the lottery L itself.

Definition: A person is *globally risk-seeking* if, for any lottery L, she prefers the lottery L over a certain payment equal to the expected value of L. A person is *locally risk-seeking over range* $[x', x'']$ if, for any lottery L with support a subset of $[x', x'']$, she prefers the lottery L over a certain payment equal to the expected value of L.

Definition: A person is *globally risk-neutral* if, for any lottery L, she is indifferent between the lottery L and a certain payment equal to the expected value of L. A person is *locally risk-neutral over range* $[x', x'']$ if, for any lottery L with support a subset of $[x', x'']$, she is indifferent between the lottery L and a certain payment equal to the expected value of L.

Given these definitions, the following results are well known:

Result: Under the EU model, a person is globally risk-averse if and only if $u(\cdot)$ is globally concave, and she is locally risk-averse over range $[x', x'']$ if and only if $u(\cdot)$ is concave over range $[x', x'']$.

Result: Under the EU model, a person is globally risk-seeking if and only if $u(\cdot)$ is globally convex, and she is locally risk-seeking over range $[x', x'']$ if and only if $u(\cdot)$ is convex over range $[x', x'']$.

Result: Under the EU model, a person is globally risk-neutral if and only if $u(\cdot)$ is globally linear, and she is locally risk-neutral over range $[x', x'']$ if and only if $u(\cdot)$ is linear over range $[x', x'']$.

Hence, the EU model is well defined for $u(\cdot)$ with any shape, and indeed is well defined if the shape of $u(\cdot)$ changes with x. For instance, if $u(\cdot)$ were concave for small x, linear for intermediate x, and convex for large x, then the person's behavior would be risk-averse when choosing between lotteries that involve only small outcomes, risk-neutral when choosing between lotteries that involve only intermediate outcomes, and risk-seeking when choosing between lotteries that involve only large outcomes. Nonetheless, EU was originally introduced to capture the intuition that humans are generally risk-averse, and most applications of EU assume that $u(\cdot)$ is globally concave. The remainder of our discussion in this section considers this case.

Figure 1 provides two graphical depictions of indifference curves under EU. Panel A provides an example of the Marschak-Machina triangle that depicts preferences for three-outcome lotteries when the three outcomes are fixed and the probabilities vary. Consider prospects of the form $L = (x_1, p_1; x_2, 1 - p_1 - p_3; x_3, p_3)$ with $x_1 < x_2 < x_3$. Fixing x_1, x_2, and x_3, all such prospects can be represented in (p_1, p_3) space. Under EU, indifference curves in this space have slope

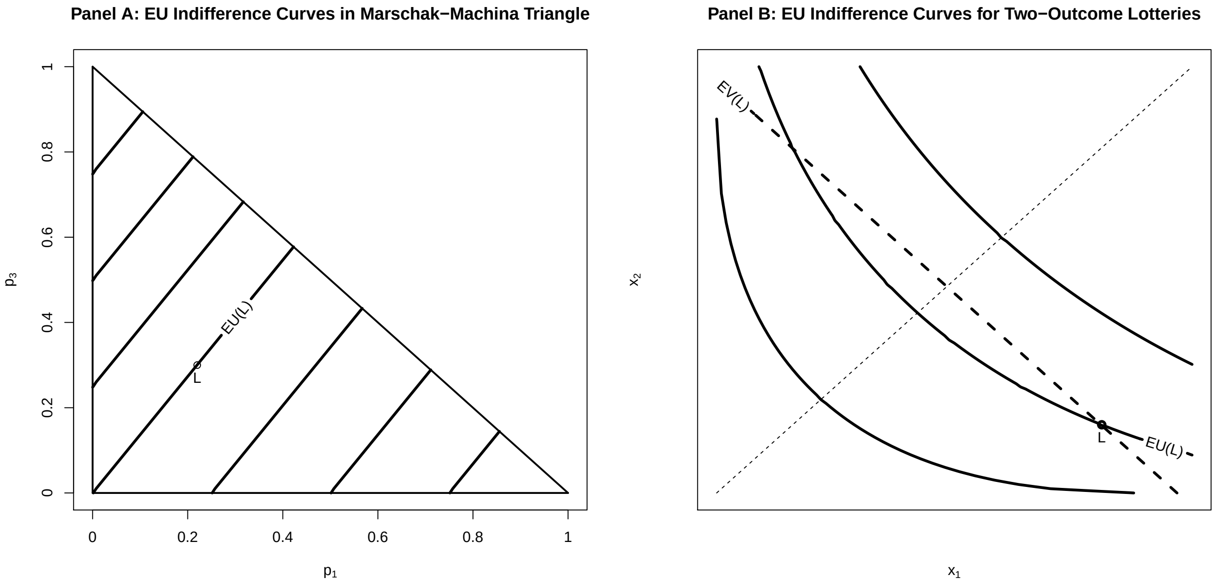
$$\left.\frac{dp_3}{dp_1}\right|_{U(L)=\bar{U}} = \frac{u(x_2) - u(x_1)}{u(x_3) - u(x_2)}.$$

Because the slope is independent of p_1 and p_3, these indifference curves are linear and parallel. Moreover, risk aversion is captured in the slope of these indifference curves. If the person is risk-neutral and thus $u(\cdot)$ is linear, the slope is $\frac{x_2 - x_1}{x_3 - x_2}$. In contrast, if the person is risk-averse and thus $u(\cdot)$ is concave, the slope is steeper than $\frac{x_2 - x_1}{x_3 - x_2}$, and a more risk-averse person will have steeper indifference curves.

Panel B of Figure 1 considers preferences over binary prospects when the probabilities are fixed and the outcomes vary. Consider prospects of the form $L = (x_1, p_1; x_2, 1 - p_1)$. Fixing p_1, all such prospects can be represented in (x_1, x_2) space. Under EU and assuming that $u(\cdot)$ is twice differentiable, indifference curves in this space have slope

$$\left.\frac{dx_2}{dx_1}\right|_{U(L)=\bar{U}} = -\frac{p_1 \; u'(x_1)}{(1 - p_1) \; u'(x_2)}.$$

Here, risk aversion is captured in the shape of these indifference curves. If the person is



Note: Panel A presents EU indifference curves in the Marschak-Machina triangle for prospects of the form $L = (x_1, p_1; x_2, 1 - p_1 - p_3; x_3, p_3)$ with $x_1 < x_2 < x_3$. Panel B presents EU indifference curves for prospects of the form $L = (x_1, p_1; x_2, 1 - p_1)$. Dashed line reflects line of constant expected value (risk neutrality), which has slope $-\frac{p_1}{1-p_1}$.

Figure 1: EU Indifference Curves

risk-neutral and thus $u(\cdot)$ is linear, indifference curves are linear with slope $-\frac{p_1}{1-p_1}$. If the person is risk-averse and thus $u(\cdot)$ is concave, indifference curves are convex, and a more risk-averse person will have more convex indifference curves.

The two panels in Figure 1 represent benchmarks against which we'll later compare models of reference-dependent preferences.

2.3 Some Issues for EU

Having described the EU model, we next describe a series of issues that played a prominent role in the development of the literature on reference-dependent preferences.

Outcome domain: Our treatment of the EU model above is silent on the domain for outcomes (i.e., on what the x's are). When applying EU, it is virtually always assumed that the proper domain is to apply EU to final wealth states (or, outside the domain of monetary gambles, to final consumption bundles). This assumption is often referred to as *integration* because it says that the incremental outcomes associated with a chosen prospect must be integrated with everything else going on in the decision maker's life.

For monetary gambles, integration is typically implemented by assuming the decision maker has some prior wealth w and that prospect outcomes are expressed as increments that are added to or subtracted from that prior wealth to yield a final wealth (to be consumed). Hence, a prospect $L \equiv (x_1, p_1; x_2, p_2; ...; x_N, p_N)$ is evaluated according to

$$U(L) \equiv \sum_{n=1}^{N} p_n u(w + x_n).$$

One reason to assume integration is to make EU a portable model that can be applied in many domains—from portfolio choice and insurance purchasing to human capital accumulation and gambling—without the need to assume a new utility function in every domain. A second reason to assume integration is because economists often believe that one's risk preferences over incremental prospects are likely to vary with one's wealth, and in particular that people become less risk-averse as their wealth goes up. If so, then we must include one's initial wealth into the EU calculus. For the remainder of this chapter, we shall assume integration whenever we discuss EU.

Probabilities: Our treatment of EU above does not specify what the probabilities (the p_n's) represent. When applying EU, it is typically assumed that if there are objective probabilities known to the decision maker, then those objective probabilities will be used. Otherwise, the decision maker is assumed to formulate a set subjective probabilities that are then used.[3] In fact, empirical applications of EU often combine these approaches along with an assumption of rational expectations—specifically, by using data to generate an econo-

[3]When probabilities are subjective, the EU analog is known as subjective expected utility (SEU), with an analogous utility formulation that is linear in subjective probabilities. For axiomatic developments of SEU, see Savage (1953, 1954); Anscombe and Aumann (1963).

metrician's best estimate for the objective probabilities, and then assuming that a person's subjective beliefs correspond to those estimated objective probabilities.

Small-stakes risk neutrality, calibration, and aggregation: Under EU, a person must be approximately risk-neutral for sufficiently small stakes (see, e.g., Pratt 1964; Arrow 1971; Segal and Spivak 1990). One manifestation of this fact appears in Panel B of Figure 1: as indifference curves intersect the 45-degree line of certainty (on which $x_1 = x_2$), the slope is $-\frac{p_1}{1-p_1}$, exactly equal to the slope of the line of constant expected value. A second manifestation appears in attitudes toward positive expected value prospects: if a prospect L has a positive expected value, then any EU maximizer would always want to take a stake in that prospect as long as that stake is small enough.

Rabin (2000a) provides a calibration argument indicating that, for any plausible EU parameters, this approximate risk neutrality must hold even for somewhat larger stakes. Specifically, he proves that noticeable risk aversion over any set level of stakes implies absurd risk aversion over stakes only one or two orders of magnitude larger. He provides a series of striking examples: for instance, if an EU maximizer turns down a 50-50 bet to lose \$100 or gain \$110 at all wealth levels, then that person must also turn down any 50-50 bet to lose \$1000 or gain $\$X$ for any $\$X$, no matter how large. Intuitively, turning down the small-stakes bet at wealth w implies significant local curvature of $u(\cdot)$ around w; however, turning down that same bet over a range of wealth levels implies the same significant local curvature holds for all w, resulting in massive curvature for larger stakes.

The Rabin critique implies that EU is inconsistent with many plausible patterns of risk preferences. For instance, consider a person who is indifferent to accepting vs. rejecting a 50-50 bet to lose \$100 or gain \$110 for a range of wealths and who is also indifferent to accepting vs. rejecting a 50-50 bet to lose \$1000 or gain \$1100 for that same range of wealths.[4] This pattern seems perfectly plausible—and it would certainly seem that we'd want a theory that would at least permit this pattern as a possibility. However, Rabin's

[4]This range need not cover all wealths—e.g., some range $[\bar{w}, \bar{w} + \$5000]$ will do.

theorem reveals that this pattern is inconsistent with EU.

The implications of small-stakes risk aversion for larger-stakes prospects was also noted by Samuelson (1963). He starts with an interesting historical anecdote in which he offered lunch colleagues a simple bet on a fair coin: \$200 if the side of their choosing comes up, −\$100 if not. A colleague responded, "I won't bet because I would feel the \$100 loss more than the \$200 gain. But I'll take you on if you promise to let me make 100 such bets" (p.2). Samuelson proved that his colleague's response is inconsistent with EU. Specifically, he proved that, under EU, if one rejects a prospect, then one must also reject an aggregate bet in which that prospect is independently played multiple times (as long as the rejection of the prospect would occur at any wealth level that could obtain in the aggregate bet). The logic is similar to that of the Rabin critique.

Rabin (2000a) and Samuelson (1963) provide different interpretations for their results. Samuelson argued that his colleague had a mistaken view of the aggregate bet and wouldn't want it if he truly understood it.[5] Rabin, in contrast, viewed noticeable but reasonable small- and moderate-stakes risk aversion as a typical behavior. Indeed, Rabin and Thaler (2001) argue that the behavior of Samuelson's colleague is a pattern we might expect many people to exhibit. One might reject the single bet for exactly the reason Samuelson's colleague expressed. At the same time, an aggregate bet made up of 100 independent plays is a quite attractive bet that offers a tiny chance (0.00044) of losing any money, and a large chance of earning a significant amount of money—e.g., the probability of winning at least \$2000 is 0.982.

Allais' paradoxes: Allais (1953b) presented two famous "paradoxes"—that is, patterns of choice inconsistent with EU. Both revolve around violations of the independence axiom and the corresponding consequence of linearity in probabilities.[6] Allais' more famous paradox—

[5]Samuelson (1963) emphasizes that, under EU, his colleague should not have wanted to aggregate the prospect but rather he should have wanted to subdivide the risk, say, into 100 bets of −\$1 against \$2.

[6]Interestingly, Allais' presentation of Allais (1953a) was in the same session as Samuelson's presentation of Samuelson (1953) and the day after Savage's presentation of Savage (1953) at the Colloque Internationale d'Econométrie in Paris in May of 1952 (*Colloques Internationaux du Centre National de la Recherche*

often labeled the "common-consequence paradox"—poses a pair of choices:

Choice 1: $A \equiv (100M, 1)$ vs. $B \equiv (500M, .10; 100M, .89; 0, .01)$.

Choice 2: $C \equiv (100M, .11; 0, .89)$ vs. $D \equiv (500M, .10; 0, .90)$.

When people introspect about these choices, many conclude that they would prefer A over B and that they would prefer D over C. However, this pattern is inconsistent with EU. In particular, according to the independence axiom, taking the common consequence of a .89 chance of $1M$ in A and B and converting it into a common consequence of a .89 chance of 0 to create C and D should not change one's preference.

Allais' second paradox—often labeled the "common-ratio paradox"—also poses a pair of choices:

Choice 1: $A \equiv (100M, 1)$ vs. $B \equiv (500M, .98; 0, .02)$.

Choice 2: $C \equiv (100M, .01; 0, .99)$ vs. $D \equiv (500M, .0098; 0, .9902)$.

Again, when people introspect about these choices, many conclude that they would prefer A over B and D over C, and again this pattern is inconsistent with EU. The independence axiom implies that scaling down the likelihood of non-zero amounts by a common ratio should not change one's preference—here, the likelihoods of winning $100M$ and $500M$ in A and B are both scaled down by a factor .01 to create C and D.

Moving forward: Behavioral economics has been deeply influenced by the issues highlighted by Rabin, Samuelson, and Allais. A major concern is having a model that is consistent with (i) noticeable and reasonable small- and moderate-stakes risk aversion; (ii) the behavior of Samuelson's colleague; and (iii) independence violations à la Allais. Hence, we'll return to these phenomena as we develop models of reference-dependent preferences.

Scientifique (Econometrie) 40 1953).

3 Reference-Dependent Preferences

3.1 Overview and Some History

As the EU model was taking its place as the standard economic approach to decision making under uncertainty, other researchers started to highlight two types of potential deviations from EU: those that relate to linearity in probabilities, and those that relate to the utility function.

In parallel to work in economics by Allais, psychologists began to explore the possibility that people might hold "subjective probabilities" that need not correspond to objective probabilities.[7] Preston and Baratta (1948) and Edwards (1955) used experiments to study how subjective probabilities compare to objective probabilities. Later, Edwards (1962) described a process by which objective probabilities were replaced by decision weights which need not respect EU's linearity requirements.

At the same time, economists were speculating on deviations from a globally concave utility function defined over final wealth states. Friedman and Savage (1948) pointed to the existence of decision makers who simultaneously purchase lottery tickets at less than fair odds and insurance for moderate risks. To explain this behavior, they suggested that the utility function over final wealth states might have both concave and convex regions. Motivated by Friedman and Savage (1948), Markowitz (1952) suggested an alternative solution: perhaps instead of utility being defined over final wealth states, utility is defined over gains and losses relative to present wealth.[8] Based on thought experiments—specifically, how people would choose between gaining a certain amount $\$x$ vs. a one in ten chance of gaining $\$10x$, and how people would choose between losing a certain amount $\$x$ vs. a one in ten chance of losing $\$10x$—Markowitz hypothesized a pattern of risk seeking followed by risk aversion as

[7] In this work, the usage of "subjective probability" differs from that among economists in that (i) subjective probabilities might deviate from objective probabilities even when the latter are known to the decision maker, and (ii) subjective probabilities are permitted not to sum to one.

[8] Markowitz (1952) permits that reference wealth might shift if the person has recently experienced a windfall gain or loss.

the stakes increased for gains, and the opposite pattern for losses. He further hypothesized that people dislike symmetric fair bets, and thus suggested that $|u(-x)| > u(x)$.

In their model of "prospect theory," Kahneman and Tversky (1979) combined and built upon early thinking on both lines of inquiry. As a first step, they collected data on people's responses to hypothetical choice problems, from which they stipulated a set of properties that choices seemed to satisfy. Much of the data they gathered was inspired by prior work—e.g., their paired problems 1 and 2 are a variant of Allais' common-consequence example, their paired problems 3 and 4 are a variant of Allais' common-ratio example, and their various paired problems x and x' in the gain vs. loss domain are analogous to Markowitz's thought experiments.

Kahneman and Tversky then set out to develop a model that would accommodate all of these properties. Specifically, they propose that gambles of the form $L = (x_1, p_1; x_2, p_2; 0, p_3)$ are evaluated according to

$$V(L) \equiv \pi(p_1)v(x_1) + \pi(p_2)v(x_2), \tag{2}$$

where $\pi(p)$ represents a probability weighting function that converts objective probabilities into decision weights, and $v(x)$ is a value function defined over gains and losses (note that $v(0) = 0$ is assumed—see below).[9]

In the years that followed, the literature found promise in exploring both nonlinear probability weighting and a value function defined over gains and losses, although often independently. In particular, gain-loss utility received prominent attention in the behavioral-economics literature, where research on topics such as small-stakes risk aversion, exchange behavior, and labor supply drew out the implications of loss aversion. Probability weighting received prominent attention in the decision-theory literature, where research on EU devi-

[9] In the original formulation, Kahneman and Tversky (1979) explicitly limit their domain to gambles with at most two non-zero outcomes, and they also assume a slightly different functional for gambles that are sure to yield a gain or sure to yield a loss. The primary reason for such restrictions was to eliminate clearly incorrect predictions due to features of the probability weighting function. Later variants of probability weighting eliminate the need for such fixes, as we discuss in Section 8.

ations à l'Allais and other violations of independence were rationalized by models of probability distortion. The approach of isolating the predictions of one particular phenomenon (e.g., the predictions of gain-loss utility or the predictions of probability weighting) is quite common in economics, and we follow the same path here. The remainder of this section and Sections 4 through 6 will focus on the reference-dependent aspects of the model, while Section 8 will provide discussion and details on nonlinearity in probabilities.[10]

3.2 Risky Choice and the Value Function

We begin with a simple model of reference-dependent preferences in the spirit of prospect theory. Again, consider prospects of the form $L \equiv (x_1, p_1; x_2, p_2; ...; x_N, p_N)$, and let r be a reference point around which a person defines gains and losses—i.e., $x_n > r$ is a gain while $x_n < r$ is a loss. Prospect L is evaluated according to

$$V(L|r) \equiv \sum_{n=1}^{N} p_n v(x_n - r). \tag{3}$$

Kahneman and Tversky (1979) suggest that a natural candidate for the reference point, especially in simple experimental gambles, is one's initial wealth. Given that the x_n's are defined as increments to wealth, this implies $r = 0$ as in equation 2 above. Equation 3 permits other reference points as well, and, as we shall see, the question of what is the reference point has been a major topic in the literature.[11]

Kahneman and Tversky (1979) argue that the value function, $v(\cdot)$, is an increasing function with three key features:

[10]Some research explores probability weighting and gain-loss utility in concert, often by applying directly the model of cumulative prospect theory developed in Tversky and Kahneman (1992). Several prominent examples come from the field of finance—see Chapter X of this handbook for a detailed summary of this work.

[11]Kahneman and Tversky (1979) recognize this issue, as they write:

> "So far ... the reference point was taken to be the status quo, or one's current assets. Although this is probably true for most choice problems, there are situations in which gains and losses are coded relative to an expectation or aspiration level that differs from the status quo."

1. *Zero value of reference point:* $v(0) = 0$.

2. *Diminishing sensitivity:* $v''(x) < 0$ for $x > 0$, but $v''(x) > 0$ for $x < 0$.

3. *Loss aversion:* For $x > 0$, $v(x) < -v(-x)$ and $v'(x) < v'(-x)$.

The first feature is really just a normalization. The second feature says that people have a diminishing sensitivity to the magnitude of gains and losses. For gains ($x > 0$), this property is analogous to the assumption of diminishing marginal utility for final wealth in EU. For losses ($x < 0$), in contrast, to have diminishing sensitivity to marginal losses as losses get larger is equivalent to the value function being convex in this domain.

The third feature says that, when comparing gains and losses, losses loom larger than commensurate gains. The assumption of loss aversion has two features: a loss has larger magnitude utility consequences, and larger marginal utility consequences than an equally sized gain. Interestingly, while loss aversion has played an especially important role in the literature, it was the piece of the theory about which Kahneman and Tversky were least confident (see Kahneman (2000)). Moreover, whereas their exposition merely posits $v(x) < -v(-x)$ and $v'(x) < v'(-x)$ as above, their Figure 3 further suggests a kink at $x = 0$. This kink has been assumed in most of the research that followed.

To illustrate, consider two functional forms that might be used:

<ins>Tversky and Kahneman (1992)</ins>:

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda(-x)^{\beta} & \text{if } x \leq 0 \end{cases}$$

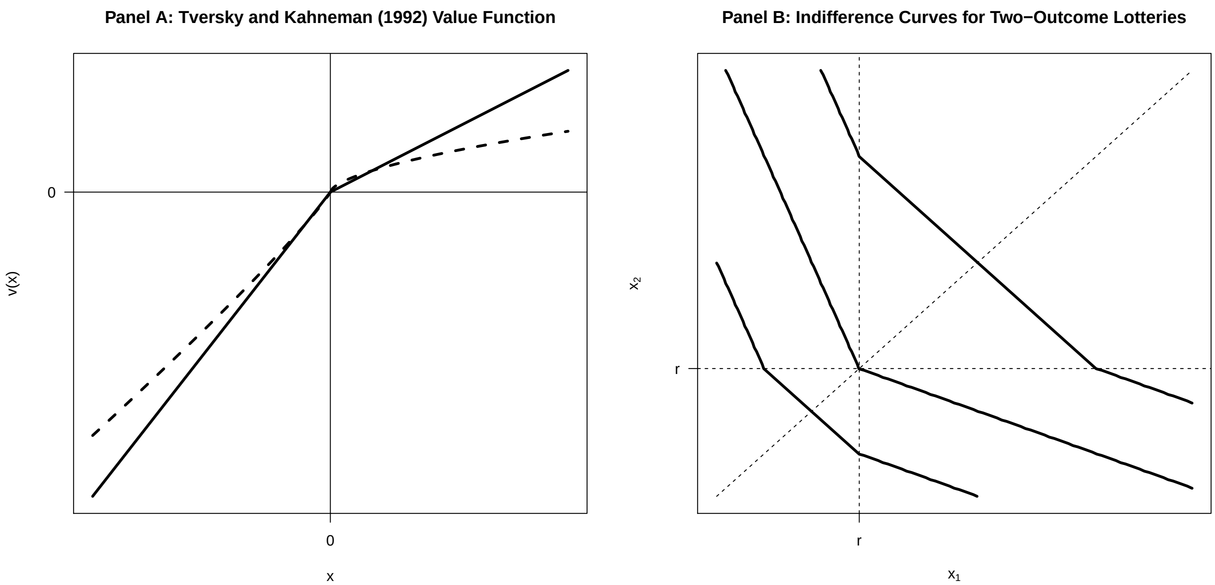
where $\alpha, \beta \in (0, 1]$ and $\lambda \geq 1$

<ins>Two-part linear</ins>:

$$v(x) = \begin{cases} x & \text{if } x \geq 0 \\ \lambda x & \text{if } x \leq 0 \end{cases}$$

where $\lambda \geq 1$

Figure 2, Panel A depicts the Tversky and Kahneman functional form. This functional form captures all three properties from Kahneman and Tversky (1979). Specifically, the parameter α captures the degree of diminishing sensitivity in the gain domain ($\alpha = 1$ implies no diminishing sensitivity), the parameter β captures the degree of diminishing sensitivity in the loss domain ($\beta = 1$ implies no diminishing sensitivity), and the parameter λ captures the degree of loss aversion when comparing gains to losses ($\lambda = 1$ implies no loss aversion).



Note: Panel A depicts the Tversky and Kahneman (1992) value function given a reference point r and loss aversion $\lambda > 1$. Dashed line corresponds to formulation with $\alpha < 1$ and $\beta < 1$; solid line corresponds to formulation with $\alpha = \beta = 1$. Panel B presents indifference curves for prospects of the form $L = (x_1, p_1; x_2, 1 - p_1)$ when $\alpha = \beta = 1$.

Figure 2: Reference-Dependent Preferences

A value function with these properties has important implications for risky choice. Diminishing sensitivity in the gain domain implies that, when making choices between prospects where all possible outcomes are (weakly) in the gain domain, the person will be risk-averse. Intuitively, the structure of equation 3 is identical to equation 1, and thus because the value function is locally concave in the gain domain, the person will be locally risk-averse in the gain domain. Analogously, diminishing sensitivity in the loss domain implies that, when making choices between prospects when all possible outcomes are (weakly) in the loss domain, the person will be risk-seeking.[12] Importantly, loss aversion is irrelevant to both of

[12]An often emphasized implication of prospect theory is the "four-fold pattern of risk preferences": (i) people are risk-averse over moderate-probability gains, (ii) people are risk-seeking over moderate-probability losses, (iii) people are risk-loving over small-probability gains, and (iv) people are risk-averse over small-probability losses. Under prospect theory, (i) and (ii) are driven by diminishing sensitivity in the value

these implications. Rather, loss aversion becomes relevant only when making choices between prospects when gains and losses are both possible, and it creates an additional source of risk aversion for such choices.

Much of the literature that has applied prospect theory has focused on the specific implications of loss aversion. The two-part-linear functional form assumes there is no diminishing sensitivity, and thus isolates these implications. In most of our development, we follow this tradition and focus only on the two-part linear form. In Section 9 we will return to diminishing sensitivity in our discussion.

Figure 2, Panel B further highlights the implications of the two-part-linear functional form by presenting its indifference curves in binary lottery (x_1, x_2) space given an exogenous reference point r. Consider first the indifference curve through (r, r). Note that this indifference curve is kinked at (r, r), implying that, unlike under EU, a loss-averse individual need not be risk-neutral over infinitesimally small stakes. Next, consider indifference curves to the right or the left of (r, r). In each case, there are three regions. In the middle region, where either $x_1, x_2 > r$ or $x_1, x_2 < r$, the slope is equal to $-\frac{p_1}{(1-p_1)}$ and thus the person would be locally risk-neutral. In contrast, in the region where $x_1 > r > x_2$, the slope is equal to $-\frac{p_1}{\lambda(1-p_1)}$, which is smaller (in magnitude) than $\frac{-p_1}{(1-p_1)}$, reflecting that losses are felt more severely than commensurate gains. Analogously, in the region where $x_1 < r < x_2$, the slope is equal to $-\frac{\lambda p_1}{(1-p_1)}$, which is larger (in magnitude) than $\frac{-p_1}{(1-p_1)}$.

3.3 Riskless Choice

Although prospect theory was initially formulated as a model for decision making under uncertainty, it was quickly recognized that a reference-dependent value function would have a number of direct applications for riskless choice (Thaler (1980)). Indeed, many applications of reference-dependent preferences focus on the riskless evaluation of gains and losses (see

function (as described in the text), while (iii) and (iv) are driven by overweighting of small probabilities in the probability-weighting functions (as described in Section 8).

Section 4).

Extending the theory into the domain of riskless choice requires additional modeling choices.[13] First, whereas models of decision making under uncertainty focus on how people trade off consumption of a single good (money) across different states, models of riskless choice focus on how people trade off consumption of one good for consumption of another good. In other words, both consumption and reference points might be multi-dimensional. Second, the fact that consumption and reference points might be multi-dimensional requires the researcher to decide how to model sensations of gain and loss across dimensions. Third, in the domain of riskless choice, it is important to incorporate the intrinsic utility of different goods—i.e., the utilities that would fully determine behavior of a standard agent.

These modeling exigencies have led to the following formulation of reference dependence for riskless choice. Suppose a person is choosing between riskless consumption bundles of the form $\mathbf{x} \equiv (x^1, ..., x^N)$ with reference bundle $\mathbf{r} \equiv (r^1, ...r^N)$, where each x^n represents a different good (e.g., mugs, candy, cars, money) and each r^n represents the corresponding reference point. The decision maker evaluates bundle $\mathbf{x}$ according to

$$U(\mathbf{x}|\mathbf{r}) \equiv u(\mathbf{x}) + v(\mathbf{x}|\mathbf{r})$$

where $u(\cdot)$ represents intrinsic utility and $v(\cdot)$ represents reference-dependent sensations of gain and loss. While more general approaches are possible, the typical approach assumes additive separability across goods. In other words, $u(\mathbf{x}) \equiv \sum_{n=1}^{N} u^n(x^n)$ where u^n is an intrinsic utility function for good n, and $v(\mathbf{x}|\mathbf{r}) \equiv \sum_{n=1}^{N} v^n(x^n|r^n)$ where v^n is a gain-loss utility function for good n.

[13] Tversky and Kahneman (1991) provide the first formal treatment of gain-loss utility for riskless choice. Later approaches evolved through specific applications and new theories.

For $v^n(x^n|r^n)$, Tversky and Kahneman (1991) suggest functional form

$$v^n(x^n|r^n) \equiv \begin{cases} \eta^n(x^n - r^n) & \text{if} \quad x^n \geq r_n \\ \eta^n \lambda^n (x^n - r^n) & \text{if} \quad x^n \leq r^n \end{cases} \tag{4}$$

where $\eta^n \geq 0$ captures the importance of gain-loss utility relative to intrinsic utility, and $\lambda^n \geq 1$ captures the degree of loss aversion.[14] Though restricted to be the two-part linear form, this formulation retains a great degree of flexibility. The magnitude of loss aversion (λ^n) is permitted to differ across goods—e.g., when studying the willingness to pay for a mug, this formulation permits that one could have gain-loss utility over mugs but not over money. In addition, the overall magnitude of gain-loss utility (η^n) is also permitted to differ across goods—e.g., one could assume that the gain utility associated with obtaining a mug is larger than the gain utility associated with obtaining a car. As we'll see in Section 4, applications of gain-loss utility have taken advantage of both forms of flexibility.

Koszegi and Rabin (2006) restrict this freedom, suggesting instead that

$$v^n(x^n|r^n) \equiv \mu(u^n(x^n) - u^n(r^n)) \quad \text{where } \mu(z) = \begin{cases} \eta z & \text{if } z \geq 0 \\ \eta \lambda z & \text{if } z \leq 0. \end{cases} \tag{5}$$

This formulation makes two key disciplining assumptions. First, it assumes that gains and losses are defined in terms of intrinsic utilities—because it assumes that the argument in μ is the intrinsic-utility difference between consuming x^n and consuming the reference point r^n. Second, it assumes a universal gain-loss utility function—in particular, η and λ are the same for all dimensions of consumption. Together, these assumptions eliminate the flexibility discussed above.[15] The Koszegi-Rabin gain-loss function has the virtue that, relative to any intrinsic utility function that one might assume in a standard model, there are only

[14] Tversky and Kahneman (1991) do not include intrinsic utility in their formulation, and thus do not include η^n. In addition, the λ^n here is equivalent to the $1/\lambda^n$ in their formulation.

[15] Most applications of the Koszegi-Rabin utility model assume this two-part-linear functional form for μ, and thus focus on the implications of loss aversion. However, it would be straightforward to permit diminishing sensitivity in μ.

two additional free parameters, η and λ. Moreover, the model nests the standard model whenever $\eta = 0$ or $\lambda = 1$.[16]

3.4 Editing and Mental Accounting

Kahneman and Tversky (1979) incorporate two stages in the choice process: an editing stage and an evaluation stage. Everything detailed above refers to the evaluation stage, which describes how a person makes a choice from a well defined set of prospects. The editing stage happens prior to the evaluation stage, and reflects how a person converts a choice problem that is presented to her into a well defined set of prospects from which she must make a choice (which she'll then do at the evaluation stage).

Kahneman and Tversky (1979) suggest a large number of operations that might occur in the editing stage. Some important examples are "coding" (redefining outcomes as gains or losses), "combination" (combining the probabilities for identical outcomes), "cancellation" (discarding of shared components), and "detection of dominance" (eliminating dominated options). In the original theory, the editing stage served two purposes. First, some of their motivating data could only be explained by people interpreting two arguably identical problems in different ways (e.g., the isolation-effect example described below). Second, the evaluation stage generates some "perverse" predictions (especially related to probability weighting—see Section 8), and they correct for these at the editing stage (e.g., by eliminating dominated options).

The second purpose is ad hoc, and fortunately subsequent work has made it less necessary. The first purpose, in contrast, may reflect a real psychology in decision making. In the world, people are presented with complex choice problems, and before one can apply any type of evaluative procedure, one must first convert the complex choice problem into a well defined

[16]With $\lambda = 1$, the model reduces to

$$v(x|r) = u(x) + \eta(u(x) - u(r)) = (1+\eta)u(x) - \eta u(r),$$

a monotone (affine) transform of $u(x)$.

set of options to evaluate. This mental procedure is often labelled "mental accounting," and any investigation of decision making must make a set of assumptions about mental accounting (such assumptions are sometimes made explicitly, but they are more often made implicitly). When applying gain-loss utility, there are two key dimensions on which one must make mental-accounting assumptions: coding and bracketing.

Coding refers to how a person encodes outcomes as gains vs. losses—or, equivalently, what is the reference point around which gains and losses are defined. To illustrate the importance of coding, consider an example of the isolation effect from Kahneman and Tversky (1979):

> Problem 11: You get 1000 for sure. In addition, choose between $(1000, .5)$ and $(500, 1)$.
>
> Problem 12: You get 2000 for sure. In addition, choose between $(-1000, .5)$ and $(-500, 1)$.

Relative to one's prior wealth, both problems involve a choice between $(2000, .5; 1000, .5)$ and $(1500, 1)$. However, the majority choose the latter in Problem 11 and the former in Problem 12.[17] Kahneman and Tversky interpret this behavior as people ignoring the shared statement of getting a fixed amount for sure and focus on the choice between the stated gambles (i.e., they view Problem 11 as a choice between $(1000, .5)$ and $(500, 1)$, ignoring the 1000 for sure).[18] This is equivalent to people using a reference point $r = w+1000$ in Problem 11 and a reference point $r = w + 2000$ in Problem 12.

When one applies gain-loss utility in complex economic applications, coding—and in particular the question of what is the reference point—is an important issue, and is left unspecified by the model. Kahneman and Tversky note that in some decisions it is appropriate to use the status quo or current asset levels, whereas in other situations an expectation or

[17]Markowitz (1952) also discussed paired problems of this type, although he introspected that people would behave the same in the two problems.

[18]With this interpretation, the majority behavior in Problems 11 and 12 is then explained by diminishing sensitivity in the value function.

an aspiration level might be more appropriate. Without a structured mechanism for determining the reference point, the model benefits from a powerful degree of freedom with many behaviors across environments potentially accommodated by a suitable definition of the reference point. As we'll see, applications often take advantage of this degree of freedom.

Bracketing refers to how a person delineates the set of choices that she faces. As a simple example, suppose a person is asked to accept or reject a lottery L (or a consumption bundle $\mathbf{x}$). On one hand, the person could evaluate the two options—accept or reject—in isolation from all the other choices that she faces in her life (narrow bracketing). On the other hand, the person could integrate this choice with all the other choices that she faces in her life and make one grand choice (broad bracketing). Or, the person could do something in between. A second dimension of bracketing revolves around the question of which options are relevant. A standard approach would assume that a person considers all options that are available. However, in complex choice problems, it seems plausible that a person might consider only a subset of the available options.

It is worth highlighting that many standard economic analyses implicitly assume narrow bracketing—e.g., any industrial organization analysis that estimates underlying preferences for an individual good while ignoring all the other goods that people consume is effectively assuming narrow bracketing. How a person brackets can of course matter in standard economic models, but it matters even more in models of gain-loss utility. The sharp sensitivity of behavior around the reference point may be muted by broader bracketing of outcomes. Much as for coding, when one applies gain-loss utility in complex economic applications, bracketing is an important issue left unspecified by the model. While there are a few papers that focus explicitly on bracketing (e.g., in the domains of labor supply and repeated risks), the issue of bracketing has been less discussed in the literature than the issue of coding.

Before leaving the topic of mental accounting, it is worth briefly distinguishing the psychological assumption of mental accounting from mere simplifying assumptions. There is a long tradition in economics of making simplifying assumptions about complex environments so as to permit tractable analysis. Typically, this is viewed as a modeling technique, and not as something that the actual agents would be doing. Mental-accounting assumptions often appear at first glance merely to be an example of the modeler making simplifying assumptions. However, the mental-accounting assumptions discussed above are often meant to reflect psychological assumptions about how the agents themselves think. Given that assumptions about the choice set can impact predictions, it is worth paying more attention to whether such assumptions reflect a plausible underlying psychology (in which case they might be justified) or whether they are really being made for simplicity (in which case we should worry about bias).

4 Applications With Exogenous Reference Points

The model of reference-dependent preferences in Section 3 applies for static choice with fixed, exogenous reference points. We now consider applications and evidence based on this approach.[19]

Some early applications of reference-dependent preferences pointed to well-known behavioral anomalies and described how such choices could be rationalized by reference-dependent preferences. One prominent domain for such efforts is simple risk preferences. For instance, after proving that EU is inconsistent with a person exhibiting noticeable and reasonable small- and moderate-stakes risk aversion, Rabin (2000a) suggests reference-dependent preferences as a natural explanation. To illustrate, consider a decision maker with a reference point at current wealth w deciding to accept or reject a 50-50 prospect over $-\$Y$ and $+\$X$. Applying equation 3 with a two-part linear value function, a reference-dependent agent

[19] For an early summary of field applications of prospect theory—both gain-loss utility and probability weighting—see Camerer (2000).

would, for any current wealth, w, accept such a bet if

$$\begin{aligned}
0.5 \cdot v(X) + 0.5 \cdot v(-Y) &> v(0) \\
0.5 \cdot (X) + 0.5 \cdot (-\lambda Y) &> 0 \\
X &> \lambda Y.
\end{aligned}$$

Hence, an agent with $\lambda = 2$, for instance, would, for all values of current wealth w, be indifferent between accepting and rejecting the following 50-50 bets:

$$\left(-\$1, \frac{1}{2}; +\$2, \frac{1}{2}\right), \left(-\$10, \frac{1}{2}; +\$20, \frac{1}{2}\right), \left(-\$100, \frac{1}{2}; +\$200, \frac{1}{2}\right), \left(-\$1000, \frac{1}{2}; +\$2000, \frac{1}{2}\right).$$

While this pattern of preferences might seem plausible, recall that Rabin (2000a) proved such preferences are inconsistent with EU. For instance, indifference to $\left(-\$1, \frac{1}{2}; +\$2, \frac{1}{2}\right)$ for all current wealths implies starkly diminishing marginal utility to the point where, if the magnitude of the potential loss is increased only a little (even \$4 is enough), the decision maker would reject any 50-50 bet no matter the potential gain.[20] In contrast, the reference-dependent agent can exhibit noticeable risk aversion over small stakes for all current wealths, while also exhibiting reasonable risk aversion over larger stakes. As such, Rabin (2000a) proposes reference dependence as a potential rationalization of noticeable and reasonable small- and moderate-stakes risk aversion.

Benartzi and Thaler (1995) demonstrate that a similar exercise can be conducted for the phenomena described by Samuelson (1963). Though Samuelson (1963) described the decisions of his colleague as an error, it seems quite plausible that many individuals would

[20] For instance, indifference to $(-1, \frac{1}{2}; +2, \frac{1}{2})$ at wealth w implies $u(w+2) - u(w) = u(w) - u(w-1)$. Defining $\Delta(x) \equiv u(w+x) - u(w+x-1)$, this becomes $\Delta(2) + \Delta(1) = \Delta(0)$, and concavity implies $\Delta(2) < \Delta(1)$ and thus $2\Delta(2) < \Delta(0)$. Indifference to this bet for all w implies $2\Delta(x+2) < \Delta(x)$ for all x. Now consider bet $(-l, \frac{1}{2}; +g, \frac{1}{2})$ with l and g even (for simplicity). The person will accept this bet only if $\sum_{x=1}^{g} \Delta(x) \geq \sum_{x=1}^{l} \Delta(1-x)$. However, $\sum_{x=1}^{g} \Delta(x) < \sum_{x=1}^{g/2} 2\Delta(2x-1) < 2\Delta(1) \sum_{x=1}^{g/2} \left(\frac{1}{2}\right)^{x-1}$ and $\sum_{x=1}^{l} \Delta(1-x) > \sum_{x=1}^{l/2} 2\Delta(2-2x) > 2\Delta(0) \sum_{x=1}^{l/2} (2)^{x-1}$. Hence, even for only $l = 4$, $\sum_{x=1}^{l} \Delta(1-x) > 6\Delta(0)$, and since $\lim_{g\to\infty} \sum_{x=1}^{g} \Delta(x) < 4\Delta(1)$, the person would reject the bet no matter how large g is.

reject a 50-50 prospect over $-\$100$ and $+\$200$, but accept multiple repetitions thereof. While such behavior is inconsistent with EU, it arises naturally under reference dependence. A reference-dependent decision maker would reject the single prospect if

$$\begin{aligned} v(0) &> 0.5 \cdot v(-100) + 0.5 \cdot v(200) \\ \text{or } \lambda &> 2. \end{aligned}$$

This same decision maker would accept two independent repetitions provided

$$\begin{aligned} v(0) &< 0.25 \cdot v(-200) + 0.5 \cdot v(100) + 0.25 \cdot v(400) \\ \text{or } \lambda &< 3. \end{aligned}$$

For $\lambda \in (2, 3)$, the decision maker rejects the single prospect and accepts the combined prospect. The key intuition is that a reference-dependent person is particularly influenced by losses, and as the bet is aggregated, the probability of a loss gets smaller and smaller.

As highlighted by Benartzi and Thaler, this application requires a bracketing assumption that the two independent repetitions are treated as a single prospect. If instead the two independent repetitions are treated as two separate decisions whether to take the single prospect, the person would still reject them. In fact, an experimental literature has directly tested the impact of this type of bracketing. Experimental subjects regularly show a substantially greater willingness to accept Samuelson-style bets when explicitly bracketed together than when explicitly bracketed separately (Gneezy and Potters 1997; Bellemare et al. 2005; Haigh and List 2005).

While rationalizations of these well-known behavioral anomalies are valuable, they do not constitute true tests of reference-dependent preferences because these anomalies were, at least partially, desiderata in the construction of these preferences. We next turn to applications that are focused on novel predictions of the model.

Readers will note that the literature and this chapter focus a great deal of attention on two domains: the endowment effect and daily labor supply. This focus does not reflect a view that these domains are particularly important. Rather, it reflects that these domains permit sharp tests of model predictions, thereby providing a solid foundation and yielding insights that can be applied broadly. Indeed, as we will also see, reference-dependent preferences have been usefully applied in a number of arguably more important domains.

4.1 Endowment Effect

Perhaps the most-prominent early example ascribed to loss aversion is the "endowment effect," the frequent finding in both experimental and survey research that the willingness to pay to obtain a good (WTP) is significantly lower than the willingness to accept to give up that the same good (WTA). The term was coined by Thaler (1980) based on the following anecdote:

> Mr. R bought a case of good wine in the late '50's for about \$5 a bottle. A few years later his wine merchant offered to buy the wine back for \$100 a bottle. He refused, although he has never paid more than \$35 for a bottle of wine.

That Mr. R's WTA (which is at least \$100) should so substantially exceed his WTP (which is at most \$35) strains the standard model. Apart from wealth effects—i.e., that owning the bottle makes him wealthier and thus willing to pay more for objects—the two values should coincide. Thaler (1980) argued that loss aversion provides a natural interpretation for such differences, both because the loss associated with giving up a bottle of wine is larger than the gain associated with obtaining the same bottle of wine, and because the loss associated with spending money to buy a bottle of wine is larger than the gain associated with obtaining money when selling a bottle of wine.

A few years later, Thaler and coauthors pursued experimental demonstrations of the endowment effect (Knetsch (1989), Kahneman et al. (1990), Kahneman et al. (1992)). The body of evidence on endowment effects takes two prominent forms: choice from valuation tasks and choice from exchange tasks.

In valuation tasks, the goal is to elicit a monetary valuation for an item (such tasks are most similar to Thaler's wine-bottle anecdote). The focus has been on three groups of subjects:

- *Buyers* are not endowed with the item, and their WTP is elicited.
- *Sellers* are endowed with the item, and their WTA is elicited.
- *Choosers* are not endowed with the item, and asked whether they would prefer to obtain an amount of money or the item. A chooser's valuation is labelled the *equivalent gain* (EG).

Kahneman et al. (1990) elicited valuations from buyers and sellers for mugs and pens, and found (consistent with Thaler's anecdote) that the WTA was significantly larger than the WTP, typically two or three times larger. To control for wealth effects, they also elicited valuations for choosers—there is no wealth difference between sellers and choosers. They found that the WTA was also significantly larger than the EG, which was only a little larger than the WTP. There were numerous replications and extensions in the years which followed, and this pattern frequently emerges in valuation tasks (Horowitz and McConnell (2002) provide a survey of 50 studies and find a median ratio of mean WTA to mean WTP of 2.6).

In exchange tasks, all subjects face a choice between item A and item B, but subjects differ in their endowment. Group A is endowed with item A and given the opportunity to exchange it for item B. Group B is endowed with item B and given the opportunity to exchange it for item A. Group C is not endowed with either item, and merely chooses between the two items. Note that, in such exchange tasks, there is no difference in wealth across subjects. Knetsch (1989) ran this paradigm with all three groups choosing between a mug and a chocolate bar, and found that group A was most likely to choose item A, group

B was least likely to choose item A, and group C was in the middle. This pattern frequently emerges in exchange tasks.

It is instructive to develop a model of the loss-aversion account of the endowment effect (we expand on this model in Section 6.1 when we discuss the more recent endowment-effect literature in detail). Suppose there are two goods, good M and good P, and consumption is a bundle (x_M, x_P). A person with reference bundle (r_M, r_P) makes choices to maximize utility

$$U((x_M, x_P)|(r_M, r_P)) \equiv v_M x_M + v_P x_P + \mu(v_M x_M - v_M r_M) + \mu(v_P x_P - v_P r_P)$$

where v_M and v_P reflect the intrinsic marginal utility from goods M and P, and μ is a universal gain-loss utility function as in equation 5.

Consider a valuation task for a mug. Let $x_M \in \{0, 1\}$ denote mug consumption, so v_M is the intrinsic value of a mug. Let $x_P \in \mathcal{R}$ denote money spent or received during the experiment, and assume $v_P = 1$ (reflecting quasi-linear utility).[21] The valuations of the three types can be derived as follows:

- Sellers have reference points $r_M = 1$ and $r_P = 0$, and thus sell the mug when

$$\begin{aligned} U((1,0)|(1,0)) &\leq U((0,x_P)|(1,0)) \\ v_M &\leq x_P - \eta\lambda v_M + \eta x_P \qquad \Leftrightarrow \qquad x_P \geq \frac{1+\eta\lambda}{1+\eta} v_M \equiv WTA. \end{aligned}$$

- Buyers have reference points $r_M = 0$ and $r_P = 0$, and thus buy a mug when

$$\begin{aligned} U((0,0)|(0,0)) &\leq U((1,-x_P)|(0,0)) = \\ 0 &\leq v_M - x_P + \eta v_M - \eta\lambda x_P \qquad \Leftrightarrow \qquad x_P \leq \frac{1+\eta}{1+\eta\lambda} v_M \equiv WTP. \end{aligned}$$

[21] Given linear intrinsic utility for money, without loss of generality we can set prior wealth equal to zero.

- Choosers have reference points $r_M = 0$ and $r_P = 0$, and thus choose a mug when

$$\begin{aligned} U((1,0)|(0,0)) &\geq U((0,x_P)|(0,0)) \\ v_M + \eta v_M &\geq x_P + \eta x_P \qquad \Leftrightarrow \qquad x_P \leq v_M \equiv EG. \end{aligned}$$

In this model, loss aversion over mugs implies $WTA > EG$, and loss aversion over money implies $EG > WTP$. Hence, much as Thaler (1980) intuited, there are two reasons why $WTA > WTP$. Moreover, because we have assumed a universal gain-loss utility function, the model predicts $WTA/EG = EG/WTP$. The fact that early experiments typically found $WTA/EG > EG/WTP$ was taken as evidence that loss aversion over a good is significantly larger than loss aversion over money, and indeed the fact that EG is close to WTP was taken as evidence that there might be no loss aversion over money (see discussion in Tversky and Kahneman (1991)).

Next consider an exchange task of a mug vs. a pen. Again let $x_M \in \{0,1\}$ denote mug consumption, but now let $x_P \in \{0,1\}$ denote pen consumption, so v_M and v_P are the intrinsic values of a mug and a pen. The behavior of the different groups can be derived as follows:

- People endowed with a pen have reference points $r_M = 0$ and $r_P = 1$, and thus choose pen when

$$\begin{aligned} U((0,1)|(0,1)) &\geq U((1,0)|(0,1)) \\ v_P &\geq v_M + \eta v_M - \eta\lambda v_P \qquad \Leftrightarrow \qquad v_P \geq \frac{1+\eta}{1+\eta\lambda} v_M. \end{aligned}$$

- People endowed with a mug have reference points $r_M = 1$ and $r_P = 0$, and thus choose pen when

$$\begin{aligned} U((0,1)|(1,0)) &\geq U((1,0)|(1,0)) \\ v_P - \eta\lambda v_M + \eta v_P &\geq v_M \qquad \Leftrightarrow \qquad v_P \geq \frac{1+\eta\lambda}{1+\eta} v_M. \end{aligned}$$

- People not endowed with either item have reference points $r_M = 0$ and $r_P = 0$, and thus choose pen when

$$\begin{aligned} U((0,1)|(0,0)) &\geq U((1,0)|(0,0)) \\ v_P + \eta v_P &\geq v_M + \eta v_M \qquad \Leftrightarrow \qquad v_P \geq v_M. \end{aligned}$$

Because $\lambda > 1$ implies $\frac{1+\eta}{1+\eta\lambda} v_M < v_M < \frac{1+\eta\lambda}{1+\eta} v_M$, loss aversion implies that people endowed with a pen have the lowest cutoff and thus are most likely to choose pens, and people endowed with a mug have the highest cutoff and thus are least likely to choose pens. Hence, this model cleanly illustrates the loss-aversion account for endowment effects in exchange tasks.

While valuation tasks and exchange tasks are often viewed as distinct, they are both ways of asking people to choose between two goods while varying their endowments. Indeed, the model is much the same whether the second good is money or pens, as is clear from the inequalities above for the two cases. In practice, the main difference is that money is more divisible, and thus valuation tasks reveal finer grained deviations from the standard model with greater statistical precision.

Though WTP-WTA disparities in valuation experiments and exchange asymmetries in exchange experiments are widely documented, an important body of research over the past 15-20 years has questioned the usual loss-aversion interpretation reflected above. In Section 6.1, we describe this debate in detail.

4.2 Labor Supply

A second prominent early example revolved around workers who choose their labor supply on a daily basis—such as taxi drivers—and the possibility that they might be using a daily income target. More precisely, the hypothesis is that workers are bracketing their labor-supply decisions at the daily level, and on each day they experience reference-dependent gain-loss utility depending on how that day's earnings compare to some reference point (target earnings).

We begin with a simple model of this domain (framed in a way that can also be applied to other domains later). Consider a worker choosing an effort level e that generates cost $c(e)$ that is increasing and (weakly) convex. Effort level e also generates an observable outcome, $x(e)$, that is increasing and (weakly) concave. Letting $u(x(e))$ denote the intrinsic utility associated with $x(e)$, a standard agent would choose e to maximize $u(x(e)) - c(e)$. In what follows, we assume for simplicity that intrinsic utility is linear.[22] Hence, a standard agent would choose e to maximize $x(e) - c(e)$, yielding e^* such that

$$MB(e^*) \equiv x'(e^*) = c'(e^*) \equiv MC(e^*).$$

In the domain of daily labor supply, e can be interpreted as the number of hours worked. Given an hourly wage w, $x(e) = we$. Standard agents, then, merely choose effort such that $c'(e^*) = w$, and, given that c is convex, a drop in the wage leads to a drop in labor supply.[23]

Consider instead an agent who experiences gain-loss utility based on how $x(e)$ compares to some exogenous reference point r_x. Applying μ as in equation 5, this agent chooses e to maximize

$$U(e) \equiv x(e) + \mu(x(e) - r_x) - c(e).$$

[22]Models of this form would typically assume either that x is linear and c is convex or that c is linear and x is concave, depending on the interpretations of x and c.

[23]The assumption that intrinsic utility is linear corresponds to an assumption that any standard income effect associated with earnings on a single day are negligible.

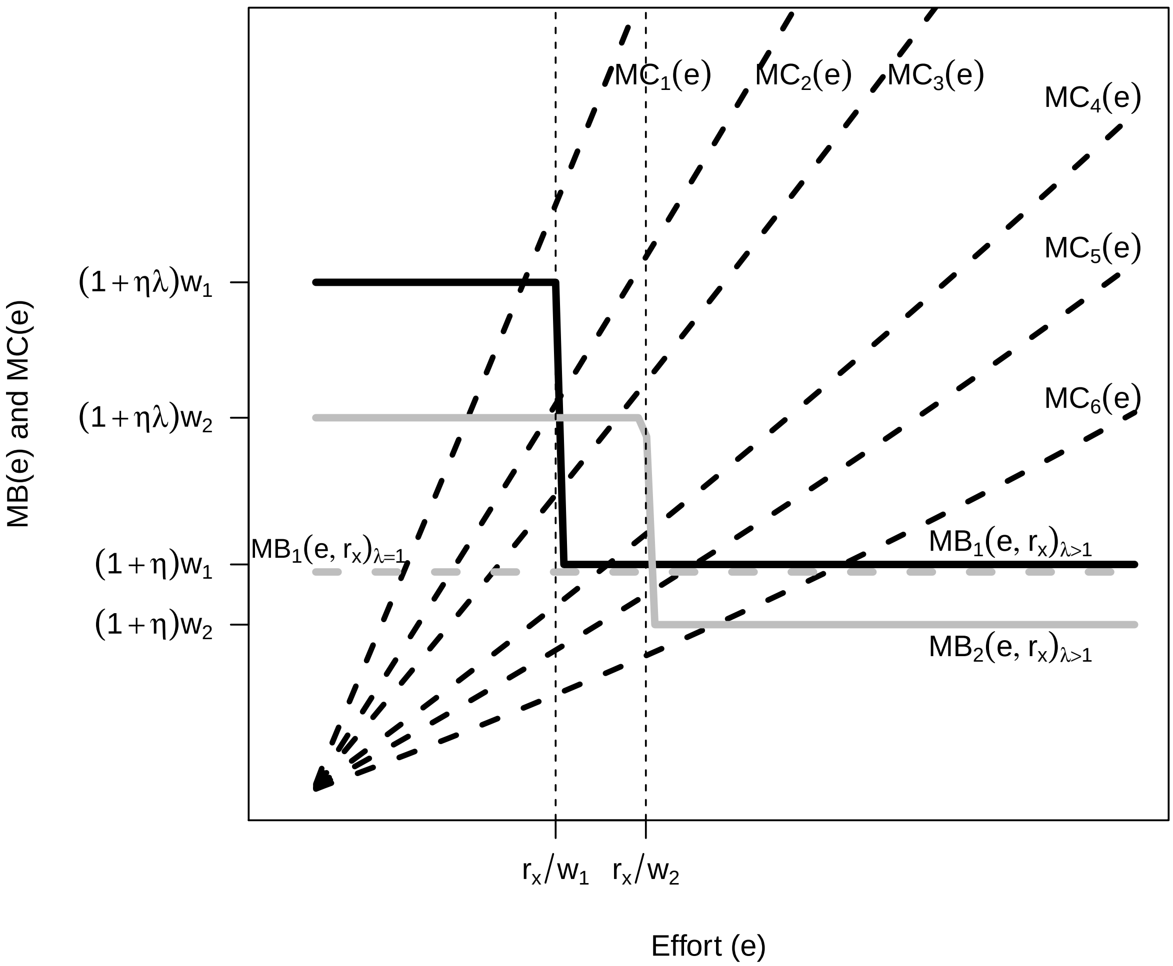
Note that, for simplicity, we assume that the agent experiences no gains and losses associated with the cost of effort, $c(e)$. As such, marginal costs remain $MC(e) = c'(e)$, while marginal benefits become

$$MB(e) = \begin{cases} (1+\eta)x'(e) & \text{if } x(e) > r_x \\ (1+\eta\lambda)x'(e) & \text{if } x(e) < r_x. \end{cases}$$

Hence, marginal benefits drop discontinuously at r_x.

Figure 3 depicts marginal benefits and marginal costs for a reference-dependent agent. Consider first the dark boldfaced marginal benefit curve labeled $MB_1(e, r_x)_{\lambda>1}$ which is associated with benefit function $x_1(e) = w_1 e$. At low effort levels, the person has earned less than her target, and thus has high marginal benefit $(1+\eta\lambda)w_1$. At high effort levels, the person has earned more than her target, and thus has low marginal benefit $(1+\eta)w_1$. At effort level e such that $x_1(e) = r_x$, which in the labor supply domain is $e_1^r = r_x/w_1$, the marginal benefit discontinuously drops. Figure 3 presents marginal cost curves for six different people (who have the same income target). We see that person 1 chooses not to make it to the target, persons 2 and 3 choose to work until they exactly hit the target, and persons 4, 5, and 6 work so as to make more than the target.

The key comparative static in the early labor-supply literature is how do people react to a change in the daily wage. In Figure 3, the light boldfaced marginal benefit curve labeled $MB_2(e, r_x)_{\lambda>1}$ reflects what the benefit curve would look like for daily wage $w_2 < w_1$. This drop in the wage increases the location of the discontinuity to $e_2^r = r_x/w_2$, and it also decreases the marginal benefit on either side of the discontinuity. Whereas standard agents would react to a wage decrease by simply reducing their labor supply, reference-dependent workers have a more mixed response. For some marginal cost curves, workers will show the expected response of reducing their labor supply (persons 1, 2, 5, and 6). For other marginal cost curves, however, workers react to a lower wage by increasing labor supply (persons 3 and 4). The intuition is simple: for $e \in (e_1^r, e_2^r)$, the lower wage actually increases the marginal benefit of effort because this range of effort is now in the loss domain.



Note: Figure presents marginal benefit and cost of effort. Solid black line $MB_1(e, r_x)_{\lambda>1}$ reflects marginal benefit for a loss-averse agent who faces wage w_1. Solid gray line $MB_2(e, r_x)_{\lambda>1}$ reflects marginal benefit for a loss-averse agent who faces wage $w_2 < w_1$. Dashed gray line $MB_1(e, r_x)_{\lambda=1}$ reflects marginal benefit for an agent who has no loss aversion and who faces wage w_1. $MC_1(e)$, $MC_2(e)$, ...,and $MC_6(e)$ reflect marginal costs for agents 1, 2, ..., and 6.

Figure 3: Reference Dependence and Effort Provision

The early literature on daily labor supply focused on this anomalous possibility that loss aversion could possibly lead to negative wage elasticities. The seminal paper by Camerer et al. (1997) studies behavior of New York City cab drivers. They first analyze hourly wages (hourly earnings) for drivers, and conclude that hourly wages are highly correlated within a day and relatively uncorrelated across days. Based on this preliminary finding, they take their unit of observation to be a day, and estimate a standard daily wage equation of the form

$$\ln H_t = \gamma W_t + \beta X_t + \varepsilon_t$$

where H_t is hours worked on day t, and W_t is the daily wage on day t (derived as driver-specific daily earnings divided by the hours worked). Using three different datasets and multiple specifications, Camerer et al. (1997) consistently find evidence of negative wage elasticities—inconsistent with a standard model but consistent with daily income targeting.

The specification of Camerer et al. (1997) suffers from a potential division bias because wages are constructed by dividing daily earnings by hours. Though they attempt to control for this bias with an instrumental-variables approach, the conclusions may be challenged by the lack of exogenous variation in wages. In a prominent follow-up study, Fehr and Goette (2007) implement exogenous (experimental) variation in wages for bicycle messengers. They find that higher wages lead bicycle messengers to choose to work more shifts, but also to earn less per shift (on net, total earnings increased). Fehr and Goette also conduct experiments to elicit loss aversion for a sample of the bicycle messengers, and find that the negative treatment effect on shift earnings is predicted by their experimental measures of loss aversion. Taken together, these results suggest a model in which loss aversion does not have much impact on the extensive margin of whether to work on a given day, but does impact the intensive margin of how long or how hard to work once one has shown up.[24]

The initial work on taxi drivers spurred a number of additional studies in that domain, along with a substantial discussion between labor and behavioral economists. In Section 6.2, we revisit the topic of labor supply and discuss this debate in detail.

[24]The extensive margin finding is echoed by Oettinger (1999) who analyzes stadium vendors over the course of a baseball season. Consistent with the findings of Fehr and Goette (2007) and standard labor-supply models, he shows that vendors are more likely to work when the predictable earnings are larger.

4.3 Other Forms of Effort Provision

The model of effort provision in Section 4.2 can also be applied to other domains. Here we describe three applications that study effort to reduce one's income taxes due, to run faster in a marathon, and to make a putt in professional golf.

Figure 3 highlights two additional implications of reference dependence in the domain of effort provision. First, we should observe bunching at the reference point. As a benchmark, consider the light dashed marginal benefit curve labeled $MB_1(e, r_x)_{\lambda=1}$, which reflects what the marginal benefit curve would look like for a standard agent. For standard agents, if people differ in only their marginal cost curves, each type would choose a different effort level, and we should thus observe a population distribution of outcomes that roughly corresponds to the population distribution of marginal cost curves. In contrast, for reference-dependent agents who face the dark boldfaced $MB_1(e, r_x)_{\lambda>1}$, we should observe a bunching of agents at e_1^r (e.g., persons 2 and 3 in the figure). Second, we should observe differential behavior on either side of the reference point, and in particular we should see people exert additional effort when they are in the loss domain. For instance, if the reference point were to move far to the left so as to put person 1 well above the reference point, there would be a significant reduction in her effort.

Rees-Jones (2018) identifies both implications—bunching and differential behavior—in the domain of income-tax preparation. He conceptualizes people as putting together initial tax information to formulate an initial balance due, and then deciding how much extra effort to exert to identify additional deductions and credits. He posits that, when people make these decisions, it is natural to use a reference point of a zero balance due. If so, then loss aversion implies that we should observe bunching of final balance due at zero, and moreover we should observe more additional deductions and credits for people whose final balance due is positive (in the loss domain) than we do for people whose final balance due is negative (in the gain domain). Using IRS data from 1979-1990, Rees-Jones (2018) estimates a model using both of these implications. The raw data show clear signs of bunching in the neighborhood of zero balance due, and his estimates suggest that, on average, people pursue an additional $34 of tax reduction when they owe money relative to when they are receiving a refund.

Allen et al. (2017) investigate bunching behavior in the domain of marathon running. They posit that marathon runners have reference points of round-number finishing times—e.g., finishing in better than 4 hours (4:00:00), or better than 3.5 hours (3:30:00). Using data covering 9,789,093 finishing times from 6,888 marathons run over the period 1970-2013, they find clear evidence of bunching at finishing times just better than various salient round numbers. Indeed, a simple glance at their raw data (in their Figure 2) reveals striking evidence of bunching. The most extreme bunching occurs in the minutes just before 4:00:00—300,324 finishing times fall in the interval 3:57:00-3:59:59, whereas only 212,477 finishing times fall in the interval 4:00:00-4:02:59. Their analysis finds statistically significant bunching around most other 10-minute marks as well.

Pope and Schweitzer (2011) investigate differential behavior on either side of the reference point in the domain of professional (PGA Tour) golf. In most professional golf tournaments, players play 72 holes, and the order of finish—and hence earnings—is entirely determined by the total number of shots taken over those 72 holes (with the lowest total being the best). In addition, each individual hole has a suggested score ("par"). For standard agents, because this suggested score is completely irrelevant to the order of finish and earnings in the tournament, it should be irrelevant to their effort. Pope and Schweitzer posit that professional golfers experience gain-loss utility on each hole, with par serving as a natural reference point. Applying the model above, we would predict different behavior on either side of the reference point—e.g., one should expect more effort and thus a higher proportion of putts made when one is putting for par (when trying to avoid a loss) than when one is putting for birdie (when one is trying to achieve a gain). Pope and Schweitzer analyze data from 239 PGA Tour tournaments completed between 2004 and 2009, focusing on all golfers who attempted at least 1000 putts—yielding data on 421 golfers and over 2.5 million putts. The data also contain the exact locations of the ball and the hole. Controlling for the distance of the putt, they indeed find that on average golfers are about 2 percentage points more likely to make a par putt than they are to make a birdie putt.

4.4 Finance

A robust literature has applied prospect theory to finance, and this literature has in fact highlighted the implications of all three of the main components of prospect theory: loss aversion (as highlighted in the non-finance applications above), diminishing sensitivity, and probability weighting. For a detailed coverage, see Chapter X in this volume. Here we focus on three examples that apply gain-loss utility.

One prominent example is the "equity premium puzzle," the well known finding that the observed premium that equity pays relative to bonds is too large to be consistent with EU and plausible levels of risk aversion (Mehra and Prescott (1985)). Benartzi and Thaler (1995) suggest reference-dependent loss aversion as a potential explanation (although there exist other explanations as well). The basic mechanism proposed by Benartzi and Thaler (1995) is that, with some regularity (perhaps each year), people look at their financial-portfolio statement and experience gain utility if their portfolio has gone up in value since the prior statement, and loss utility if it has gone down. If people's expectations of such feelings impact their portfolio allocation, and if losses loom larger than gains, then this utility associated with paper gains and losses will lead to increased risk-aversion in portfolio choice. Barberis et al. (2001) further develop this hypothesis in a formal asset-pricing model.

A second prominent example is the "disposition effect," the tendency of retail investors to be more prone to sell their winners than their losers. Odean (1998) documents this tendency in a sample of retail investors, and further demonstrates that their behavior is inconsistent with several standard explanations. This behavior is often interpreted as being due to a "realization utility" wherein traders experience gain-loss utility associated with whether realized capital gains are positive or negative (Barberis and Xiong (2012)).

Realization utility—even if it is linear and thus doesn't involve loss aversion—combined with discounting can generate a disposition effect. For instance, suppose a person is deciding whether to sell a winner vs. a loser today (e.g., because she needs liquidity). If she sells the winner, she experiences a gain now instead of in the future; if she sells the loser, she experiences a loss now instead of in the future. Without discounting, there is no reason to prefer one vs. the other. With discounting, however, this realization utility creates a preference to sell the winner, both to accelerate the gain and to delay the loss. Again, this effect does not require loss aversion. That said, loss aversion can magnify the effect because the negative feeling of loss would then be even more aversive and thus there would be a stronger incentive to delay it.

Even if there is no discounting, realization utility can also generate a disposition effect when combined with diminishing sensitivity. Again, consider a person who is deciding whether to sell a winner vs. a loser today. Ceteris paribus, the former involves comparing a sure gain today versus an uncertain gain in the future, whereas the latter involves comparing a sure loss today versus an uncertain loss in the future. Because diminishing sensitivity creates risk aversion for gains and risk seeking for losses, it encourages the person to sell the winner and hold the loser.

A third example is evidence from Genesove and Mayer (2001) that can be interpreted as reflecting a disposition effect in housing markets. Housing markets are much more complicated than stock markets. In the latter, a seller faces a decision whether to sell at a known market price, and the transaction can take place relatively quickly. In housing markets, a seller must first set a list price and then wait for offers, where the choice of list price is likely to impact the frequency with which offers come in as well as the nature of those offers. Once an offer comes in, frequently there is a period a negotiation, and eventually there is a sale at a transaction price that need not be the same as either the original list price or the initial offer price.

Within this domain, realization utility combined with either discounting or diminishing sensitivity will make sellers subject to a loss less willing to sell than sellers subject to a gain. The primary prediction of this mechanism is that sellers will set higher list prices when they are subject to a loss than when they are subject to a gain. A secondary prediction is that sellers will have larger thresholds when deciding whether to accept offers.[25]

Genesove and Mayer (2001) use data from the Boston condominium market in the 1990's. Because this market experienced rising prices over 1982-1989, falling prices over 1989-1992, and then rising prices again over 1992-1997, they observe sellers selling similar houses under similar market conditions but who purchased their houses at different prices. Hence, they can investigate the extent to which seller behavior depends on whether they are subject to a gain or a loss.[26] Because list prices are observable in the data, it is possible to conduct a direct test of the primary prediction above. Genesove and Mayer indeed find that, relative to sellers who are subject to an expected nominal gain, sellers who are subject to an expected nominal loss set higher list prices. Additionally, higher expected losses are associated with higher list prices. For the secondary prediction, a more indirect test is required, specifically testing whether sellers who are subject to a loss also experience larger final transaction prices and longer time to sale. Genesove and Mayer also find evidence of both.[27]

4.5 Coding and Bracketing in Applications

As we leave these applications, it is important to reiterate the importance of coding and bracketing.

On the former, note that these applications use a wide range of reference points: one's

[25]A possible third prediction is bunching of transaction prices on the prior purchase price. However, whether a model of reference-dependent realization utility would generate such a prediction would depend upon how one models the details of how offers are generated and how negotiations evolve.

[26]A major part of the analysis revolves around defining what it means for a seller to be subject to a gain or a loss. To determine whether a seller is likely to be experiencing a gain or a loss relative to the purchase price, they estimate a predicted sales price and compare it to the purchase price. The predicted sales price is generated from hedonic price regressions that control for various features of the house, although they must also deal with the bias created by unobserved quality. See the paper for details.

[27]Although Genesove and Mayer (2001) argue that their empirical findings are due to loss aversion, they might better be attributed to realization utility combined with either discounting or diminishing sensitivity (although loss aversion might enhance the magnitude).

prior wealth, one's initial endowment of a mug or pen, one's daily target for labor earnings, the salient benchmark of a zero balance due on taxes, salient round numbers for marathon finishing times, a suggested score (par) on a golf hole, a prior value of one's portfolio, and the price at which an asset is purchased. The choice of reference point represents a powerful degree of freedom when one takes reference-dependent preferences to applications. Recently, interest has turned to disciplining the model with structured mechanisms for the formulation of reference dependence (as we'll see in the next section).

These applications similarly make a range of bracketing assumptions. For instance, individuals are assumed to treat an aggregate bet as one grand bet, to treat each day's labor-market decision as independent from other days, and to treat each golf hole separately. As such, tests of reference dependence in these settings should be properly recognized as tests of joint hypotheses related to loss aversion and correct formulation of the bracket.

5 Expectations-Based Models

5.1 Overview and Some History

Reference dependence around an appropriately defined exogenous reference point has proven a useful model for explaining a variety of behavior. From risky choice, to financial behavior, to labor supply, the model provides a coherent account of decision making. At various times, however, researchers have taken a different approach to gain-loss utility wherein gains and losses are defined relative to an expectations-based referent.[28]

The first work in this direction was carried out by Bell (1985) and Loomes and Sugden (1986) under the label "disappointment aversion" (DA). They posit that as one considers a lottery, one forms a prior expectation of its outcome, and then, when the uncertainty is

[28] We make use of the terms "reference point" and "referent" in this chapter to describe the locations around which gains and losses are defined. The term "referent" is less frequently used in the literature, but is often helpful terminology when there is a distribution of possible reference points to emphasize that the referent is not a single point.

resolved, one experiences elation or disappointment depending on how the realized outcome compares to that prior expectation. While this intuition is clearly quite similar to the loss-aversion intuition from prospect theory, this work focused more on anomalies à la Allais and subsequently did not receive much attention in the early behavioral economics literature.[29]

Roughly two decades later, Koszegi and Rabin (2006, 2007) (KR) proposed a model with an expectations-based referent that was motivated by a desire to impose some discipline on models of reference dependence. Koszegi and Rabin realized that if one uses an expectations-based referent combined with an assumption that expectations must be rational, the model's flexibility is dramatically reduced. It turns out that the KR model shares some features with DA models.

When working with an expectations-based referent, there are two key modeling questions that one must address. First, given that expectations will likely involve uncertainty, how should one incorporate this uncertainty into the referent? Second, what determines expectations? We examine these two questions in the next two subsections, highlighting key differences between the DA and KR formulations.

5.2 Formulation of the Expectations-Based Referent

The premise of an expectations-based approach is that the referent depends on one's expectations over outcomes. Because expectations will typically involve uncertainty, we need to develop a model in which the referent is a lottery rather than an outcome. We again consider prospects of the form $L \equiv (x_1, p_1; x_2, p_2; ...; x_N, p_N)$ over N potential outcomes with $\sum_{n=1}^{N} p_n = 1$. However, we now assume a person has a reference lottery of the form $R = (r_1, q_1; r_2, q_2; ...; r_M, q_M)$ over M potential reference outcomes with $\sum_{m=1}^{M} q_m = 1$. Given

[29] Rabin (1998) and Camerer and Loewenstein (2003), two prominent early overviews of behavioral economics, do not discuss or cite models of disappointment aversion.

a reference lottery R, the person evaluates lottery L according to

$$U(L|R) \equiv \sum_{n=1}^{N} p_n \left[u(x_n) + v(x_n|R)\right]$$

where $u(x_n)$ denotes the intrinsic utility from outcome x_n, and $v(x_n|R)$ denotes the gain-loss utility associated with getting outcome x_n given a reference lottery R.[30]

The literature has pursued two approaches to model $v(x_n|R)$. Models of disappointment aversion (Bell 1985; Loomes and Sugden 1986) assume that x_n is compared to a single summary statistic from the reference lottery R. Specifically, such models assume:

$$\text{DA Approach: } v(x_n|R) \equiv \mu\left(u(x_n) - \left(\sum_{m=1}^{M} q_m u(x_m)\right)\right). \tag{6}$$

If $u(x) = x$, then x_n is compared to the mean of R. More generally, the intrinsic utility from x_n is compared to the expected intrinsic utility from the reference lottery.[31]

KR propose an alternative approach in which x_n is compared to every outcome that might have occurred in the reference lottery R. Specifically, they assume:

$$\text{KR Approach: } v(x_n|R) \equiv \sum_{m=1}^{M} q_m \mu\left(u(x_n) - u(x_m)\right). \tag{7}$$

Hence, if x_n is larger than some x_m in R but smaller than other x_m in R, it will be experienced as a gain relative to the former and a loss relative to the latter. With this formulation, KR assume that, when the person aggregates these feelings of gains and losses across all comparisons, the person uses weights that correspond to the probabilities in R.

[30] Our treatment here develops gain-loss utility with an expectations-based referent for the case where each x_n is a scalar—that is, the case where there is one dimension of consumption. If there are multiple dimensions of consumption, it is straightforward to extend the analysis here if one further assumes additive separability across dimensions (as in Section 3.3 and 5.4).

[31] In his main example, Bell (1985) assumes μ takes a form analogous to the two-part linear value function discussed in Section 3.2 with losses (disappointment) looming larger than gains (elation). In contrast, Loomes and Sugden (1986) assume that μ is symmetric around zero (no loss aversion), convex over gains, and concave over losses—i.e., a form that is quite different from the value function proposed by Kahneman and Tversky (1979).

These two approaches reflect different psychologies for how people experience sensations of gain and loss. Moreover, they also generate different predictions. Perhaps most notably, consider what happens when there is increased risk in the reference lottery that does not impact the expected utility of the reference lottery. Under the DA approach, such increased risk in the reference lottery has no impact on behavior. In contrast, under the KR approach, such increased risk in the reference lottery actually makes the person more willing to bear risk—a kind of endowment effect for risk.[32]

Example 1: Consider a choice between $L_1 \equiv (z+\gamma, \frac{1}{2}; -z, \frac{1}{2})$ vs. $L_2 \equiv (0,1)$, where $z > 0$ and $\gamma \geq 0$. The reference lottery is $R \equiv (\phi, \frac{1}{2}; -\phi, \frac{1}{2})$ where $\phi \in [0, z]$.

- Under the DA approach reflected in equation 6:

$$\begin{aligned} U(L_1|R) &= \frac{1}{2}(1+\eta)\gamma - \frac{1}{2}\eta(\lambda-1)z \\ U(L_2|R) &= 0 \\ &\iff \text{choose } L_1 \text{ if } \gamma > \frac{\eta(\lambda-1)z}{1+\eta}. \end{aligned}$$

- Under the KR approach reflected in equation 7:

$$\begin{aligned} U(L_1|R) &= \frac{1}{2}(1+\eta)\gamma - \frac{1}{2}\eta(\lambda-1)z \\ U(L_2|R) &= -\frac{1}{2}\eta(\lambda-1)\phi \\ &\iff \text{choose } L_1 \text{ if } \gamma > \frac{\eta(\lambda-1)(z-\phi)}{1+\eta}. \end{aligned}$$

---

[32] While, for completeness, the presentation above permits $u(x)$ to be non-linear, Example 1 and all subsequent examples assume $u(x) = x$.

In Example 1, the person must choose whether to accept or reject a 50-50 gamble with a (weakly) positive expected value. Under the DA approach, gain-loss utility around $ER = 0$ creates risk aversion, and thus the expected value of the gamble (as captured by γ) must be sufficiently large for the person to take the gamble. The amount of risk in R (as captured by ϕ) does not impact behavior because it does not change ER. Under the KR approach, gain-loss utility again creates risk aversion. Notice, however, that the more risk there is in R (the larger is ϕ), the smaller need be the expected value of the gamble. Indeed, for $\phi = z$, the person becomes risk-neutral and accepts L_1 for any $\gamma > 0$. Intuitively, under the disappointment-aversion approach, L_2 has a major advantage over L_1 because it involves no gain-loss utility. Under the KR approach, in contrast, L_2 loses this advantage, because even the certain outcome generates gain-loss utility relative to the outcomes in R that could have occurred.

Sprenger (2015) provides a test of DA preferences vs. KR preferences by investigating whether people in fact exhibit an endowment effect for risk. The design is based on an assumed framing effect: that when subjects are presented with a series of binary choices in which one option is fixed and the other varies in a way clearly designed to identify an indifference point, the fixed option will be used as a referent. Hence, one group of subjects is presented tasks with a fixed certain amount and asked to choose an equivalent gamble, while a second group is presented tasks with a fixed gamble and asked to choose an equivalent certain amount. The results support KR preferences: subjects given a fixed certain amount are around 75% more likely to exhibit risk aversion and 30% less likely to exhibit risk neutrality than subjects given a fixed gamble.[33]

[33] Sprenger (2015) also documents the same phenomena within-subjects. Fixing a choice between \$10 for sure and a 50-50 gamble over \$0 and \$30, if the \$10 is fixed 54% of subjects prefer the \$10 while if the 50-50 gamble is fixed only 32% prefer the \$10.

5.3 What Determines Expectations?

Equations 6 and 7 provide models of gain-loss utility for any expectations-based referent (i.e., for any reference lottery R). One must next make assumptions regarding what determines expectations. Two approaches appear in the literature. First, we might take expectations to be exogenous (to the current choice), much as is done in the applications in Section 4. Second, we might take expectations to be endogenous to the current choice. One such approach might permit a person to "choose" the expectations that she'd like as part of the choice problem. This approach, however, is not entirely satisfactory as it seems unlikely a person could really come to hold expectations that differ dramatically from what is actually likely to occur. Moreover, permitting "choice" in one's beliefs introduces another degree of freedom into the model. Hence, as a disciplining assumption, much of the literature has assumed that expectations must be "rational" in the sense that they must be consistent with the person's own behavior.

5.3.1 Exogenous Expectations and Surprise Choice

In Example 1, the referent, R, was considered exogenous to choice. Applying expectations-based models in such a way is effectively no different from applying standard models of reference dependence. In their work, Koszegi and Rabin discuss how this approach might be appropriate in surprise choice situations—that is, when expectations were likely determined by other experiences completely unrelated to a choice that has suddenly been thrust upon the person. In such settings, the model is applied exactly as in the examples above. As we'll see, some empirical applications pursue this approach, taking advantage of natural arguments for what expectations might be.

5.3.2 Endogenous Expectations Induced By Choice

A first rational-expectations approach assumes that when a person chooses a particular option, that option fully determines the referent around which she experiences gain-loss utility when the final outcome is realized. In other words (and using the notation from Section 5.2), if the person chooses lottery L, the reference lottery becomes $R = L$, and thus her utility would be $U(L|L)$. The person would then choose the lottery that maximizes $U(L|L)$.

This assumption was originally incorporated into the models of disappointment aversion (DA) of Bell (1985) and Loomes and Sugden (1986). The motivation there is much as in the paragraph above: when considering a lottery L, the person forms a "prior expectation" about that lottery, and then once the uncertainty is resolved, the person experiences elation or disappointment depending on how the realized outcome compares to that "prior expectation".[34] More recently, Koszegi and Rabin (2007) incorporate this assumption into their solution concept "choice-acclimating personal equilibrium" (CPE). Their motivation is ever-so-slightly different, and goes as follows: if person commits to a lottery L well in advance of the resolution of uncertainty, then by the time the uncertainty is resolved the person will have come to expect the lottery L, and thus it becomes the referent around which gains and losses are defined.

Hence, DA and CPE both assume a person chooses the lottery that maximizes $U(L|L)$, and differ only in their assumption about how a person compares a realized outcome to the reference lottery. DA defines gain-loss utility according to equation 6, while CPE defines gain-loss utility according to equation 7. It follows that under these two models a person

[34] Gul (1991) provides an alternative formulation of disappointment aversion in which the referent is the certainty equivalent of the gamble, including disappointment when formulating that certainty equivalent. Bell (1985) and Loomes and Sugden (1986) effectively assume the referent is the certainty equivalent using only intrinsic utility when formulating that certainty equivalent.

evaluates lotteries according to:

$$\text{DA:} \qquad U(L|L) = \sum_{n=1}^{N} p_n \left[u(x_n) + \mu \left(u(x_n) - \sum_{m=1}^{N} p_m u(x_m) \right) \right]$$

$$\text{CPE:} \qquad U(L|L) = \sum_{n=1}^{N} p_n \left[u(x_n) + \sum_{m=1}^{N} p_m \mu(u(x_n) - u(x_m)) \right].$$

For the CPE equation (and applying equation 5), if we order outcomes such that $x_1 \leq x_2 \leq ... \leq x_N$, one can derive that the CPE equation becomes

$$\text{CPE:} \quad U(L|L) = \sum_{n=1}^{N} p_n u(x_n) - \sum_{n=1}^{N-1} \left[\sum_{m=n+1}^{N} p_n p_m \eta(\lambda - 1)(u(x_m) - u(x_n)) \right]. \qquad (8)$$

From this equation, one can see that η and λ only matter through the product $\eta(\lambda - 1) \equiv \Lambda$. Hence, many applications of CPE in fact use the parameter Λ instead of η and λ, and we often use Λ below.

For certain outcomes or for binary lotteries, DA and CPE yield identical equations for $U(L|L)$.[35] For binary gambles, the reason is that the higher outcome will always be perceived as a gain and the lower outcome as a loss regardless of whether the referent is the other outcome or the gamble's certainty equivalent. Hence, with a two-part-linear gain-loss utility function, if the choice set contains only certain outcomes and binary lotteries, the two models make identical predictions.

Panel A of Figure 4 depicts indifference curves for these models in the domain of binary lotteries. For comparison, Panel A also presents indifference curves for an exogenous certain reference point, r, while applying the same $\mu(\cdot)$. Just like for a certain reference point, DA/CPE generates kinked utility around certainty. Moreover, DA/CPE yields a larger kink than a certain reference point, and thus generates a stronger preference for certain outcomes.

[35] For either model, one can derive that, for any binary lottery $L = (x_1, p_1; x_2, 1-p_1)$ with $x_1 > x_2$, utility is

$$U(L|L) = p_1 u(x_1) + (1 - p_1)u(x_2) - p_1(1 - p_1)\Lambda[u(x_1) - u(x_2)].$$

Intuitively, relative to having a certain outcome, when one chooses a lottery L with positive expected value, not only does one now generate gain-loss utility (as would happen with a certain reference point), but this gain-loss utility is magnified by the fact that choosing L also means creating a disadvantageous reference lottery (e.g., under DA, a reference lottery with a higher expected value against which gains and losses are defined). Hence, DA/CPE is consistent with small- and moderate-stakes risk aversion (as discussed in Sections 2 and 4) with an even smaller extent of loss aversion.

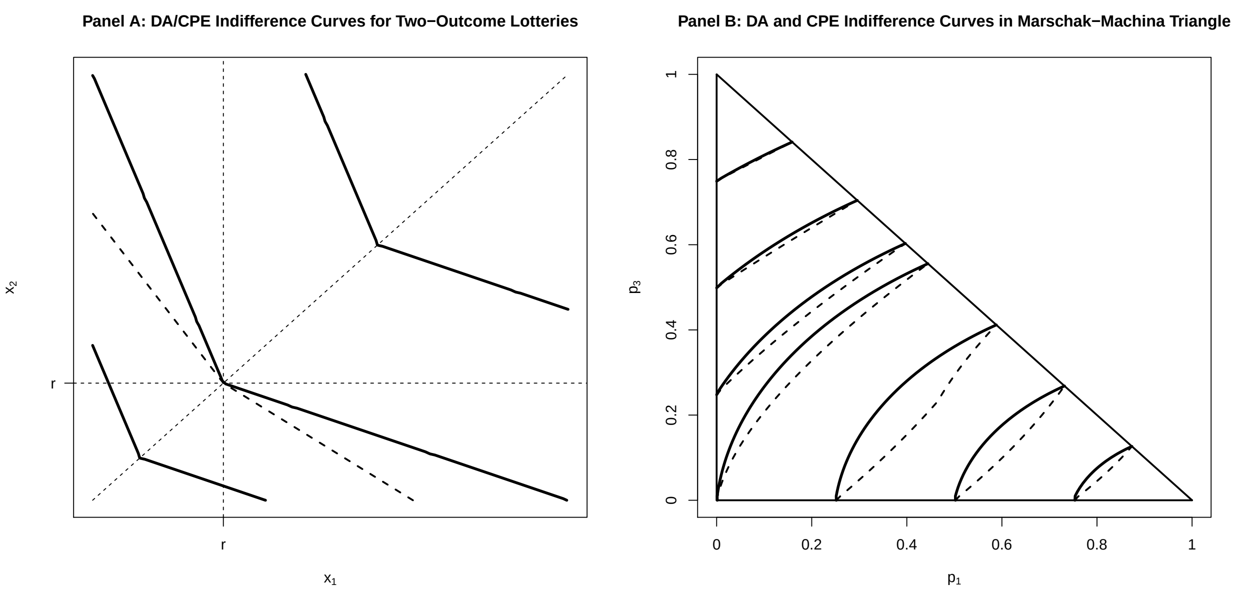
It is also worth noting that both DA and CPE can generate violations of first-order stochastic dominance.[36] To illustrate, consider a choice between $L_1 \equiv (\$10, 1)$ and $L_2 \equiv (\$10, .9; \$20, .1)$. It is straightforward to derive that $U(L_1|L_1) = 10$ while $U(L_2|L_2) = 11 - .9\Lambda$, and thus the person would choose the dominated lottery L_1 if $\Lambda > 1.1$—e.g., if $\eta = 1$ and $\lambda = 2.5$. (In Panel A of Figure 4, this possibility implies that it is possible for indifference curves to be positively sloped.) Though such violations of first-order stochastic dominance are viewed as a normatively unappealing prediction for a model of decision making (for relevant discussion, see Quiggin 1982), several leading empirical examples of dominance violations do exist (see, e.g., Gneezy et al. 2006).

For lotteries with more than two outcomes, DA and CPE differ. Panel B of Figure 4 starts to reveal the nature of this difference by presenting indifference curves in the Marschak-Machina triangle for three-outcome prospects $L = (x_1, p_1; x_2, 1 - p_1 - p_3; x_3, p_3)$ with $x_1 < x_2 < x_3$. Panel B reveals that, even though these models do not assume any probability weighting, in both models $U(L|L)$ will be nonlinear in probabilities.[37] The source of this nonlinearity is simple: probabilities are present both in taking the expectation over utility outcomes and in the referent itself. These nonlinearities reflect important implications. On one hand, they reflect that these expectations-based models are consistent with some prominent EU anomalies (that are inconsistent with status quo reference dependence). For instance, these models can accommodate Allais' two paradoxes. On the other hand, these

[36] The formulation in Gul (1991) does not have this property.

[37] The formulation in Gul (1991) does not have this property.

nonlinearities reflect that these models can violate certain "normative" properties that we might expect people to satisfy. We have already discussed possible violations of first-order stochastic dominance. Another important example is violations of betweenness, wherein a convex combination of two indifferent gambles should be indifferent to both. CPE decision makers will generally be averse to such convex combinations as they induce greater potential for losses.[38]



Note: Panel A presents DA/CPE indifference curves for prospects of the form $L = (x_1, p_1; x_2, 1 - p_1)$, for which the two models make identical predictions. Dashed line reflects indifference curve through (r, r) for an agent with reference-dependent preferences and exogenous reference point r. Panel B presents DA and CPE indifference curves in the Marschak-Machina triangle for prospects of the form $L = (x_1, p_1; x_2, 1 - p_1 - p_3; x_3, p_3)$ with $x_1 < x_2 < x_3$, for which the two models make different predictions. Solid lines correspond to CPE indifference curves; dashed line corresponds to DA indifference curves.

Figure 4: DA and CPE Indifference Curves

[38]For DA decisionmakers, the direction of violations of betweenness will depend on the EU certainty equivalents of the indifferent gambles and that of the convex combination, making general conclusions more challenging to draw. Under KR, the gain-loss comparisons are fixed, only their relative probabilities change, yielding more clear predictions.

5.3.3 Endogenous Expectations Induced by Planned Choice

Koszegi and Rabin (2007) motivate CPE as relevant for situations in which a person commits to a choice well in advance of the resolution of uncertainty. In their original formulation, Koszegi and Rabin (2006), they propose a different approach based on the following type of situation: a person is confronted with a choice situation and given some time to formulate a plan for which option they will choose; however, the person does not actually make (commit to) that choice until shortly before the uncertainty is resolved. In such situations, it seems reasonable to assume that, by the time one needs to make a choice, the person's *planned choice* becomes the referent.

Imposing rational expectations in this environment requires an equilibrium assumption. Specifically, rational expectations require that planned choice be identical to actual choice. At the same time, the person will in fact make that planned choice only if it is indeed optimal given the expectations induced by that choice. Formally, they define a notion of *personal equilibrium (PE)* to be a choice that a person would want to make given that he expects it.

Definition: Given a choice set $\mathcal{L}$, a lottery $L \in \mathcal{L}$ is a *personal equilibrium (PE)* if $U(L|L) \geq U(L'|L)$ for all $L' \in \mathcal{L}$.

Given its equilibrium nature, PE can lead to multiplicity. For instance, it could easily be the case that L is better than L' when expecting L, while L' is better than L when expecting L'. Example 2 illustrates this possibility:

Example 2: Consider a choice set $\mathcal{L} \equiv \{L_1, L_2\}$ where $L_1 = (0, 1)$ and $L_2 = (-10, \frac{1}{2}; 10 + x, \frac{1}{2})$.

- Given a reference lottery L_1, $U(L_1|L_1) = 0$ while $U(L_2|L_1) = \frac{x}{2} - \frac{1}{2}\eta\lambda(10) + \frac{1}{2}\eta(10+x)$, and thus $U(L_1|L_1) \geq U(L_2|L_1)$ if and only if $x \leq \frac{\eta(\lambda-1)}{1+\eta}10$. In other words, L_1 is a PE when $x \leq \frac{\eta(\lambda-1)}{1+\eta}10$.

- Given a reference lottery L_2, $U(L_1|L_2) = 0 + \frac{1}{2}\eta(10) - \frac{1}{2}\eta\lambda(10+x)$ while $U(L_2|L_2) = \frac{x}{2} - \frac{1}{4}\eta(\lambda-1)(20+x)$, and thus $U(L_2|L_2) \geq U(L_1|L_2)$ if and only if $x \geq 0$. In other words, L_2 is a PE when $x \geq 0$.

- Hence, for any $x \in [0, \frac{\eta(\lambda-1)}{1+\eta}10]$, L_1 and L_2 are both PE.

Example 2 not only illustrates the multiplicity of PE, but also highlights that this multiplicity shows up even in straightforward examples such as a decision whether to accept a 50-50 gamble. On one hand, this multiplicity might provide an explanation for certain behaviors if features of the environment lead one to select one PE over another. On the other hand, however, Koszegi and Rabin point out that there is a natural selection criterion to use. In particular, when there are multiple PE, it means that there are multiple consistent plans. Since, when formulating a plan, the person is of course free to choose which of these consistent plans to follow, it seems natural to assume that the person would choose the consistent plan that yields the largest ex-ante utility (as reflected by $U(L|L)$). KR formalize this as a *preferred personal equilibrium (PPE)*:

Definition: Given a choice set $\mathcal{L}$, a lottery $L \in \mathcal{L}$ is a *preferred personal equilibrium (PPE)* if it is a PE and if for any other PE L', $U(L|L) \geq U(L'|L')$.

To illustrate, we continue with Example 2:

Example 2 (continued):

- Applying the equations above, $U(L_1|L_1) \geq U(L_2|L_2)$ for any $x \leq \frac{\eta(\lambda-1)}{1-\eta(\lambda-1)/2}10$. Because $\frac{\eta(\lambda-1)}{1-\eta(\lambda-1)/2}10 > \frac{\eta(\lambda-1)}{1+\eta}10$, $U(L_1|L_1) > U(L_2|L_2)$ for any $x \in [0, \frac{\eta(\lambda-1)}{1+\eta}10]$. It follows that L_1 is the unique PPE for all $x \leq \frac{\eta(\lambda-1)}{1+\eta}10$, while L_2 is the unique PPE for all $x > \frac{\eta(\lambda-1)}{1+\eta}10$.

- Finally, note that if we apply CPE to this example, L_1 is a CPE for all $x \leq \frac{\eta(\lambda-1)}{1-\eta(\lambda-1)/2}10$, while L_2 is a CPE for all $x \geq \frac{\eta(\lambda-1)}{1-\eta(\lambda-1)/2}10$.

Several features of PE and PPE warrant attention. First, as highlighted in Example 2, often the determinant of when a lottery is a PPE is not ex ante utility, but rather when it is a PE. In Example 2, even though the certain L_1 continues to yield higher ex ante utility for some $x > \frac{\eta(\lambda-1)}{1+\eta}10$, it stops being a PE—that is, even if the person planned on choosing L_1, she wouldn't carry out that plan. Second, and closely related, note the key difference between PPE and CPE: under CPE, the chosen option need not be a PE. In Example 2, for $x \in (\frac{\eta(\lambda-1)}{1+\eta}10, \frac{\eta(\lambda-1)}{1-\eta(\lambda-1)/2}10]$ we have that L_1 is a CPE but not a PPE or PE. Third, and again closely related, in Example 2 CPE yields more risk aversion than PPE. In fact, this reflects a more general result (see Proposition 8 in Koszegi and Rabin (2007)) that CPE yields more risk aversion than PPE.

Fourth, unlike CPE, PE and PPE cannot yield violations of first-order stochastic dominance. For any fixed referent R, $U(L|R)$ is monotonic. Hence, if lottery L' dominates lottery L, then $U(L'|L) > U(L|L)$, and therefore L could not be a PE or a PPE. Finally, because PE is defined relative to the specific options in the choice set, PE and PPE are not easily represented by indifference sets as we have done with models in prior sections.[39]

5.4 Applying Expectations-Based Models

Before turning to specific applications, we briefly discuss several additional issues that emerge when one applies the expectations-based models described above.

[39] Also note that, unlike for CPE, for PE or PPE η and λ enter in ways that cannot be reduced to $\Lambda \equiv \eta(\lambda - 1)$.

First, our discussion describes several different assumptions one might make about the determinants of the referent, and one must make a decision about which approach to apply. Koszegi and Rabin suggest a psychological basis for making this decision as a function of the timing of the decision at hand. In situations where one commits to a choice well in advance of the realization of outcomes, they suggest that the DA/CPE approach is appropriate. Insurance choices might fall into this category. In situations where one makes a plan well in advance of the realization of outcomes but commits to a choice only shortly before the realization of outcomes, they suggest that PE or PPE is appropriate. Shopping decisions might fall into this category (as described in Section 6.4). Finally, in surprise situations where a person is presented with a choice shortly before the realization of outcomes, they suggest using an exogenous reference lottery based on prior expectations.

The potential slipperiness of having multiple solution techniques was not lost on the authors, who note the additional explanatory power of this freedom (see Koszegi and Rabin 2006, pg. 1141 for discussion). In applications of expectations-based loss aversion, researchers often have not been as attentive as they ought to be to the question of what is the proper solution concept. Indeed, because CPE is significantly more tractable and straightforward to work with, it is the solution concept that is most frequently used, including in situations where it might not be appropriate. In Section 7.3, we'll discuss another potential issue for both PPE and CPE: they are static concepts that ignore possible "news" utility associated with making a plan (for PPE) or choice (for CPE).

Second, our presentation above considers the case of choices over simple monetary gambles. In many applications, there will be multiple dimensions of consumption (in their original work, Koszegi and Rabin (2006) frame their model for this domain). It turns out that, if as in Section 3.3 one assumes that intrinsic utility and gain-loss utility are both additively separable across goods, then it is straightforward to extend the approaches in Section 5.2. Specifically, consider lotteries of the form $L \equiv (\mathbf{x}_1, p_1; ...; \mathbf{x}_N, p_N)$ where $\mathbf{x}_n \equiv (x_n^1, ..., x_n^K)$ is a consumption bundle over K dimensions of consumption, and consider reference lotteries of the form $R \equiv (\mathbf{r}_1, q_1; ...; \mathbf{r}_M, q_M)$ where $\mathbf{r}_m \equiv (r_m^1, ..., r_m^K)$ is a reference consumption bundle over the same K dimensions of consumption. Then utility under the KR-approach and the DA-approach are:

$$\text{KR-approach: } U(L|R) \equiv \sum_{n=1}^{N} p_n \left[\sum_{k=1}^{K} \left(u_k(x_n^k) + \sum_{m=1}^{M} q_m \mu(u_k(x_n^k) - u_k(r_m^k) \right) \right]$$

$$\text{DA-approach: } U(L|R) \equiv \sum_{n=1}^{N} p_n \left[\sum_{k=1}^{K} \left(u_k(x_n^k) + \mu \left(u_k(x_n^k) - \sum_{m=1}^{M} q_m u_k(r_m^k) \right) \right) \right]$$

Given $U(L|R)$, we can apply the different solution concepts from Section 5.3

Third, when one applies the PE or PPE solution concepts, it might be natural to extend the model to incorporate exogenous uncertainty that is resolved between the time that plans are formulated and the time that a choice is made (again, in their original work, Koszegi and Rabin (2006) frame their model for this domain). When shopping, for instance, it could be that one forms an ex ante plan for whether to purchase an item as a function of the market price, but does not learn the market price until arriving at the store. It is straightforward to extend the model to such situations. Following Koszegi and Rabin (2006), let the exogenous uncertainty be reflected in a state s with distribution $Q(s)$, where state s generates a choice set $\mathcal{L}(s)$. Because a final choice is made after s is revealed, an ex ante plan is a strategy L^* where $L^*(s) \in \mathcal{L}(s)$ is the planned choice for state s. Finally, let G_{L^*} denote the ex ante distribution over outcomes (i.e., over the $\mathbf{x}_n$'s) induced by Q and L^*. Then a strategy L^* is a PE if, for every state s, $U(L^*(s)|G_{L^*}) \geq U(L|G_{L^*})$ for all $L \in \mathcal{L}(s)$, and L^* is a PPE if it is a PE and for any other PE $L^{*'}$ we have $U(G_{L^*}|G_{L^*}) \geq U(G_{L^{*'}}|G_{L^{*'}})$. We consider an example of this type in the next section.

6 Applications of Expectations-Based Models

Koszegi and Rabin's work inspired a wave of applied work that both develops implications of expectations-based reference dependence and tests many of those implications. Once again, much of this work focuses on the domains of the endowment effect and daily labor supply as these domains again permit sharp tests of model predictions. We begin our discussion with these two domains, but then move on to work in other, more economically important domains. As we'll see, not surprisingly, the evidence has not provided clean answers. At the same time, however, this work has sharpened researchers' thinking about reference-dependent preferences.

6.1 Endowment Effect

Section 4.1 describes early evidence on the endowment effect, and develops a simple model that illustrates the loss-aversion interpretation of the endowment effect. Numerous studies prior to the turn of the century seemed to find robust support for the endowment effect—indeed, Knetsch et al. (2001) write (p. 257), "The endowment effect and loss aversion have been among the most robust findings of the psychology of decision making." However, research since 2000 has questioned this robustness, which in turn has led to a reexamination of the simple loss-aversion interpretation, including an assessment of whether models of expectations-based reference dependence might better explain behavior.

A first line of inquiry into the robustness and field validity of the endowment effect revolved around the role of market experience. In two influential papers, List (2003, 2004) demonstrates that market experience can reduce and perhaps even eliminate the endowment effect. Specifically, he considers subjects at a sportscard show, where he compares behavior of professional dealers, experienced nondealers, and inexperienced nondealers.[40] List investigates whether these subjects exhibit an endowment effect for sports memorabilia and for

[40]The partition of nondealers into experienced vs. inexperienced is done using survey data on the average number of sports memorabilia trades they make per month.

mugs vs. candy bars. In both cases, inexperienced nondealers exhibit significant endowment effects, people with more trading experience exhibit smaller endowment effects, and professional dealers seem to exhibit no endowment effect. These results suggest that the endowment effect might have limited relevance for field exchange behavior if exchanges are generally undertaken by experienced agents.

A second line of inquiry into the robustness and field validity of the endowment effect investigated whether the basic result itself is "real" or instead merely an artifact of experimental protocols. Plott and Zeiler (2005, 2007) argue that the language used when providing subjects with their endowments, visual and physical cues, and the language used to describe the choice options can all serve to create experimenter demand effects. They demonstrate how procedural variations along these dimensions can alter the magnitude of the endowment effect, and in fact under their "full set of procedural controls" they are able to eliminate the endowment effect entirely.

Early responses to these critiques argued that the additional context—i.e., market experience or procedural controls—served to alter people's reference points. Perhaps people with market experience do not code a newly obtained object as something to be consumed, and thus don't feel a sense of loss when they give it up. Perhaps Plott and Zeiler's procedural controls actually served to undermine a feeling of endowment. With the reemergence of expectations-based loss aversion in Koszegi and Rabin (2006), these arguments became naturally framed in terms of expectations.

It is instructive to formalize these arguments. For comparison to List (2003, 2004), Plott and Zeiler (2007), and the more recent work below, we do so in the context of the exchange paradigm. Consider the mug vs. pen version of the endowment effect model from Section 4.1 in which final consumption is a vector (x_M, x_P), where $x_M \in \{0, 1\}$ is mug consumption and $x_P \in \{0, 1\}$ is pen consumption. Recall that v_M and v_P are the intrinsic valuations for the two items, and we apply the two-part linear universal gain-loss utility function in equation 5.

Suppose that standard procedures in endowment-effect experiments naturally lead a person endowed with a mug to relatively quickly expect to leave the experiment with that mug—i.e., to adopt a reference bundle $\mathbf{r}^{MUG} \equiv (1,0)$. Analogously, a person endowed with a pen might quickly adopt a reference point $\mathbf{r}^{PEN} \equiv (0,1)$. If subjects then make choices with these (now exogenous) reference points, the predictions are equivalent to the endowment-effect model in Section 4.1—specifically, subjects endowed with a pen choose to keep the pen if $v_P > \frac{1+\eta}{1+\eta\lambda} v_M$, whereas subjects endowed with a mug choose to trade for a pen if $v_P > \frac{1+\eta\lambda}{1+\eta} v_M$.

Next, consider how market experience or procedural controls might alter behavior. An interpretation of the intuition above is that market experience or procedural controls can break the link between being endowed with an object and an expectation to consume that object. In other words, they might lead subjects to adopt a reference point $\mathbf{r}^0 \equiv (0,0)$ regardless of endowment, in which case they would choose pen when $v_P > v_M$ regardless of endowment.[41] Koszegi and Rabin (2006) explicitly suggest these interpretations, and Knetsch and Wong (2009) conduct an experiment with procedures explicitly designed to alter expectations and find evidence consistent with these interpretations.

The logic above assumes exogenous (at the time of choice) expectations determined by a combination of past experience (e.g., market experience) and experimental procedures. One might instead consider endogenous expectations as in Section 5.3.2 and 5.3.3. However, because endogenous expectations are forward-looking and thus depend on treatment only if treatment impacts the choice set, endogenous expectations have trouble generating an endowment effect. For example, in the standard mugs-vs.-pens paradigm, the choice set is $\{\mathbf{c}^{MUG} \equiv (1,0), \mathbf{c}^{PEN} \equiv (0,1)\}$ regardless of endowment. Example 3 works out predictions under various models of endogenous expectations.

---

[41]One could also imagine intermediate cases. For instance, if upon being endowed with a mug a person comes to expect there is a 60% chance of leaving with a mug and a 40% chance of something else, then she might adopt a reference lottery $R \equiv ((1,0), .6; (0,0), .4)$.

Example 3: Consider a choice set $\{\mathbf{c}^{MUG} \equiv (1,0), \mathbf{c}^{PEN} \equiv (0,1)\}$.

- CPE: $U(\mathbf{c}^{MUG}|\mathbf{c}^{MUG}) = v_M$ and $U(\mathbf{c}^{PEN}|\mathbf{c}^{PEN}) = v_P$. $\mathbf{c}^{PEN}$ is a CPE if $U(\mathbf{c}^{PEN}|\mathbf{c}^{PEN}) \geq U(\mathbf{c}^{MUG}|\mathbf{c}^{MUG})$, or $v_P \geq v_M$. Analogously, $\mathbf{c}^{MUG}$ is a CPE if $v_P \leq v_M$.
- PE: $U(\mathbf{c}^{MUG}|\mathbf{c}^{PEN}) = v_M + [-\eta\lambda v_P + \eta v_M]$. $\mathbf{c}^{PEN}$ is a PE if $U(\mathbf{c}^{PEN}|\mathbf{c}^{PEN}) \geq U(\mathbf{c}^{MUG}|\mathbf{c}^{PEN})$, or $v_P \geq \frac{1+\eta}{1+\eta\lambda} v_M$. Analogously, $\mathbf{c}^{MUG}$ is a PE if $v_P \leq \frac{1+\eta\lambda}{1+\eta} v_M$.
- PPE: For $v_P \in \left[\frac{1+\eta}{1+\eta\lambda} v_M, \frac{1+\eta\lambda}{1+\eta} v_M\right]$, $\mathbf{c}^{PEN}$ and $\mathbf{c}^{MUG}$ are both PE. Comparing $U(\mathbf{c}^{MUG}|\mathbf{c}^{MUG})$ and $U(\mathbf{c}^{PEN}|\mathbf{c}^{PEN})$, $\mathbf{c}^{PEN}$ is ex-ante preferred when $v_P \geq v_M$. It follows that $\mathbf{c}^{PEN}$ is a PPE for $v_P \geq v_M$, and $\mathbf{c}^{MUG}$ is a PPE for $v_P \leq v_M$.

From Example 3, we see that CPE and PPE permit no possibility of an endowment effect in the standard experimental design. Again, because expectations are fully forward-looking, endowment is irrelevant, and in fact the prediction under either CPE or PPE is that people choose the item that yields the largest intrinsic value. PE, in contrast, permits some scope for an endowment effect. Specifically, because for some people—those with $v_P \in \left[\frac{1+\eta}{1+\eta\lambda} v_M, \frac{1+\eta\lambda}{1+\eta} v_M\right]$—there are multiple PE, an endowment effect could emerge if endowment has a systematic effect on equilibrium selection wherein those endowed with item k are more prone to follow the PE of choosing c_k. Of course, interpreting the endowment effect as driven entirely by the needed equilibrium selection is not entirely satisfying.

Recently, researchers have attempted to directly test whether endogenous expectations are impacting behavior in the endowment-effect paradigm by altering the paradigm in a way that should impact behavior under endogenous expectations but not under exogenous expectations. One possibility—pursued by Ericson and Fuster (2011) and Heffetz and List (2014)—is to introduce a chance that trade is not permitted. In other words, subjects are endowed with one item, then told that they might be able to trade it for another item, but, in the event they choose to trade, with probability q (varied across subjects) the trade will not be permitted.

With this modification, the choice set depends on endowment. For subjects endowed with a pen, the choice set becomes $\{F_{PEN} \equiv ((0,1),1), F_{MUG} \equiv ((1,0),1-q;(0,1),q)\}$. In other words, if a subject chooses to keep the pen, she'll end up with the pen no matter whether trade is permitted, and thus she is choosing lottery F_{PEN}. If instead she chooses to trade for a mug, the trade will succeed with probability $1-q$, and thus she is choosing lottery F_{MUG}. Using an analogous logic, for subjects endowed with a mug, the choice set becomes $\{F'_{PEN} \equiv ((0,1),1-q;(1,0),q), F'_{MUG} \equiv ((1,0),1\}$. Example 4 derives predictions for subjects endowed with pens under various models of the referent. (Given the symmetry of this model, predictions for subjects endowed with mugs are analogous.)

Example 4: Consider a choice set

$$\{F_{PEN} \equiv ((1,0),1), F_{MUG} \equiv ((1,0),1-q;(0,1),q)\}.$$

- Given an exogenous reference point $\mathbf{r}^{PEN} \equiv (0,1)$, the person will choose F_{PEN} if $v_P \geq \frac{1+\eta}{1+\eta\lambda} v_M$, because $U(F_{PEN}|\mathbf{r}^{PEN}) = v_P$ and $U(F_{MUG}|\mathbf{r}^{PEN}) = qv_P + (1-q)\,[v_M - \eta\lambda v_P + \eta v_M]$.

- CPE: $U(F_{PEN}|F_{PEN}) = v_P$ and $U(F_{MUG}|F_{MUG}) = qv_P + (1-q)v_M - q(1-q)\Lambda(v_P + v_M)$. F_{PEN} is a CPE if $U(F_{PEN}|F_{PEN}) \geq U(F_{MUG}|F_{MUG})$, or $v_P \geq \frac{1-q\Lambda}{1+q\Lambda} v_M \equiv \bar{v}^{CPE}(q)$. Analogously, F_{MUG} is a CPE if $v_P \leq \bar{v}^{CPE}(q)$. Note that $\bar{v}^{CPE}(q)$ is decreasing in q.

- PE: $U(F_{MUG}|F_{PEN}) = qv_P + (1-q)v_M + (1-q)\,[-\eta\lambda v_P + \eta v_M]$. F_{PEN} is a PE if $U(F_{PEN}|F_{PEN}) \geq U(F_{MUG}|F_{PEN})$, or $v_P \geq \frac{1+\eta}{1+\eta\lambda} v_M \equiv \bar{v}^{KEEP}$. $U(F_{PEN}|F_{MUG}) = v_P + (1-q)\,[-\eta\lambda v_M + \eta v_P]$. F_{MUG} is a PE if $U(F_{MUG}|F_{MUG}) \geq U(F_{PEN}|F_{MUG})$, or

$v_P \leq \frac{1+q\eta+(1-q)\eta\lambda}{1+q\eta\lambda+(1-q)\eta} v_M \equiv \bar{v}^{TRADE}(q)$. Note that $\bar{v}^{TRADE}(0) = \frac{1+\eta\lambda}{1+\eta} v_M$, $\bar{v}^{TRADE}(q)$ is decreasing in q, and $\bar{v}^{TRADE}(1) = \bar{v}^{KEEP}$.

- PPE: For $v_P \in \left[\bar{v}^{KEEP}, \bar{v}^{TRADE}(q)\right]$, F_{PEN} and F_{MUG} are both PE. Comparing $U(F_{PEN}|F_{PEN})$ and $U(F_{MUG}|F_{MUG})$, F_{PEN} is ex-ante preferred when $v_P \geq \bar{v}^{CPE}(q)$. It follows that F_{PEN} is a PPE for $v_P \geq \max\left\{\bar{v}^{CPE}(q), \bar{v}^{KEEP}\right\} \equiv \bar{v}^{PPE}(q)$, and otherwise F_{MUG} is a PPE. Note that $\bar{v}^{CPE}(0) > \bar{v}^{KEEP}$, $\bar{v}^{CPE}(q)$ is decreasing in q, and there exists $q' \in (0, 1/2)$ such that $\bar{v}^{CPE}(q) < \bar{v}^{KEEP}$ for all $q > q'$.

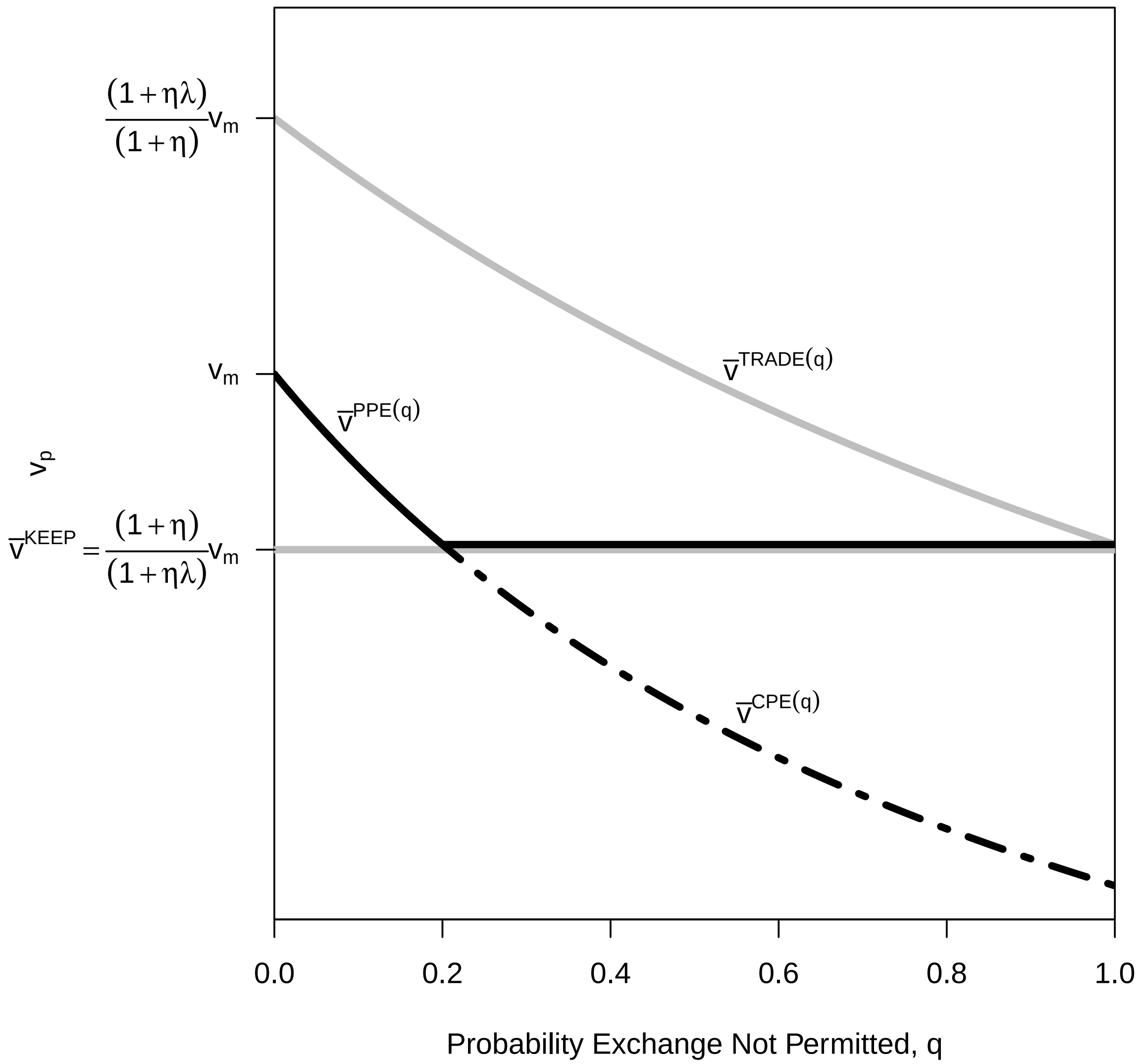
- Figure 5 depicts $\bar{v}^{CPE}(q)$, $\bar{v}^{KEEP}$, $\bar{v}^{TRADE}(q)$, and $\bar{v}^{PPE}(q)$.

From Example 4, if endowing a person with a pen leads that person to adopt a reference point $\mathbf{r}^{PEN} \equiv (0, 1)$, then trading decisions are not affected by the possibility that trade is not permitted. In contrast, under CPE or PPE, an increase in q makes people endowed with pens more likely to choose to keep the pen (because the cutoff $\bar{v}$ declines).[42] Because an analogous logic holds for people endowed with mugs, it follows that, under CPE or PPE, an increase in q should increase exchange asymmetries. Intuitively, as the probability of being stuck with the endowed item increases, the person's expectations of leaving with that item increase, and thus the person is more likely to choose to keep it.

Ericson and Fuster (2011) and Heffetz and List (2014) provide similar experimental designs that manipulate the probability with which exchange is not permitted. The results are mixed. Using a relatively small sample (45 subjects) and endowing everyone with the same good, Ericson and Fuster (2011) find support for endogenous expectations: subjects who faced $q = 10\%$ were more willing to trade than subjects who faced $q = 90\%$.[43] Across three

[42] Under PPE, this effect holds for small q. For q large enough, further increases in q do not affect trading decisions.

[43] Ericson and Fuster (2011) also use a similar approach in a valuation task, and again find support for endogenous expectations. Camerer et al. (2016) conduct a replication of this second Ericson and Fuster (2011) experiment as part of their large replication exercise of experimental economics findings. The Ericson and Fuster (2011) treatment effect was within the confidence interval of the replication and the replication's treatment effect was marginally significant.



Note: Figure presents the four thresholds defined in Example 4 for a person who is endowed with a pen when the probability that exchange is not permitted is q. Keeping the pen is a CPE for any $v_P \geq \bar{v}^{CPE}(q)$, and trading for a mug is a CPE for any $v_P \leq \bar{v}^{CPE}(q)$. Keeping the pen is a PE for any $v_P \geq \bar{v}^{KEEP}$, trading for a mug is a PE for any $v_P \leq \bar{v}^{TRADE}(q)$, and both are PE for any $v_P \in [\bar{v}^{KEEP}, \bar{v}^{TRADE}(q)]$. Keeping the pen is a PPE for any $v_P \geq \bar{v}^{PPE}(q)$, and trading for a mug is a PPE for any $v_P < \bar{v}^{PPE}(q)$.

Figure 5: Exchange Behavior When Trade Possibly Not Permitted

experiments with a substantially larger sample (a total of 560 subjects), Heffetz and List (2014) do not find a significant impact of $q = 1\%$ vs. $q = 99\%$ or of $q = 10\%$ vs. $q = 90\%$.

Moreover, unlike Ericson and Fuster (2011), Heffetz and List (2014) randomly endow subjects with one of the two items, and this endowment impacts behavior (independent of q) in a way that is consistent with standard endowment effects.

In an alternative approach, Goette et al. (Forthcoming) instead introduce a possibility that subjects are forced to trade. In other words, subjects are endowed with one item, then told that they'll have the opportunity to trade it for another item, but, in the event that they choose to keep the item, with probability q they'll be forced to trade it. Again, this modification makes the choice set depend on endowment, and, interestingly, flips the choice sets relative to the case where trade might not be permitted. Specifically, subjects endowed with pens face choice set $\{F'_{PEN} \equiv ((0,1), 1-q; (1,0), q), F'_{MUG} \equiv ((1,0), 1\}$, while subjects endowed with mugs face choice set $\{F_{PEN} \equiv ((0,1), 1), F_{MUG} \equiv ((1,0), 1-q; (0,1), q)\}$. We can thus use the derivations in Example 4, except that these now apply for subjects endowed with mugs.

Several conclusions follow. First, while not derived in Example 4, it is straightforward to derive that if endowing a person with a mug leads that person to adopt a reference point $\mathbf{r}^{MUG} \equiv (1,0)$, then trading decisions are not affected by the possibility of forced trade. Second, under CPE or PPE, an increase in q makes people endowed with mugs more likely to trade for pens.[44] Because an analogous logic holds for people endowed with pens, it follows that, under CPE or PPE, an increase in q should yield increased exchange asymmetries in the opposite direction from the endowment effect—people endowed with pens are more likely to choose mugs than people endowed with mugs. Intuitively, as the probability of being forced to trade for the alternative item increases, the person's expectations of leaving with that alternative item increase, and thus the person is more likely to trade to get it.

Goette et al. (Forthcoming) provide a set of three experiments to test this prediction. They do so in the context of a valuation task, but the predictions are analogous—e.g., under CPE or PPE, $q > 0$ implies $WTP > WTA$, and the larger is q the larger is this reverse

[44] Again, under PPE, this effect holds for small q, but for q large enough further increases in q do not affect trading decisions.

WTP-WTA disparity. However, Goette et al. (Forthcoming) find that q seems to have limited impact on behavior in two of their three experiments. Moreover, for all q that they consider, they find a standard endowment effect, and never find a reverse endowment effect. WTA exceeds WTP by around 50 percent when $q = 0$ and similar magnitude gaps are observed across a range of probabilities in two of their three experiments. In one experiment the directional effects of the KR predictions are supported.

In a recent working paper, Heffetz (2018) raises the possibility that beliefs might need to "sink in" before they become the referent, and further suggests that the different results in the literature might be due to experimental differences that impact the extent to which beliefs sink in before choices are made. To investigate this possibility, Heffetz develops a sink-in manipulation in which subjects experience the relevant probability (a 1-in-6 chance) through a series of 18 die rolls before making the choice of interest. When Heffetz (2018) applies this manipulation in the Heffetz and List (2014) paradigm, he indeed finds evidence consistent with the CPE/PPE prediction. (However, this sink-in manipulation does not have an analogous impact in the labor-supply paradigm discussed in the next subsection.)

Taken together, the combined findings of Heffetz and List (2014) and Goette et al. (Forthcoming) perhaps lend support to the original interpretation of Koszegi and Rabin (2006). In other words, perhaps subjects in endowment-effect experiments are not heavily influenced by forward-looking reference points. Rather, in many instances, their behavior is quite consistent with, upon being endowed with an item, quickly expecting to leave the experiment with that item. Of course, the results of List (2003, 2004), Plott and Zeiler (2005, 2007), Ericson and Fuster (2011), and Heffetz (2018) suggest that it is sometimes possible to break this expectation. Clearly more work is needed to fully understand behavior in this domain.

6.2 Labor Supply

One of the most prominent debates surrounding reference-dependent preferences revolves around the possibility of daily income targeting in the domain of labor supply. In this section, we describe this debate and the role of expectations-based loss aversion in this debate. In addition, we also describe recent experiments on labor supply in real-effort tasks that were designed to explicitly test expectations-based loss aversion.

In Section 4.2, we described the basic hypothesis of daily income targeting, and how it could sometimes—but not always—lead to negative wage elasticities. We also described the seminal work by Camerer et al. (1997) who indeed found evidence of negative wage elasticities using three datasets of New York City cab drivers.

Farber (2005) points out two issues in the approach of Camerer et al. (1997). First, the approach relies on an assumption that there exists a "daily wage," and Farber presents evidence suggesting substantial variation in each driver's hourly wage over the course of a day. Second, Farber revisits the issue of division bias due to hourly wages being calculated using earnings divided by hours. Camerer et al. (1997) recognize and attempt to address this issue using an instrumental variables strategy, but Farber argues that the instrument is not effective. But Farber's main contribution is to point out a better approach that avoids both problems: treat the unit of observation as a trip, and estimate a discrete-choice stopping model.

Specifically, Farber (2005) uses a reduced-form approach that assumes driver i will stop after trip τ if $R_{i\tau} \geq 0$ where

$$R_{i\tau} \equiv \gamma_h h_{i\tau} + \gamma_y y_{i\tau} + \mathbf{X}_i \beta + \delta_i + \varepsilon_{i\tau}. \tag{9}$$

In this specification, $h_{i\tau}$ and $y_{i\tau}$ are cumulative hours worked and cumulative earnings after driver i completes trip τ, $\mathbf{X}_i$ is a vector of control variables, and δ_i captures individual fixed effects. A standard labor-supply model would imply that cumulative hours should have an impact on the stopping probabilities ($\gamma_h > 0$) because as cumulative hours increase the marginal cost of further effort also increases. In contrast, a standard model would imply that cumulative earnings should not affect stopping probabilities ($\gamma_y = 0$).[45] Farber's estimates

[45]More precisely, this conclusion holds under the reasonable assumption that daily earnings are such a

support both predictions, albeit with wide confidence intervals in the case of $\hat{\gamma}_y$.

In terms of ruling out daily income targeting, Farber (2005) has a potential problem: he is using a misspecified model of the role of cumulative earnings. Given the specification in equation 9, cumulative earnings has a linear impact on the probability of stopping. In contrast, the model from Section 4.2 suggests cumulative earnings should have a threshold impact on the probability of stopping—that is, $\gamma_y y_{i\tau}$ should be replaced with $\theta_y I(y_{i\tau} \geq \bar{y})$ where I is an indicator function and $\bar{y}$ is target earnings. Moreover, if stopping typically occurs around target earnings, then the vast majority of observations will have $y_{i\tau} < \bar{y}$. As a result, even if there is daily income targeting, one might expect to estimate $\hat{\gamma}_y \approx 0$ given the specification in Farber (2005).[46]

Farber (2008) addresses this problem by instead using a threshold specification for the impact of cumulative earnings. However, because target earnings are unobserved, he cannot take an analogous reduced-form approach. Instead, he estimates a structural model in which target earnings is a latent variable that varies across drivers and across days—specifically, he assumes the target, $\bar{y}$, is normally distributed. When Farber estimates this model, he finds significant evidence of income targeting ($\hat{\theta}_y > 0$). At the same time, he finds evidence of substantial day-to-day and between-driver variation in target earnings, and moreover that most shifts appear to stop before target earnings are met. From these findings, Farber concludes that, while income targeting may occur, it has limited value for predicting behavior both in the cross-section and over time, and thus has limited value for economics.

Crawford and Meng (2011) incorporate theoretical insights from expectations-based loss aversion into this debate. In particular, they argue that we need not treat target earnings as a latent variable, because we can assume that target earnings $\bar{y}$ are equal to recent expectations about earnings (operationalized as average earnings on days in the recent past). In effect,

small part of lifetime earnings that they have no impact of the marginal utility of lifetime income.

[46]The logic behind this claim combines the selection problem of seeing few observations beyond the target and the misspecification of functional form. For a reference-dependent agent, earnings should have zero impact on stopping for all $y_{i\tau} < \bar{y}$, and θ_y impact on stopping for all $y_{i\tau} > \bar{y}$. If the vast majority of observations have $y_{i\tau} < \bar{y}$, the linear regression line would have approximately zero slope.

the challenge of predicting the target by Farber (2008) is overcome via the assumption of rational expectations. Moreover, they argue that drivers should exhibit expectations-based loss aversion over hours worked as well, where target hours $\bar{h}$ are equal to recent expectations about hours (operationalized as average hours on days in the recent past).

Based on this logic, Crawford and Meng (2011) use a reduced-form approach that assumes driver i will stop after trip τ if $R_{i\tau} \geq 0$ where

$$R_{i\tau} \equiv \gamma_h h_{i\tau} + \theta_h I(h_{i\tau} > \bar{h}) + \gamma_y y_{i\tau} + \theta_y I(y_{i\tau} > \bar{y}) + \mathbf{X}_i \beta + \delta_i + \varepsilon_{i\tau}. \tag{10}$$

They find (see their Table 3) that, consistent with expectations-based loss aversion, the probability of stopping depends on cumulative earnings only through the threshold effect (i.e., $\hat{\theta}_y > 0$ and $\hat{\gamma}_y \approx 0$). Moreover, they also find that the probability of stopping depends on cumulative hours through the linear effect ($\hat{\gamma}_h > 0$)—as one would expect from a standard labor-supply model—and also through the threshold effect ($\hat{\theta}_h > 0$), which is consistent with expectations-based loss aversion over hours worked.[47] Crawford and Meng (2011) go on to estimate a structural model that permits an estimate of the degree of loss aversion. Interestingly, their estimates suggest stronger loss aversion over hours worked, although they cannot reject equal loss aversion along the two dimensions.[48]

Thakral and To (2017) provide a new perspective on this debate by investigating the importance of the timing of earnings within a day. Using a dataset of all New York City taxi trips in 2013, and using a more flexible version of the approach in Farber (2005), Thakral and To first find that the probability of stopping depends on cumulative earnings for the day as predicted by income targeting. More importantly, when, unlike the prior literature, they

[47] Agarwal et al. (2015) apply a similar approach to data on Singapore taxicab drivers, and find similar results.

[48] Farber (2015) revisits labor supply decisions with a much larger and richer dataset. Interestingly, while he attempts to explicitly test the possibility of expectations-based loss aversion, he does so within the context of estimating wage elasticities—while still acknowledging that this approach might be problematic due to there not being a "daily wage." Farber (2015) also considers a stopping-model approach, and while he uses a more flexible approach than in Farber (2005) for how cumulative hours and cumulative earnings might impact the probability of stopping, he does not permit a threshold based on expected hours or earnings.

allow for a differential impact of earnings from different times in the day, they find that the probability of stopping is most sensitive to earnings in the most-recent past. This finding is inconsistent with the typical assumption in the prior literature that there is a daily referent (that might be exogenous or based on expectations for that day). Instead, to explain their findings, Thakral and To develop a model in which the referent adjusts over the course of a shift in response to realized earnings—so that by, say, the eighth hour of the shift, earnings from the first few hours of the shift are mostly incorporated into the referent and thus have little impact on stopping, while earnings from the most-recent hours of the shift are mostly not incorporated into the referent and thus have a significant impact on stopping.

In parallel to the field analyses above, several experiments have studied expectations-based loss aversion in the domain of labor supply from a different angle: they use real-effort tasks as a means to conduct explicit tests of the theory. The general approach asks subjects to engage in some sort of experimental task that requires effort, where subjects are paid as a function of that effort. The key feature is that payments are delivered in the form of a lottery, designed such that if that lottery becomes a subject's reference lottery, then features of that lottery will have a predictable impact on the subject's behavior.

Abeler et al. (2011) use an individual decision problem in which subjects exert effort to count the number of zeros in tables that consist of 150 randomly ordered zeros and ones. Subjects were permitted to continue as long as they liked, up to a maximum of 60 minutes (on average subjects worked roughly 35 minutes). Letting e denote the number of tables solved correctly, subjects were paid according to the lottery $(we, \frac{1}{2}; F, \frac{1}{2})$. In other words, there was a 50% chance that a subject was paid according to her acquired earnings at some wage rate w (20 cents in the experiment), and there was a 50% chance she was paid a fixed payment F.[49]

The payment scheme was chosen to test for expectations-based loss aversion. For standard

[49] This lottery was implemented by having subjects choose, before starting the task, between two envelopes, one which contained a card saying "Acquired earnings" and one which contained a card saying "F euros." The chosen envelope was opened after the subject decided to stop working.

agents (including standard reference-dependent agents), the fixed payment F should be irrelevant to the choice of when to stop as it does not alter marginal benefits. In contrast, for agents with expectations-based loss aversion—in particular for whom the lottery $(we, \frac{1}{2}; F, \frac{1}{2})$ becomes their reference lottery—F influences their choice. Analogous to the logic from Figure 3, there exists a kink in marginal benefits of effort at exactly $we = F$. For any e such that $we \leq F$, increases in e reduce gain-loss disutility, and thus the marginal benefit from e is larger than it would be if the person only cared about intrinsic utility. Analogously, for any e such that $we \geq F$, increases in e increase gain-loss disutility, and thus the marginal benefit from e is smaller than it would be if the person only cared about intrinsic utility.

To test for expectations-based loss aversion, Abeler et al. (2011) consider two treatments, one with $F = 3$ euros, and one with $F = 7$ euros. They focus on two predictions of the intuition above. First, the discrete drop in the marginal benefit at $we = F$ implies that there should be bunching at $we = 3$ in the $F = 3$ treatment and bunching at $we = 7$ in the $F = 7$ treatment. Second, for a person who would choose $we \in [3, 7)$ in the $F = 3$ treatment, the increase in the marginal benefit in this range should lead them to choose a larger we in the $F = 7$ treatment.[50] As a result, on average we should be larger in the $F = 7$ treatment. They indeed find support for both predictions.[51] Camerer et al. (2016) conduct a replication of the Abeler et al. (2011) experiment as part of their large replication exercise of experimental economics findings. While the Abeler et al. (2011) treatment effect was within the confidence interval of the replication, the replication's treatment effect was not statistically significant.

In independent and contemporaneous work, Gill and Prowse (2012) test expectations-based loss aversion using a different payoff scheme. They consider a two-player sequential-

[50] Abeler et al. (2011) formally derive these predictions for CPE (which is equivalent to DA in this domain given the binary lottery). One can show that the predictions are qualitatively the same under PPE.

[51] The data are not consistent with the more precise prediction that the two treatments should not impact the percent of participants who choose $we > 7$, as only 42% do so in the $F = 3$ treatment while 55% do so in the $F = 7$ treatment (percentages calculated from Figure 1 in Abeler et al. (2011)). The data also show suggestive evidence of subjects using focal points—e.g., choosing we equal to 5 or 10—raising the question whether the treatment is serving as a focal point rather than changing the reference lottery.

move game in which subjects exert effort to adjust sliders to a target position using their computer mouse. First, player 1 has 120 seconds to correctly adjust as many sliders as possible (out of 48). Next, player 2 observes player 1's performance and then carries out the same task. Finally, one player earns a prize v, where the probability that each player wins the prize is determined by their performances—specifically, letting e_i denote the number of correctly adjusted sliders by player i, player i's probability of winning the prize is

$$P_i(e_i, e_j) = \frac{50 + e_i - e_j}{100}.$$

Throughout their analysis, Gill and Prowse focus on the behavior of player 2, because the goal is to analyze player 2's behavior as a function of the observed e_1.[52] For standard agents (including standard reference-dependent agents), e_1 should have no impact on e_2, because the marginal benefit from increasing e_2 is independent of e_1—i.e., $(\partial P_2/\partial e_2)v$ is independent of e_1. Gill and Prowse show that, for agents who behave according to DA/CPE (which are the same in this domain), player 2 should react to an increase in e_1 by reducing e_2. In their experiment, they indeed find evidence of this discouragement effect. They further use their data to estimate a structural model, and they estimate an average $\Lambda = 1.7$ (although with significant heterogeneity, including 17% with $\Lambda < 0$).

Gneezy et al. (2017) expand on the analysis in Abeler et al. (2011). In particular, they use the same experimental structure, except that subjects were paid according to the lottery $(we, \frac{1}{2}; F, p; 0, \frac{1}{2}-p)$ where $F > 0$ and $p \leq 1/2$. The two treatments in Abeler et al correspond to $p = 1/2$ and $F = 3$ or 7. Gneezy et al. consider those treatments, and further consider $p = 1/2$ and $F = 0$ or 14, and thus they consider further variation in the fixed payment. Moreover, Gneezy et al. also consider treatments with $F = 14$ and $p = .125$, .250, and .375. They show that models of expectations-based loss aversion imply that effort should be monotonically increasing in the fixed payment when $p = 1/2$ (just as in Abeler et al),

[52] Hence, in principle, the authors could have conducted the same experiment with only one subject while varying e_1 exogenously.

and that effort should be monotonically increasing in p when $F = 14$. The data contradict these predictions of monotonicity. Specifically, when $p = 1/2$ they find that average effort is lowest when $F = 3$, and when $F = 14$, they find that average effort is lowest when $p = .25$.

To investigate whether the results in Gneezy et al. (2017) might be in part due to beliefs not sinking in and thereby becoming the referent, Heffetz (2018) also incorporates his sink-in manipulation (discussed in Section 6.1) to a variant of the Gneezy et al. (2017) paradigm. Unlike in the endowment-effect paradigm, here the sink-in manipulation had no impact on behavior, and the results are consistent with the results in Gneezy et al. (2017)—and inconsistent with the predictions of CPE or PPE.

As with exchange experiments, the evolving literature on expectations-based reference dependence and effort provision requires further development. In field settings, the plausibility of endogenous expectations influencing labor supply is clear. Setting these expectations in experimental studies seems to be more tenuous. It may be hard to establish relevant expectations with other aspects of the decision environment serving as salient points of reference. In Section 9, we return to such issues in our concluding thoughts.

6.3 Job Search

Recently, DellaVigna et al. (2017) investigate reference dependence in job search. Interestingly, their work seems to suggest that, in this domain, a referent based on recent experience better explains the data than a referent based on forward-looking rational expectations. In addition, their work highlights some additional issues that can emerge as reference-dependent preferences are taken to more sophisticated economic applications.

The starting point is a well-known empirical finding from labor economics. In many countries, unemployment insurance follows a two-step design where there is a constant replacement rate for some fixed period, followed by lower benefits. In such countries, the hazard rate from unemployment typically follows a U-shaped pattern: high hazard rates in the days immediately following the loss of a job, then a decline in hazard rates, followed by an increase that peaks when the higher benefit level expires. In standard job-search models, this pattern has been attributed to unobserved heterogeneity wherein the initial decline is caused by selection out of the most productive job searchers, and then the later peak is driven by the remaining less productive job searchers wanting to find a job before the higher benefit level expires.

DellaVigna et al. (2017) suggest an alternative, reference-dependent interpretation of this behavior that does not require unobserved heterogeneity. If a person's reference point is her past earnings, then a newly unemployed person will feel a strong feeling of loss and thus will exert substantial search effort to eliminate this loss. As time passes, however, the worker's reference point will decline toward the benefit level, and as it does, so will her incentive to search. Finally, as the person approaches the end of the higher benefits, she'll want to find a job before the higher benefit level expires.

DellaVigna et al. (2017) do not merely propose an alternative account for a known empirical result. In fact, they identify a country (Hungary) that implemented a regime shift from a standard two-step design to a three-step design in which the initial period is split into two sub-periods, the first with a higher benefit than before and the second with a lower benefit than before. DellaVigna et al. (2017) demonstrate that the standard model and the reference-dependent model make distinct predictions for how hazard rates should react to this regime change, and they demonstrate that the observed (reduced-form) empirical reaction is much more in line with the predictions of the reference-dependent model.

An important modeling decision for DellaVigna et al. (2017) revolves around what the referent is. Because job-search models are dynamic in the sense that an individual makes multiple decisions over time, one must model how the referent might change over time, and the extent to which individuals correctly understand how the referent changes over time. Given that their hypothesis about the role of reference dependence relies on people having a backward-looking referent that adapts over time, DellaVigna et al. assume that the referent is an average of recent earnings, and they further assume that people fully understand how the referent changes over time. They also consider a more forward-looking rational-expectations referent—specifically, the referent for period t is expected period-t earnings from the perspective of period $t-1$. This formulation does not perform as well as the referent based on recent experience.

In terms of dealing with a more sophisticated application, there are two additional notable features of DellaVigna et al. (2017) that do not frequently appear in the literature on reference-dependent preferences. First, because they are dealing with a domain in which unobserved heterogeneity is acknowledged to be an important factor, they expand beyond their reduced-form empirical analysis to estimate a structural model that incorporates unobserved heterogeneity. Doing so permits them to assess the value of reference dependence in interpreting behavior while permitting the standard model to have a fair amount of flexibility. They in fact show that the reference-dependent model without unobserved heterogeneity outperforms a standard model with unobserved heterogeneity.

Second, DellaVigna et al. (2017) permit multiple departures from the standard model. Although their estimated structural model of reference dependence outperforms a standard model with unobserved heterogeneity, the resulting estimates are not entirely satisfactory because people are estimated to be extremely impatient. DellaVigna et al. (2017) demonstrate that if in addition people are permitted to exhibit present bias (as in Laibson (1997) or O'Donoghue and Rabin (1999)), the estimated model fits the data even better with arguably reasonable levels of patience. Hence, DellaVigna et al. (2017) illustrates how, as economists take behavioral ideas to more sophisticated applications, it may be important to permit a variety of phenomena.[53]

[53]This point is not really novel, as it is consistent with the usual economics approach of analyzing an application while incorporating whatever factors are likely to be important in that application.

6.4 Consumer Choice and Market Prices

While the applications above test predictions using experimental or field data, other applications focus more on developing predictions in economic contexts and comparing those predictions to stylized empirical facts. A prominent example is the domain of consumer choice.

In their original work, Koszegi and Rabin (2006) apply their model of expectations-based loss aversion to this domain. Consider an adapted version of their model. Suppose a consumption bundle takes the form $\mathbf{x} \equiv (x^1, x^2)$ where $x^1 \in \{0, 1\}$ is consumption of some consumer good (say, shoes) and $x^2 \in \Re$ is money spent on that good. Hence, if a person buys a pair of shoes at price π, then $\mathbf{x} \equiv (1, -\pi)$, and if she does not buy shoes, then $\mathbf{x} \equiv (0, 0)$. Suppose the person's intrinsic utility is $u(\mathbf{x}) = x^1\gamma + x^2$, where $\gamma > 0$ reflects the intrinsic consumption utility from shoes. The person behaves according to the KR-model described above with a two-part-linear universal gain-loss utility function.

Now suppose that, before going to the store, the person faces price uncertainty: with probability p_l the price will be π_l, and with probability p_h the price will be $\pi_h > \pi_l$. This is a natural domain in which to apply PPE, wherein in advance the person formulates a plan for what to do at the store as a function of the price, but does not make an actual decision until she is at the store. Applying the structure from Section 5.4, state $s \in \{l, h\}$ generates a choice set $\mathcal{L}(s) \equiv \{(0, 0), (1, -\pi_s)\}$, and a strategy L^* must specify a planned choice for each of the two possible market prices.

To simplify this example, suppose that $\pi_l < \frac{1+\eta}{1+\eta\lambda}\gamma$. At this price, the person will buy no matter her beliefs.[54] Hence, $L^*(l) = (1, -\pi_l)$, and we merely must derive the $L^*(h)$ consistent with PPE. If $L^*(h) = (1, -\pi_h)$, then the ex ante distribution over outcomes is $G_{BUY} \equiv ((1, -\pi_l), p_l; (1, -\pi_h), p_h)$. If instead $L^*(h) = (0, 0)$, then the ex ante distribution over outcomes is $G_{NOT} \equiv ((1, -\pi_l), p_l; (0, 0), p_h)$. With a little work, one can derive

[54]Intuitively, this claim follows from deriving when the person would want to buy despite having the beliefs least conducive to buying—which is when the money spent is perceived as a loss and when obtaining shoes is perceived as a gain. See Koszegi and Rabin (2006) for the full argument.

that there exists $p' \in (0,1)$ such that for all $p_l < p'$ the PPE involves buying whenever $U(G_{BUY}|G_{BUY}) \geq U(G_{NOT}|G_{NOT})$,[55] which yields buying whenever

$$\pi_h \leq \gamma + \frac{2p_l\pi_l\Lambda}{1+p_l\Lambda}$$

Koszegi and Rabin (2006) emphasize several features of this example. First, note that $\bar{\pi}_G > \gamma$, meaning that the person is willing to pay more than her valuation for the good. This result follows from an attachment effect: because the person plans to buy when the price is low, she is attached to buying even when the price is high in the sense that not buying will feel like a loss. Hence, she has an extra reason to buy beyond the intrinsic value γ. Second, as p_l increases, $\bar{\pi}_G$ increases. Intuitively, as the probability of the low price increases, the attachment to buying grows stronger, thus increasing the willingness to pay when the price is high. Third, as π_l increases, $\bar{\pi}_G$ increases. Intuitively, buying at price π_h will feel like a loss relative to buying at price π_l, and this comparison effect counteracts the attachment effect. The larger is π_l, the smaller is this countervailing effect.

In this example, the person might pay more than her intrinsic valuation when the price is high, but this is only because of an attachment from the possibility of buying at a low price. Could it be that in expectation the person plans to pay more than her valuation? It turns out that the answer is yes. More generally, Heidhues and Koszegi (2014) consider the case of a profit-maximizing monopolist facing a KR consumer (following PPE), where the monopolist can commit to any distribution of prices. Consistent with the intuition in the simple example above, they show that the optimal pricing strategy involves a high regular price π_{reg} that is larger than the consumer's valuation, with some chance of a lower sales price drawn from the range $\pi \in [\frac{1+\eta}{1+\eta\lambda}\gamma, \pi']$ for some $\pi' \in (\frac{1+\eta}{1+\eta\lambda}\gamma, \pi_{reg})$. As in the example above, such a pricing strategy induces the consumer to buy no matter the realized price, and

[55] Formally, define $\bar{\pi}_N$ such that $U((0,0)|G_{NOT}) = U((1,-\bar{\pi}_N)|G_{NOT})$, and $\bar{\pi}_B$ such that $U((0,0)|G_{BUY}) = U((1,-\bar{\pi}_B)|G_{BUY})$. One can show $\bar{\pi}_N < \bar{\pi}_B$, so there are multiple PE for any $\pi \in [\bar{\pi}_N, \bar{\pi}_B]$. One can further show that $\bar{\pi}_G > \bar{\pi}_N$ and that there exists $p' \in (0,1)$ such that $\bar{\pi}_G < \bar{\pi}_B$ if and only if $p_l < p'$. Finally, $\bar{\pi}_N < \bar{\pi}_G < \bar{\pi}_B$ implies that the PPE involves buying whenever $\pi_h \leq \bar{\pi}_G$.

to pay an expected price larger than her valuation—thus yielding larger profits than merely setting a certain $\pi = \gamma$.

Heidhues and Koszegi (2014) describe how this optimal pricing pattern is consistent with some stylized facts that we observe in real-world markets. Specifically, for many goods, prices seem to alternate between higher regular prices and lower sales prices (and, indeed, are often labelled regular prices and sales prices). Moreover, just as predicted by the model, regular prices tend to be pretty stable over time, whereas sales prices tend to vary from sale to sale. However, Heidhues and Koszegi (2014) also point out two caveats. First, they highlight how the consumer would be better off if she could commit not to visit the store. Second, they highlight how when additional firms are added to the model, competition eliminates this result. Hence, the prediction of regular prices and sales depends on consumers being unable to commit not to shop, and depends on firms having some market power—both of which seem likely true for products sold at supermarkets.

Heidhues and Koszegi (2008) consider a different implication of expectations-based loss aversion for market prices: when firms compete in differentiated products, loss aversion tends to reduce and can even eliminate price variation in response to cost or demand changes, and can also generate "focal-price equilibria" in which firms that face different costs or demands all charge the same "focal" price. To see the intuition, consider a possible equilibrium in which all firms choose a focal price π^* regardless of any firm-specific cost shocks, and consider the incentive for one firm to deviate from this focal price. Because in such an equilibrium consumers expect to pay a price π^*, if the firm chooses a higher price, paying that price will be perceived as a loss relative to paying π^*, while if the firm chooses a lower price, paying that price will be considered a gain. Because losses loom larger than gains, it follows that the firm's equilibrium residual demand curve is kinked around the focal price π^*. This kink can serve to induce the firm to indeed choose price π^* for a range of cost realizations—thus supporting the equilibrium. Heidhues and Koszegi (2008) derive conditions under which focal-price equilibria exist, and conditions under which all equilibria are focal price equilibria.

They further argue that these results might help to explain stylized facts about real-world pricing—especially the stickiness of retail prices.

For more work on how expectations-based loss aversion might impact market prices, see Herweg and Mierendorff (2013), who demonstrate that, for consumers who face uncertainty about their own demand, expectations-based loss aversion generates a preference for flat-rate tariffs so as to avoid negative gain-loss utility that would arise if there were equilibrium uncertainty about how much money they would spend.

6.5 Mechanism Design: Auctions

Another natural application of expectations-based loss aversion is to auctions. When one bids in an auction, one generates expectations for the likelihood of winning the auction and for the amount that one needs to pay in the event of winning the auction, and it seems natural to think that the realized outcome might be judged relative to those expectations.

Lange and Ratan (2010) apply expectations-based loss aversion to first-price and second-price auctions. They use the KR approach to gain-loss utility from equation 7, except that they do not assume a universal gain-loss function. Rather, when one bids for a commodity, they permit a different degree of loss aversion (different λ's) for the commodity and for money (analogous to equation 4 in the domain of riskless choice). Finally, they effectively assume CPE is the relevant solution concept, motivated by the fact that in sealed-bid auctions expectations become fixed once the bid is submitted.

Lange and Ratan focus their analysis around the difference between commodity auctions and "induced-value auctions," where the latter is defined to be bidding for something that is readily converted into money—this might be induced-value tokens in laboratory experiments, or financial assets in the field. They show, for instance, that expectations-based loss aversion over money tends to decrease bids in first-price commodity auctions and increase bids in first-price induced-value auctions. Intuitively, losses on the money dimension occur when winning a commodity auction because that's when one must pay money (thus depressing bids), whereas they occur when losing an induced-value auction because that's when one does not obtain the money prize (thus increasing bids).

Expectations-based loss aversion over the commodity has more nuanced predictions: in both first-price and second-price auctions, it pushes bids higher for people with high valuations and pushes bids lower for people with low valuations. Intuitively, since obtaining vs. not obtaining the commodity is a binary outcome, gain-loss disutility is reduced if there is less uncertainty about which of these two outcomes will occur—i.e., if the probability of winning the auction is pushed closer to zero or one. Hence, if a person has a high valuation such that she is likely to win the auction, loss aversion over the commodity has an upward impact on her bid so as to push the likelihood of winning closer to one. Analogously, if a person has a low valuation such that she is unlikely to win the auction, loss aversion over the commodity has a downward impact on her bid so as to push the likelihood of winning closer to zero.

Lange and Ratan emphasize several implications of their analysis. Perhaps most important, their analysis highlights how the existence of loss aversion on both the money and commodity dimensions can lead to very different bidding behavior in commodity auctions vs. induced-value auctions. Hence, one must be careful in extrapolating from laboratory induced-value auctions to predict how people behave in field commodity auctions.

Banerji and Gupta (2014) build on the theoretical analysis from Lange and Ratan (2010) to develop a test of expectations-based loss aversion. Specifically, they consider a special case of a second-price auction: the Becker-DeGroot-Marschak elicitation procedure (see Becker et al. (1964)), which is essentially a second-price auction where the other bid is determined randomly. As in Lange and Ratan, they consider commodity auctions vs. induced-value auctions, where they apply the KR approach to gain-loss utility from equation 7. Unlike Lange and Ratan, however, they do assume a universal gain-loss function. Finally, they assume CPE is the relevant solution concept.

Banerji and Gupta consider BDM auctions in which the distribution of the second bid is uniform on the interval $[0, K]$. Their focus is on how the bid distribution changes as a function of K. Standard agents should merely bid their value, and thus K is irrelevant (except for censoring of people with a value larger than K). In induced-value auctions, expectations-based loss aversion does not change this prediction because in such auctions gain-loss disutility is minimized when one bids one's value. However, in commodity auctions an increase in K leads a loss-averse consumer to reduce her bid (as long as her initial bid was smaller than K).

Banerji and Gupta test these predictions experimentally by conducting BDM auctions for a candy bar and for an induced-value token. In each case, they examined bid distributions when $K = 150$ vs. when $K = 200$. For the commodity (candy bar) auction, they indeed find that the distribution of bids for $K = 150$ first-order stochastically dominates the distribution of bids for $K = 200$. For the induced-value auction, the two bid distributions are quite similar, and moreover the majority of subjects do indeed bid their value. On net, these results are quite consistent with the model of expectations-based loss aversion.

6.6 Mechanism Design: Optimal Contracting

Herweg et al. (2010) consider the implications of expectations-based loss aversion for incentive contracts. They are motivated by a puzzle in the agency literature. Standard agency models predict that optimal contracts should take advantage of all available information. As one prominent example, if there is an observable performance measure, and if different realizations of that performance measure yield different inferences about the amount of effort that an agent exerted, then the optimal contract should pay the agent different amounts for each realization of the performance measure. However, real-world contracts seem not to take advantage of all available information, and indeed they frequently take a simple binary form in which the agent receives a base amount plus a bonus if the performance measure exceeds some critical level (see Herweg et al for a discussion of this puzzle).

Herweg et al suggest that expectations-based loss aversion might be the source of such contracts. Specifically, they assume an agent who behaves according to CPE with linear utility for money and a two-part linear gain-loss utility function.[56] For such an agent, the optimal contract cannot pay a uniform wage no matter the outcome because then there would be no incentive to exert effort (just as for a standard agent). However, the more variation there is in the set of possible wages that might be paid, the larger is the gain-loss disutility. Herweg et al identify conditions under which it is in fact optimal to limit the wage to two possible values, which can be interpreted as a base wage and a bonus payment that is paid if the performance measure is high enough.

It is worth highlighting that this result depends on assuming the KR approach to gain-loss utility. Under the KR approach, making additional use of information in the performance measure means introducing more ex ante uncertainty into the agent's final wage, which in turn means the principal must pay a higher expected wage in order to induce the agent to sign the contract. Under the DA approach, creating spread in wages that doesn't change the expected wage and that doesn't change which wages are above or below the expected wage has no impact on gain-loss utility.

7 Reference-Dependent "News" Utility

7.1 Overview and Some History

Up to now, our entire focus has been on contemporaneous gain-loss utility—that is, gain-loss utility experienced in period t as a result of comparing realized period-t consumption to a referent. As we've seen, period-t contemporaneous gain-loss utility matters prior to period t because a person takes it into account when making choices prior to period t that impact period-t consumption.

This focus on consumption yielding utility when that consumption occurs is the norm for

[56] It is unclear whether CPE is an appropriate solution concept for agency applications. In particular, it would be appropriate in situations where the performance measure is realized well after the effort is exerted.

most economics. However, there is also research on various forms of "anticipatory" utility wherein a person experiences utility now from thinking about (anticipating) consumption in some future period.[57] While most anticipatory-utility models lie outside the domain of reference-dependent preferences, Koszegi and Rabin (2009) develop a model of reference-dependent "news" utility. Specifically, when a person receives information ("news") that changes her beliefs about future consumption, she experiences gain-loss utility now derived from comparing these new beliefs to her prior beliefs (her referent).

7.2 News Utility

Consider news utility experienced in period $\tau \leq t$ related to consumption in period t, and suppose beliefs about period-t consumption take the form $L^t \equiv (x_1^t, p_1; x_2^t, p_2; ...; x_N^t, p_N)$.[58] Let $L_{\tau-1}^t$ denote the person's prior beliefs entering period τ, and suppose the person receives news in period τ that leads her to change her beliefs to L_τ^t. Koszegi and Rabin (2009) assume that the news utility associated with this change in beliefs is

$$N(L_\tau^t | L_{\tau-1}^t) = \gamma_{\tau,t} \int_0^1 \mu(\bar{x}_{L_\tau^t}(p) - \bar{x}_{L_{\tau-1}^t}(p))dp$$

where μ is a gain-loss utility function and, for each $p \in (0,1)$, $\bar{x}_L(p)$ is such that $\Pr_L(x \leq \bar{x}_L(p)) \geq p$ and $\Pr_L(x \leq \bar{x}) < p$ for any $\bar{x} < \bar{x}_L(p)$. $\gamma_{\tau,t} \in [0,1]$ reflects the weight that the news gets relative to the actual consumption experience in period t (that gets weight 1). Koszegi and Rabin suggest that $\tau < \tau'$ implies $\gamma_{\tau,t} \leq \gamma_{\tau',t}$, so that news gets (weakly) less weight the further in the future is the actual consumption.

This formulation of news utility implies $N(L^t | L^t) = 0$, and thus it captures that no news means no news utility.[59] Example 5 (adapted from Section III in Koszegi and Rabin (2009))

---

[57] See, for instance, Loewenstein (1987), Caplin and Leahy (2001), Brunnermeier and Parker (2005), and Koszegi (2006).

[58] To simplify the exposition, this treatment considers the case where consumption is a scalar. Koszegi and Rabin (2009) develop their model for the case where consumption is a vector.

[59] Given the motivation from gain-loss utility, one might have thought it natural to assume $N(L_\tau^t | L_{\tau-1}^t) =$

illustrates how this formulation works when there is actual news.

Example 5: Suppose $L^t_{\tau-1} \equiv (\underline{x}, \frac{1}{2}; \bar{x}, \frac{1}{2})$ with $\bar{x} > \underline{x}$, and suppose further that the person receives a signal in period t that outcome $x_n \in \{\underline{x}, \bar{x}\}$ now has probability $q \geq 1/2$.

If the signal is good news, new beliefs are $L^t_\tau \equiv (\underline{x}, 1-q; \bar{x}, q)$ and thus

$$\begin{aligned} N(L^t_\tau | L_{\tau-1}) &= \gamma_{\tau,t} \left[\int_0^{1-q} \mu(\underline{x} - \underline{x}) dp + \int_{1-q}^{1/2} \mu(\bar{x} - \underline{x}) dp + \int_{1/2}^{1} \mu(\bar{x} - \bar{x}) dp \right] \\ &= \gamma_{\tau,t} \left[(1/2 - (1-q)) \, \mu(\bar{x} - \underline{x}) \right] = \gamma_{\tau,t} \, (q - 1/2) \, \eta(\bar{x} - \underline{x}). \end{aligned}$$

If the signal is bad news, new beliefs are $L^t_\tau \equiv (\underline{x}, q; \bar{x}, 1-q)$ and thus

$$\begin{aligned} N(L^t_\tau | L_{\tau-1}) &= \gamma_{\tau,t} \left[\int_0^{1/2} \mu(\underline{x} - \underline{x}) dp + \int_{1/2}^{q} \mu(\underline{x} - \bar{x}) dp + \int_{q}^{1} \mu(\bar{x} - \bar{x}) dp \right] \\ &= \gamma_{\tau,t} \left[(q - 1/2) \, \mu(\underline{x} - \bar{x}) \right] = -\gamma_{\tau,t} \, (q - 1/2) \, \eta\lambda(\bar{x} - \underline{x}). \end{aligned}$$

In Example 5, if the person receives good news about period-t consumption, she experiences positive news utility, whereas if she receives bad news, she receives negative news utility. The magnitude of the news utility depends on a combination of how informative the news is (how much larger q is than $1/2$) and how much the person cares in period τ about period-t consumption (how much larger $\gamma_{\tau,t}$ is than 0). In the limit where the news is uninformative (when $q = 1/2$), there is no news utility—illustrating the more general point that $N(L|L) = 0$. In the other limit where the news fully resolves the uncertainty (when

$\gamma_{\tau,t} E_{L^t_\tau} v(x^t | L^t_{\tau-1})$ where v is as in equation (7). However, that formulation would yield the unnatural implication that $N(L^t | L^t) < 0$.

$q = 1$), the news utility is equivalent to what standard contemporaneous gain-loss utility would be, except that it is discounted by $\gamma_{\tau,t}$. This latter point highlights that, with this formulation of news utility, contemporaneous gain-loss utility is a merely a special case of news utility where news is received in period t about the realization of period-t consumption.

7.3 News Utility and Decision Making

Because news utility is inherently an intertemporal concept, it must be incorporated into a model of intertemporal utility. A natural approach is to assume that news utility experienced in period τ is merely an additive term in the period-τ instantaneous utility function.

To illustrate, consider a model with consumption in periods 1 and 2. As a function of period-1 consumption (x^1), prior beliefs about period-1 and period-2 consumption (L_0^1 and L_0^2), and period-1 beliefs about period-2 consumption (L_1^2), period-1 intertemporal utility might be

$$\left[u(x^1) + v(x^1|L_0^1) + N(L_1^2|L_0^2)\right] + \delta\left[E_{L_1^2}(u(x^2) + v(x^2|L_1^2))\right] \tag{11}$$

where $\delta \leq 1$ is a standard time-discounting factor.

Much as for the expectations-based models discussed in Section 5.2, this model in principle permits some flexibility in the source of beliefs. Equation (11) already incorporates an assumption that period-1 beliefs about period-2 consumption (L_1^2) become the referent in period 2. A second natural assumption—though the model would be coherent without it—is that those period-1 beliefs should be formed from the prior beliefs and the news using Bayes rule. Finally—and most analogous to static expectations-based models—the prior beliefs could in principle be exogenous or endogenous. Analogous to our discussion in Section 5.3, the former might reflect that the person is surprised in period 1 with this choice situation, whereas the latter might reflect that the person is informed of the choice situation in advance and formulates a plan. Indeed, for the latter case, Koszegi and Rabin (2009) develop a dynamic version of PPE.[60]

In their online appendix, Koszegi and Rabin (2009) highlight how their static formulations of PPE and CPE (defined in Section 5.3 above) can be thought of as special cases of their dynamic formulation of PPE where (i) all outcomes occur and all uncertainty is resolved in some period t, and (ii) the person formulates a plan (for PPE) or commits to a choice (for CPE) in some period $\tau < t$ with $\gamma_{\tau,t} = 0$. The key issue is that the act of formulating a plan (for PPE) or committing to a choice (for CPE) creates news—beliefs in period τ shift from prior beliefs to the beliefs associated with the plan or choice. This shift in beliefs is ignored in the static versions of PPE and CPE. If $\gamma_{\tau,t} = 0$, then that shift in beliefs in period τ is indeed irrelevant; otherwise, however, the plan or choice must be made while also taking into account the impact on news utility experienced in period τ. The latter case further highlights how one must be careful when applying the static version of CPE or PPE (as discussed in Section 5.4).

7.4 Applications of News Utility

Thus far, there have been fewer applications of reference-dependent news utility than there have been applications of expectations-based loss aversion. But some initial work is promising.

Koszegi and Rabin (2009) emphasize that their model has two implications for preferences over the resolution of uncertainty. First, because $\gamma_{\tau,t} < \gamma_{\tau',t}$ for $\tau < \tau'$, the model implies that people like news earlier rather than later. Second, and perhaps more important, because loss aversion implies that, in expectation, each piece of news yields negative news utility, the model implies that people do not like piecemeal news—i.e., they prefer less-frequent, more-informative signals to more-frequent, less-informative signals. Falk and Zimmermann (2017)

[60]As in most models of anticipatory utility, news utility causes intertemporal preferences to be time-inconsistent. In the two-period example, for instance, whereas the anticipatory utility $N(L_1^2|L_0^2)$ impacts the person's period-1 intertemporal preferences, when period 2 arrives $N(L_1^2|L_0^2)$ is no longer relevant. The dynamic version of PPE assumes that people must follow a consistent plan, but one could imagine alternative assumptions.

find empirical support for both implications.

Pagel (2018) develops the importance of the second implication by embedding news utility in a stochastic life-cycle model of portfolio choice.[61] Pagel demonstrates that, in order to receive news about one's portfolio via less-frequent, more-informative signals, people choose to ignore their portfolios for extended periods of time. Moreover, because people would still like to have their portfolios actively rebalanced even while they are ignoring them, people are willing to pay for a portfolio manager. Finally, because portfolio accounts are monitored infrequently while checking accounts (used for consumption) are monitored frequently, the model implies different marginal propensities to consume out of these two accounts.

Koszegi and Rabin (2009) also develop their model's implications for consumption-savings decisions using a stylized two-period model. Pagel (2017a) expands on this approach by embedding news utility in a life-cycle consumption-saving model with a single asset. She finds that news utility yields excess smoothness in consumption, a life-cycle consumption profile that is hump-shaped, and a drop in consumption at retirement. In addition, Pagel (2015) investigates the implications of news utility for asset pricing, and argues that the model can explain the historical equity premium using reasonable parameter values.

8 Probability Weighting

8.1 Overview and Some History

While our primary focus in this chapter is reference-dependent preferences, we take a short detour before we conclude to discuss probability weighting, and in particular to highlight some connections between probability weighting and expectations-based loss aversion.

Probability weighting was incorporated into prospect theory to accommodate the Al-

[61] In finance applications, uncertainty in current beliefs can be strongly correlated with uncertainty in prior beliefs. To account for this correlation, Pagel uses a different formulation of news utility that is qualitatively similar to the Koszegi and Rabin (2009) formulation but quantitatively more tractable. For a detailed discussion of this approach, see Pagel (2017b).

lais paradoxes and other EU deviations not naturally accommodated by loss aversion. The insights derived from probability weighting have been particularly valuable for analyzing decisionmaking with relatively low probability events. Insurance choice, gambling, and financial investment present natural environments for application.

In their presentation of probability weighting, Kahneman and Tversky (1979) recognize a potential problem: nonlinear probability weighting can lead to violations of dominance. They address this problem by assuming that dominated options are eliminated from consideration at the editing stage that is described in Section 3.4. Shortly thereafter, a more elegant solution to the problem of dominance was proposed by Quiggin (1982): the model of rank-dependent probability weighting (RDPW).[62] This model was quickly adopted, and later a variant was incorporated into cumulative prospect theory (Tversky and Kahneman 1992).

In recent years, RDPW has received more attention in the behavioral-economics literature for several reasons. First, it turns out that some models of expectations-based loss aversion—most notably, CPE—are indistinguishable from RDPW. Second, there has been a recognition that rank dependence has its own substantive predictions, and some initial tests of those predictions have not found support.

8.2 Simple Nonlinear Probability Weighting

Under EU, the decision weight applied to outcome n is the probability p_n of that outcome occurring. The basic idea of probability weighting is that equation 1 is replaced by

$$U(L) \equiv \sum_{n=1}^{N} \omega_n u(x_n)$$

[62]When RDPW is applied in a model with a utility function defined over final wealth (as in EU), it is often labelled "rank-dependent expected utility" (RDEU). We use the label "rank-dependent probability weighting" to refer to the weighting component of the model, recognizing that it could be applied as in RDEU, but it could also be applied in models with reference-dependent preferences. To isolate the impact of probability weighting, the majority of this section focuses on models of probability weighting combined with a utility function defined over final wealth states.

where ω_n is a weight that might differ from p_n. Original formulations assume that ω_n is simply a transformation of p_n. In other words, there exists a probability weighting function $\pi(p)$ such that $\omega_n = \pi(p_n)$.

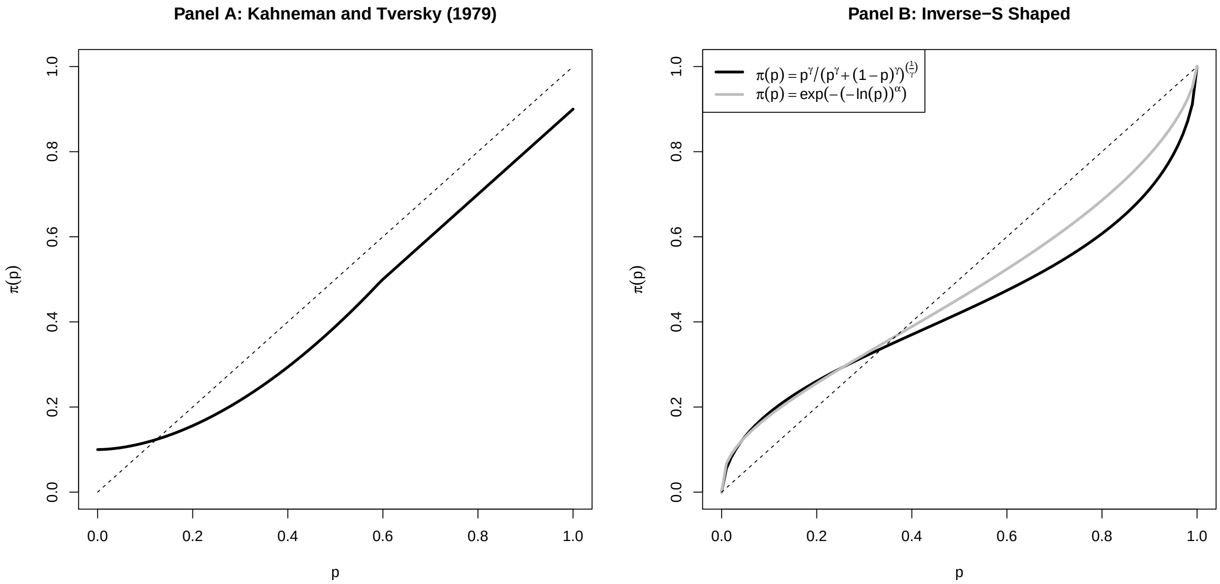
The idea of probability weighting predates prospect theory (see Section 3.1); the real contribution of Kahneman and Tversky (1979) to the literature on probability weighting was to propose—based on their experimental evidence—several features for the function $\pi(p)$. These features include:

- Overweighting ($\pi(p) > p$) of small probabilities and underweighting ($\pi(p) < p$) of large probabilities.
- Subcertainty: $\pi(p) + \pi(1-p) < \pi(1) \equiv 1$ (which is needed to accommodate the Allais common-consequence paradox).
- Subproportionality: For any $p, q, r \in (0,1)$, $\pi(pq)/\pi(p) < \pi(pqr)/\pi(pr)$ (which is needed to accommodate the Allais common-ratio paradox).

Kahneman and Tversky (1979) did not provide a parameterized functional form, but rather provided a figure like that depicted in Panel A of Figure 6 that satisfies these properties. Note that this function is discontinuous at $p = 0$ and $p = 1$, suggesting that certain outcomes are treated in a discontinuously different way from probabilistic outcomes. The literature quickly abandoned this discontinuity and instead focused on inverse-S-shaped weighting functions as depicted in Panel B of Figure 6.[63]

A problem with simple probability weighting is that it generates violations of dominance. Suppose π satisfies subcertainty and thus $\pi(p) + \pi(1-p) < 1$. If so, then there exists $\varepsilon' > 0$ such that $(x, 1) \succ (x+\varepsilon, p; x+2\varepsilon, 1-p)$ for all $\varepsilon < \varepsilon'$ because the sum of the weights applied to $x+\varepsilon$ and $x+2\varepsilon$ in the latter option is strictly smaller than the weight applied to x in the former option. Numerous additional examples can easily be generated. As another example,

[63]For more details on possible functional forms, see Table 1 and Figure 2 in Barseghyan et al. (2018).



Note: Panel A presents a probability weighting function of the form suggested by Kahneman and Tversky (1979). Panel B presents probability weighting functions corresponding to the functional forms from Tversky and Kahneman (1992) with $\gamma = 0.61$ (solid black line) and Prelec (1998) with $\alpha = 0.65$ (solid gray line). In both panels, dotted line corresponds to no probability weighting.

Figure 6: Some Probability Weighting Functions

if $\pi(1/n) > 1/n$, then then there exists $\varepsilon' > 0$ such that $(x-\varepsilon, 1/n; x-2\varepsilon, 1/n; ...x-n\varepsilon, 1/n) \succ (x, 1)$ for all $\varepsilon < \varepsilon'$.

Kahneman and Tversky (1979) recognize this problem, and address it by assuming that dominated options are eliminated from consideration at the editing stage. Shortly thereafter, Quiggin (1982) proposed a more elegant solution.

8.3 Rank-Dependent Probability Weighting

In order to avoid violations of dominance, Quiggin (1982) proposed that instead of the weight ω_n being a simple transformation of the probability p_n, it is a transformation of the cumulative probability of obtaining at most x_n. That is, when evaluating a prospect $L \equiv (x_1, p_1; x_2, p_2; ...; x_N, p_N)$, the decision maker first orders the outcomes by rank such that $x_1 \leq x_2 \leq .... \leq x_N$. The decision maker then evaluates lottery L according to

$$U(L) = \sum_{n=1}^{N} \left(\pi \left(\sum_{m=1}^{n} p_m \right) - \pi \left(\sum_{m=1}^{n-1} p_m \right) \right) u(x_n)$$

where $\pi(\cdot)$ is a weighting function that is increasing and has $\pi(0) = 0$ and $\pi(1) = 1$. In other words, under this rank-dependent probability weighting (RDPW), the weight on outcome x_n is

$$\omega_n = \pi \left(\sum_{m=1}^{n} p_m \right) - \pi \left(\sum_{m=1}^{n-1} p_m \right).$$

Several features of RDPW are worth noting. First, if $\pi(p) = p$ then $\omega_n = p_n$. Hence, RDPW deviates from standard weighting only if the function $\pi(\cdot)$ is nonlinear. Second, for any lottery that might be evaluated, the weights always sum to one. This feature eliminates the violations of dominance described in Section 8.2.

Third, the implications of RDPW depend on the functional form for $\pi(\cdot)$. Because in many examples simple probability weighting and RDPW coincide, the literature has used the same functional forms as depicted in Figure 6. Indeed, much of the research that has generated those functional forms does so in reference to certainty equivalents for gambles of the form $(x, p; 0, 1-p)$. If, as is frequently done, one further assumes $u(0) = 0$—e.g., when one applies a simple gain-loss utility function—both simple probability weighting and RDPW imply that the certainty equivalent $\bar{z}$ satisfies $u(\bar{z}) = \pi(p)u(x)$.

When applying these functional forms to gambles with more outcomes, RDPW generates a different psychology of decision making relative to simple probability weighting. Under simple probability weighting, an outcome is overweighted if the likelihood of it occurring is small enough. Under RDPW, in contrast, extreme outcomes are overweighted. To highlight this point, consider two examples:

Example 6: Consider a lottery $(x_A, .10; x_B, .45; x_C, .45)$, and consider one of the functional forms in Figure 6 panel B.

- Under simple probability weighting, the weight on outcome x_A is $\pi(.10)$, regardless of its rank, and thus would be overweighted.
- Under RDPW, if outcome x_A is the worst outcome, then it gets weight $\pi(.10)$ and is thus overweighted. If outcome x_A is the best, then it gets weight $1 - \pi(.90)$ and thus would also be overweighted. However, if outcome x_A is the middle outcome, then it gets weight $\pi(.55) - \pi(.45)$, and thus would be underweighted.[64]

Example 7: Consider a lottery with many possible outcomes, each of which occurs with a small probability, and let $F(x)$ be the cumulative probability that the lottery yields at most x. Roughly speaking, if outcome x has $\frac{d\pi(F(x))}{dp} > 1$ then the outcome will be overweighted and if outcome x has $\frac{d\pi(F(x))}{dp} < 1$ then the outcome will be underweighted. Hence, for an inverse-S-shaped probability weighting function, the smallest and the largest x's are overweighted.

When Tversky and Kahneman (1992) introduced "cumulative prospect theory" (CPT) as an improved version of prospect theory, they incorporated a variant of RDPW in which probability weighting is applied differentially to gains and losses. Specifically, if a decision maker with reference point r evaluates a prospect L where $x_1 \leq .... \leq x_{\bar{n}} < r \leq x_{\bar{n}+1} \leq ... \leq x_N$, the decision weights are:

$$\omega_n = \begin{cases} \pi^- \left(\sum_{m=1}^{n} p_m\right) - \pi^1 \left(\sum_{m=1}^{n-1} p_m\right) & \text{if } n \leq \bar{n} \\ \\ \pi^+ \left(\sum_{m=n}^{N} p_m\right) - \pi^+ \left(\sum_{m=m+1}^{N} p_m\right) & \text{if } n > \bar{n} \end{cases}$$

[64]The common functional forms all have $\pi(.10) > .10$, $\pi(.90) < .90$, and, since the slope is less than one in this range, $\pi(.55) - \pi(.45) < .10$.

In this formulation, π^+ is the probability weighting function applied to gains, and π^- is the probability weighting function applied to losses—where both would have the inverse-S-shaped functional form.

RDPW—often under the label RDEU—has been prominent in the decision-theory literature (see, e.g., Abdellaoui (2002)). In addition, there is a large experimental literature that estimates the CPT model (see, e.g., Bruhin et al. (2010)). In applied work, RDPW (often in the CPT form) has perhaps had its largest impact in the domains of finance and gambling, where researchers have developed the impact of the overweighting of extreme outcomes (see in particular Barberis (2013) and Chapter X of this volume).

8.4 RDPW vs. CPE

RDPW delivers deviations from expected utility based upon distortions of cumulative probabilities. The CPE model described in Section 5.3.2 delivers deviations from expected utility based on sensations of gain and loss. These different psychologies have a common force at their core: both rely on the ranking of outcomes. In RDPW, the rank of an outcome determines the cumulative probability of weakly inferior and superior outcomes and the corresponding distortion. In CPE, the rank of an outcome determines the number of comparisons to other outcomes that will be coded as gains and losses. Recent research has identified some tight connections between these models.

Masatlioglu and Raymond (2016) take a decision-theory approach to ask when different models have different implications for behavior. They discover that, as long as $\Lambda \leq 1$ (which rules out violations of dominance), CPE is equivalent to a special case of RDPW. To illustrate, consider a binary lottery $L \equiv (x, p; y, 1-p)$ with $x < y$. Applying equation 8,

$$U(L|L) = [p + p(1-p)\Lambda]\, u(x) + [1 - [p + p(1-p)\Lambda]]\, u(y).$$

This is equivalent to RDEU with weighting function $\pi(p) = p + p(1-p)\Lambda$ (note that $\Lambda \leq 1$ guarantees that $\pi(p)$ is increasing for all p).[65]

Barseghyan et al. (2013b) and Barseghyan et al. (2018) make the same point from the direction of identification. Specifically, they consider a model that incorporates both RDPW and CPE. For the case of binary lotteries of the form above, they show the lottery is evaluated according to

$$U(L|L) = [\pi(p) + \pi(p)(1-\pi(p))\Lambda]\, u(x) + [1 - [\pi(p) + \pi(p)(1-\pi(p))\Lambda]]\, u(y).$$

Hence, even if one were able to identify the combined weight $\Omega(p) \equiv \pi(p) + \pi(p)(1-\pi(p))\Lambda$, it would not be possible to break this apart to estimate Λ and $\pi(p)$—that is, Λ from CPE and $\pi(\cdot)$ from RDPW are not separately identified.[66]

Given this identification issue, one approach is to focus on estimating the combined weight $\Omega(p)$. Barseghyan et al. (2013a) do so using data on insurance-deductible choices for home and auto insurance. They assume that households' subjective beliefs for p correspond to an econometrician's best guess based on household characteristics and past claims data, and they take three non-parametric approaches to estimating $\Omega(p)$. All three approaches yield a function that looks qualitatively like the probability weighting from Kahneman and Tversky (1979) (depicted in Panel A of Figure 6). Because such a functional form is in fact inconsistent with the CPE implied $\pi(p) = p + p(1-p)\Lambda$, they conclude that there must be some RDPW driving behavior. Given the identification issue, however, they cannot assess whether CPE might also be playing a role.

Masatlioglu and Raymond (2016) further show that the link between RDPW and DA is less tight. For binary lotteries, DA is equivalent to CPE and thus also exhibits the equiv-

[65]More generally, Masatlioglu and Raymond prove that, for any choice set of lotteries with discrete outcomes, CPE with $\Lambda \leq 1$ is equivalent to RDEU with $\pi(p) = p + p(1-p)\Lambda$.

[66]More generally, Barseghyan et al. (2018) prove (in their appendix) that, for any choice set of lotteries with discrete outcomes, the combination of CPE with $\Lambda \leq 1$ and RDPW with probability weighting function $\pi(p)$ is equivalent to an RDEU model with probability weighting function $\Omega(p) \equiv \pi(p) + \pi(p)(1-\pi(p))\Lambda$.

alences above.[67] For lotteries with more than two outcomes, DA yields distinct predictions from RDPW and CPE. That said, given the similar intuitions behind these models, in practice it is a bit unclear how different the predictions are likely to be.

8.5 Tests of Rank Dependence

With the realization that rank dependence has real implications in terms of overweighting extreme outcomes, along with the recognition of the connection between RDPW and CPE, direct tests of rank dependence take on new importance. Indeed, one can even think of a test of rank dependence in lotteries with more than two outcomes also being a test of CPE.

In fact, rank dependence has received limited attention as a feature to be tested. Two exceptions are Wu (1994) and Wakker et al. (1994), which examine whether a preference between two lotteries is maintained when common outcomes are replaced without changing outcome ranks, a critical axiom of rank-dependent models termed "comonotonic independence." Their findings show that comonotonic independence and non-comonotonic independence are violated at similar rates in experimental samples, a prima-facie unfavorable result for rank-dependent models. The conclusions from comparing violation rates, however, are not particularly strong as random choice and near indifference between lotteries could easily rationalize the observed data patterns. It is perhaps for this reason that these early tests are not well known.

Recently, Bernheim and Sprenger (2017) provide a direct test of rank dependence based on three-outcome prospects. Consider a three-outcome prospect with $x_1 \neq x_2 > x_3$ and $p_1 = p_2 = p_3 = 1/3$. Bernheim and Sprenger's experiment increases x_2 to $x_2 + m$ and asks subjects to report the "equalizing reduction" in x_3 to $x_3 - k$ that makes the subject indifferent. The equalizing reduction concerns trade-offs between x_2 and x_3. If probability distortions are rank-dependent, these trade-offs, and hence k, should be sensitive to whether $x_1 > x_2$ or

[67] For two-outcome gambles, the Gul (1991) variant of DA is also a special case of RDEU (see Abdellaoui and Bleichrodt (2007) and Barseghyan et al. (2013a)).

$x_1 < x_2$. Bernheim and Sprenger also estimate CPT probability weighting using standard prospect theory assessment tasks. As such they can predict precisely how much k should change as ranks change. Though k is predicted to change by 20-40% across conditions, the data are tightly organized around zero change and, hence, zero rank dependence in probability distortions at both the aggregate and individual level.

Clearly more work is needed to more fully assess the assumption of rank dependence. A closely related question is to what extent the assumption of rank dependence is required in applications of probability weighting. Indeed, in applications where the small-probability events are also the extreme events, models with simple probability weighting as in Section 8.2 would likely yield similar results. But since that approach suffers from violations of dominance, it might be fruitful to develop alternative approaches that also are able to capture the key features of simple probability weighting.

9 Discussion

We have detailed in this chapter how the literature on reference-dependent preferences has progressed a long ways, especially in the past two decades, and has yielded valuable insights for interpreting a number of laboratory and field behaviors. That said, there is still much room for improvement. We conclude by discussing several unresolved questions.

Welfare implications: This chapter has taken a positive approach to reference-dependent preferences, focusing on whether models of reference-dependent preferences are consistent with observed behavior. This focus mirrors that in the literature, as there is relatively little discussion of the welfare implications of reference-dependent preferences.

When one takes a normative approach to reference-dependent preferences, a number of issues arise. Perhaps first and foremost is the question of whether gain-loss utility should be given normative weight—i.e., whether we should assume that the same preferences that rationalize behavior should also be used for welfare analysis. If gain-loss utility represents true feelings of "pleasure" and "pain," it seems natural to assume that it should be given normative weight. However, the fact that relatively innocuous changes in experimental procedures seem able to induce people to re-code gains and losses in different ways would seem to cast doubt on this conclusion.

A related normative question is whether people correctly forecast future feelings of gains and losses. In dynamic situations in which the referent adjusts over time, projection bias (Loewenstein et al. (2003)) might lead people to underappreciate those changes. For instance, if people adapt quickly to changes in one's endowment, projection bias might lead people to act as if their gain-loss utility is much larger than it actually is. More simple than projection bias, it could even be that people merely under- or over-estimate their degree of loss aversion.

As the literature on reference-dependent preferences continues to develop, it will be important to tackle these and other normative issues.

Portability and stability of preferences: An important virtue of a model of decision making is that it be portable (Rabin (2013)). In other words, it is helpful if the model is written in a way that any researcher would know how to apply the model in any application. At first glance, models of reference-dependent preferences seem highly portable. However, once one recognizes the importance of assumptions about bracketing and about the referent (more on the latter below), and the lack of systematic guidance for how to make such assumptions, the model becomes somewhat less portable—and thus requires more judgment from the modeler.

Even if a model is portable, a related issue is whether an individual's preferences within that model can be expected to be stable across contexts. For instance, even if we were confident that a person were an EU maximizer, and we quantitatively estimated the person's risk preferences from one context (say, insurance choices), it would be nice if those estimated risk preferences could be used to accurately predict behavior in another context (say, asset-investment choices). In fact, there is relatively little research on the stability of risk preferences (for an overview, see Section 7.1 in Barseghyan et al. (2018)). Chapter X of this handbook on structural behavioral economics provides additional discussion of preference stability in terms of a set of estimated parameters being able to meaningfully predict behavior in related out-of-sample environments.

Of course, given that all models in economics are incomplete—they are simplifications of reality—it would be surprising if there were complete quantitative consistency across contexts. That said, as the literature on reference-dependent preferences continues to develop, it will be useful to investigate how much we can use research in one environment to predict behavior in another.

Diminishing sensitivity: Most applications of reference-dependent preferences focus entirely on loss aversion, and ignore the possibility of diminishing sensitivity (typically by assuming a two-part linear gain-loss utility function).

For many applications, diminishing sensitivity would seem to have a relatively minor impact on predictions. Consider, for instance, the real-effort task in Abeler et al. (2011) discussed in Section 6.2. Under a two-part-linear value function, subjects who choose $e < 3$ or $e > 7$ (in either treatment) should not be influenced at all by whether $F = 3$ or $F = 7$. All the action is for those who choose $e \in [3, 7]$. With diminishing sensitivity, in contrast, those with $e < 3$ ($e > 7$) should work less in the $F = 7$ ($F = 3$) treatment because they are further in the loss (gain) domain and thus, with diminishing sensitivity, face a smaller marginal benefit from effort. Hence, diminishing sensitivity has an impact in this domain. At the same time, however, the main effects of bunching at F and a strong incentive to change effort for those with $e \in [3, 7]$ are unchanged, and we suspect these effects would swamp the more minor diminishing-sensitivity effects.

For some applications, however, diminishing sensitivity plays a critical role in delivering phenomena. As noted in Section 4.4, the disposition effect of agents wishing to realize capital gains and avoid capital losses has a natural diminishing-sensitivity interpretation. Another application related to diminishing sensitivity is Bowman et al. (1999), who consider a model of how people react to income shocks. If an income shock implies that a person will need to take losses in the form of consuming below her reference point, diminishing sensitivity implies that she would prefer to concentrate those losses in one period rather than spread them over time. Moreover, diminishing sensitivity further implies that probabilistic losses are preferred to certain losses. Hence, when people experience a negative income shock, they resist taking any losses now because they would be certain, whereas future losses are probabilistic (there might be a good income shock in the future). For positive shocks, in contrast diminishing sensitivity, like intrinsic utility, implies a desire to smooth gains. The combined prediction, then, is that we should see asymmetric responses to negative vs. positive income shocks. Bowman et al indeed point to evidence that supports this prediction of diminishing sensitivity.

These applications hint at the type of situation in which diminishing sensitivity is likely to be important—situations in which people face a choice between one large loss and multiple small losses, possibly across time, possibly across states, or possibly across different dimensions of consumption. More generally, the literature still needs to develop a better sense of when diminishing sensitivity is important.

The referent: Finally, and perhaps most important, is the issue of what the referent is. The literature on reference-dependent preferences has suggested a number of possibilities: the status quo (e.g., prior wealth), past experience, focal outcomes, aspirations, and expectations. As we've discussed, the lack of clear guidance on what the referent should be creates a major degree of freedom in the model, one that has been taken advantage of in applications.

Koszegi and Rabin (2006) developed their model of expectations-based loss aversion in large part to impose some discipline in how people apply reference-dependent preferences. In assuming that the referent is expectations and closing the model via rational expectations, they removed much of the model's flexibility (especially in PPE and CPE). Unfortunately, these more disciplined versions of the KR model have found mixed success in interpreting observed behaviors. For instance, in laboratory experiments on exchange behavior and effort provision, a reasonable reading of the literature would be that the force of endogenous expectations cannot be reliably observed. A possible interpretation is that other aspects of the decision environment, including plausible other salient referents, hamper the establishment of forward-looking expectations. While it is possible to rationalize behaviors such as the endowment effect by appealing to non-endogenous expectations, such as with an appeal to "surprise situations," once we open the door to flexibility in how expectations get established, we reintroduce a certain degree of flexibility into the model. Though principled use of such degrees of freedom may be warranted, we must proceed with caution.

In evaluating the contribution of expectations-based reference points, it is critical to take a historical perspective. In the years following the publication of Kahneman and Tversky's seminal 1979 paper, a number of shortfalls were noted (principal among them dominance). The model's subsequent refinement and implementation took at least a decade, with the body of useful applications slowly taking shape alongside this refinement. Expectations-based loss aversion must be evaluated with the same patient historical lens.

Valuable enhancements to expectations-based models could take many forms — as one example, DellaVigna et al. (2017) incorporate both habituation and forecasted gains and losses in their formulation of the referent. It seems likely that different environments will call for different refinements as environmental forces call attention to different plausible referents. Absent an overarching theory for the sensitivity of referents to environmental forces, researchers should use (and recognize that they are using) their judgment in positing which referents are likely to be important in a given environment. At the same time, they should seriously engage methods to assess that judgment—perhaps by testing the robustness of conclusions to different assumptions about the referent, and perhaps even by developing techniques to let the data "nonparametrically" reveal what the referent might be for a particular environment.

Finally, it might be that we need to start considering models with multiple referents. Existing research tends to presume that the referent is based on past experience **or** on endogenous expectations **or** on some salient benchmark. It is plausible that decision makers are simultaneously influenced by two or more referents. When assessing how happy one is with one's salary, one might in part compare it to the past, in part compare it to expectations, and in part compare it to the average of one's peers. Of course, permitting multiple referents makes the task even more difficult, and opens up further degrees of freedom. But if the goal is to best understand the behaviors that we observe in the world, researchers should not shy away from the task.

References

Abdellaoui, Mohammed, "A Genuine Rank-Dependent Generalization of the Von Neumann-Morgenstern Expected Utility Theorem," *Econometrica*, 2002, *70* (2), 717–736.

_ **and Han Bleichrodt**, "Eliciting Gul's Theory of Disappointment Aversion by the Tradeoff Method," *Journal of Economic Psychology*, 2007, *28* (631-645).

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman, "Reference Points and Effort Provision," *American Economic Review*, 2011, *101* (2), 470–492.

Agarwal, Sumit, Mi Diao, Jessica Pan, and Tien Foo Sing, "Are Singaporean Cabdrivers Target Earners?," *Working Paper*, 2015.

Allais, Maurice, "Fondements d'une Thèorie Positive des Choix Comportant un Risque," in "Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40," Paris: Centre National de la Recherche Scientifique, 1953, pp. 127–140.

_ , "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," *Econometrica*, 1953, *21* (4), 503–546.

Allen, Eric J., Patricia M. Dechow, Devin G. Pope, and George Wu, "Reference-Dependent Preferences: Evidence from Marathon Runners," *Management Science*, 2017, *63* (6), 1657–1672.

Anscombe, F.J. and R.J. Aumann, "A definition of subjective probability," *Annals of mathematical statistics*, 1963, *34* (1), 199–205.

Arrow, Kenneth J., *Essays in the Theory of Risk Bearing*, Chicago: Marham Publishing Company, 1971.

Banerji, A. and Neha Gupta, "Detection, Identification, and Estimation of Loss Aversion: Evidence from an Auction Experiment," *American Economic Journal: Microeconomics*, 2014, *6* (1), 91–133.

Barberis, Nicholas, "Thirty Years of Prospect Theory in Economics: A Review and Assessment," *Journal of Economic Perspectives*, 2013, *27* (Winter), 173–196.

_ **and Wei Xiong**, "Realization Utility," *Journal of Financial Economics*, 2012, *104*, 251–271.

_ **, Ming Huang, and Tano Santos**, "Prospect Theory and Asset Prices," *Quarterly Journal of Economics*, 2001, *116* (1), 1–53.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue, and Joshua C. Teitelbaum, "Distinguishing Probability Weighting from Risk Misperceptions in Field Data," *American Economic Review, Papers and Proceedings*, 2013, *103* (3), 580–585.

_ **,** _ **,** _ **, and** _ , "The Nature of Risk Preferences: Evidence from Insurance Data," *American Economic Review*, 2013, *103* (6), 2499–2529.

_ **,** _ **,** _ **, and** _ , "Estimating Risk Preferences in the Field," *Journal of Economic Literature*, 2018, *56* (2), 501–564.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak, "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 1964, *9* (3), 226–232.

Bell, David E., "Disappointment in Decision Making under Uncertainty," *Operations Research*, 1985, *33* (1), 1–27.

Bellemare, Charles, Michaela Krause, Sabine Kroger, and Chendi Zhang, "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility," *Economics Letters*, 2005, *87*, 319–324.

Benartzi, Shlomo and Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 1995, *110*, 73–92.

Bernheim, B. Douglas and Charles Sprenger, "Direct Tests of Cumulative Prospect Theory," *Working Paper*, 2017.

Bernouilli, Daniel, "Specimen Theoriae Novae de Mensura Sortis," *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 1738, *5*, 175–192.

_ , "Exposition of a New Theory on the Measurement of Risk," *Econometrica*, 1954, *22* (1), 22–36.

Bowman, David, Deborah Minehart, and Matthew Rabin, "Loss Aversion in a Consumption/Savings Model," *Journal of Economic Behavior and Organization*, 1999, *38*, 155–178.

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper, "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 2010, *78* (4), 1375–1412.

Brunnermeier, Markus and Jonathan Parker, "Optimal Expectations," *American Economic Review*, 2005, *95* (4), 1092–1118.

Camerer, Colin, "Prospect Theory in the Wild: Evidence from the Field," in D. Kahneman and A. Tversky, eds., *Choices, Values, and Frames*, Cambridge University Press and Russell Sage Foundation, 2000.

Camerer, Colin F., Ana Dreber, Eskil Forsell, Teck-Hua Ho, Jurgen Huber, Magnus Johanesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Felix Holzmeister, Taisuke Imai, Siri Isakksson, Gideon Nave, Thomas Pfeiffer, Michael Razen, and Hang Wu, "Evaluating Replicability of Laboratory Experiments in Economics," *Science*, 2016, *315*, 1433–1436.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics*, 1997, *112* (2), 407–441.

Caplin, Andrew and John Leahy, "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics*, 2001, *116* (1), 55–79.

Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40

Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40*, Paris: Centre National de la Recherche Scientifique, 1953.*

Crawford, Vincent P. and Juanjuan Meng*, "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," American Economic Review, 2011, 101 (5), 1912–1932.*

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

Edwards, Ward*, "The Prediction of Decisions Among Bets," Journal of Experimental Psychology, 1955, 50 (3), 201–214.*

_ , "Subjective Probabilities Inferred from Decisions," Psychological Review, 1962, 69 (2), 109–135.

Ericson, Keith M. Marzilli and Andreas Fuster*, "Expectations as Endowments: Reference-Dependent Preferences and Exchange Behavior," Quarterly Journal of Economics, 2011, 126 (4), 1879–1907.*

Falk, Armin and Florian Zimmermann*, "Beliefs and Utility: Experimental Evidence on Preferences for Information," Working Paper, 2017.*

Farber, Henry*, "Is Tomorrow Another Day? The Labor Supply of New York City Cab Drivers," Journal of Political Economy, 2005, 113, 46–82.*

_ , "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers," American Economic Review, 2008, 98, 1069–1082.

— , *"Why you Can't Find a Taxi in the Rain and Other Lessons from Cab Drivers," Quarterly Journal of Economics, 2015, 130 (4), 1975–2026.*

Fehr-Duda, Helga and Thomas Epper, *"Probability and Risk: Foundations and Economic Implications of Probability-Dependent Risk Preferences," Annual Review of Economics, 2012, 4 (567-593).*

Fehr, Ernst and Lorenz Goette, *"Do Workers Work More When Wages Are High? Evidence from a Ranomized Field Experiment," American Economic Review, 2007, 97 (1), 298–317.*

Fishburn, Peter and Peter Wakker, *"The Invention of the Independence Condition for Preferences," Management Science, 1995, 41 (7), 1130–1144.*

Friedman, Milton and Leonard J. Savage, *"The Utility Analysis of Choices involving Risk," Journal of Political Economy, 1948, 56 (4), 279–304.*

Genesove, David and Christopher Mayer, *"Loss Aversion and Seller Behavior: Evidence from the Housing Market," Quarterly Journal of Economics, 2001, 116 (4), 1233–1260.*

Gill, David and Victoria Prowse, *"A Structural Analysis of Disappointment Aversion in a Real Effort Competition," American Economic Review, 2012, 102 (1), 469–503.*

Gneezy, Uri and Jan Potters, *"An Experiment on Risk Taking and Evaluation Periods," Quarterly Journal of Economics, 1997, 112 (2), 631–645.*

— , John A. List, and George Wu, *"The Uncertainty Effect: When a Risky Prospect Is Valued Less Than Its Worst Possible Outcome," Quarterly Journal of Economics, 2006, 121 (4), 1283–1309.*

— , Lorenz Goette, Charles Sprenger, and Florian Zimmermann, *"The Limits of*

Expectations-Based Reference Dependence," Journal of the European Economic Association, 2017, 15 (4), 861–876.

Goette, Lorenz, Annette Cerulli-Harms, and Charles Sprenger, *"Randomizing Endowments: An Experimental Study of Rational Expectations and Reference-Dependent Preferences," American Economic Journal: Microeconomics, Forthcoming.*

Gul, Faruk, *"A Theory of Disappointment Aversion," Econometrica, 1991, 59 (3), 667–686.*

Haigh, Michael S. and John A. List, *"Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis," Journal of Finance, 2005, 60 (1), 523–534.*

Harless, David W. and Colin F. Camerer, *"The Predictive Utility of Generalized Expected Utility Theories," Econometrica, 1994, 62 (6), 1251–1289.*

Heffetz, Ori, *"Are Reference Points Merely Lagged Beliefs Over Probabilities?," Working Paper, 2018.*

_ and John A. List, *"Is the Endowment Effect an Expectations Effect?," Journal of the European Economic Association, 2014, 12 (5), 1396–1422.*

Heidhues, Paul and Botond Koszegi, *"Competition and Price Variation when Consumers Are Loss Averse," American Economic Review, 2008, 98 (4), 1245–1268.*

_ and _ , *"Regular Prices and Sales," Theoretical Economics, 2014, 9, 217–251.*

Herweg, Fabian and Konrad Mierendorff, *"Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs," Journal of the European Economic Association, 2013, 11 (2), 399–432.*

_ , Daniel Muller, and Philipp Weinschenk, *"Binary Payment Schemes: Moral Hazard and Loss aversion," American Economic Review, 2010, 100 (5), 2451–2477.*

Horowitz, John K. and Kenneth E. McConnell, *"A Review of WTA/WTP Studies," Journal of Environmental Economics and Management, 2002, 44 (426-447).*

Kahneman, Daniel, *"Preface," in D. Kahneman and A. Tversky, eds., Choices, Values, and Frames, Cambridge University Press and Russell Sage Foundation, 2000.*

_ and Amos Tversky, *"Prospect Theory: An Analysis of Decision under Risk," Econometrica, 1979, 47 (2), 263–291.*

_ , Jack L. Knetsch, and Richard H. Thaler, *"Experimental Tests of the Endowment Effect and the Coase Theorem," Journal of Political Economy, 1990, 98 (6), 1325–1348.*

_ , _ , and _ , *"Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," Journal of Economic Perspectives, 1992, 58, 193–206.*

Knetsch, Jack L., *"The Endowment Effect and Evidence of Nonreversible Indifference Curves," American Economic Review, 1989, 79 (5), 1277–1284.*

_ and Wei-Kang Wong, *"The Endowment Effect and the Reference State: Evidence and Manipulations," Journal of Economic Behavior and Organization, 2009, 71, 407–413.*

_ , Fang-Fang Tang, and Richard H. Thaler, *"The Endowment Effect and Repeated Market Trials: Is the Vickrey Auction Demand Revealing?," Experimental Economics, 2001, 4 (2), 257–269.*

Koszegi, Botond, *"Ego Utility, Overconfidence, and Task Choice," Journal of the European Economic Association, 2006, 4 (4), 673–707.*

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

_ and _ , *"Reference-Dependent Risk Attitudes," American Economic Review, 2007, 97 (4), 1047–1073.*

_ **and** _ , *"Reference-Dependent Consumption Plans," American Economic Review, 2009, 99 (3), 909–936.*

Kreps, David M., *A Course in Microeconomic Theory, Princeton, NJ: Princeton University Press, 1990.*

Laibson, David, *"Golden Eggs and Hyperbolic Discounting," Quarterly Journal of Economics, 1997, 112 (2), 443–477.*

Lange, Andreas and Anmol Ratan, *"Multi-Dimensional Reference-Dependent Preferences in Sealed-Bid Auctions - How (most) Laboratory Experiments Differ from the Field," Games and Economic Behavior, 2010, 68, 634–645.*

List, John A., *"Does Market Experience Eliminate Market Anomalies?," Quarterly Journal of Economics, 2003, 118 (1), 41–71.*

_ , *"Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace," Econometrica, 2004, 72 (2), 615–625.*

Loewenstein, George, *"Anticipation and the Valuation of Delayed Consumption," Economic Journal, 1987, 97 (387), 666–684.*

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

Loomes, Graham and Robert Sugden, *"Disappointment and Dynamic Consistency in Choice under Uncertainty," Review of Economic Studies, 1986, 53 (2), 271–82.*

Malinvaud, Edmond, *"Note on von Neumann-Morgenstern's Strong Independence," Econometrica, 1952, 20 (4), 679.*

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

Mas-Colell, Andreu, Michael Dennis Whinston, and Jerry R. Green, *Microeconomic Theory, Vol. Volume 1, New York: Oxford University Press, 1995.*

Masatlioglu, Yusufcan and Collin Raymond, *"A Behavioral Analysis of Stochastic Reference Dependence," American Economic Review, 2016, 106 (9), 2760–2782.*

Mehra, Rajnish and Edward C. Prescott, *"The Equity Premium: A Puzzle," Journal of Monetary Economics, 1985, 15 (2), 145–161.*

Odean, Terrance, *"Are Investors Reluctant to Realize Their Losses?," Journal of Finance, 1998, 53 (5), 177–1798.*

O'Donoghue, Ted and Matthew Rabin, *"Doing It Now or Later," American Economic Review, 1999, 89 (1), 103–124.*

Oettinger, Gerald S., *"An Empirical Analysis of the Daily Labor Supply of Stadium Vendors," Journal of Political Economy, 1999, 107 (2), 360–392.*

Pagel, Michaela, *"Expectations-based Reference-Dependent Preferences and Asset Pricing," Journal of the European Economic Association, 2015, 14 (2), 468–514.*

_ , *"Expectations-Based Reference-Dependent Life-Cycle Consumption," Review of Economic Studies, 2017, 84 (2), 885–934.*

_ , *"Prospective Gain-Loss Utility: Ordered versus Separated Comparison," Working Paper, 2017.*

_ , *"A News-Utility Theory for Inattention and Delegation in Portfolio Choice," Econometrica, 2018, 86 (2), 491–522.*

Plott, Charles R. and Kathryn Zeiler, *"The Willingness to Pay-Willingness to Accept Gap, the "Endowment Effect," Subject Misconceptions and Experimental Procedures for Eliciting Valuations," American Economic Review, 2005, 95 (3), 530–545.*

— **and** — , *"Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?," American Economic Review, 2007, 97 (4).*

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

Pratt, John W., *"Risk Aversion in the Small and in the Large," Econometrica, 1964, 32 (1/2), 122–136.*

Prelec, Drazen, *"The Probability Weighting Function," Econometrica, 1998, 66 (3), 497–527.*

Preston, Malcolm G. and Philip Baratta, *"An Experimental Study of the Auction-Value of an Uncertain Outcome," American Journal of Psychology, 1948, 61 (2), 183–193.*

Quiggin, John, *"A Theory of Anticipated Utility," Journal of Economic Behavior and Organization, 1982, 3, 323–343.*

Rabin, Matthew, *"Risk aversion and expected utility theory: A calibration theorem," Econometrica, 2000a, 68 (5), 1281–1292.*

— , *"An Approach to Incorporating Psychology into Economics," American Economic Review, 2013, 103 (3), 617–622.*

— **and Richard H. Thaler**, *"Anomalies: Risk Aversion," Journal of Economic Perspectives, 2001, 15 (1), 219–232.*

Rees-Jones, Alex, *"Quantifying Loss-Averse Tax Manipulation," Review of Economic Studies, 2018, 85 (2), 1251–1278.*

Samuelson, Paul A., *"Probability, Utility, and the Independence Axiom," Econometrica, 1952, 20 (4), 670–678.*

_ , *"Utilité, Préfèrence et Probabilité," in "Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40," Paris: Centre National de la Recherche Scientifique, 1953, pp. 141–150.*

_ , *"Risk and Uncertainty: A Fallacy of Large Numbers," Scientia, 1963, 57 (98), 1–6.*

Savage, Leonard J., *"Une Axiomatisation de Comportement Raisonnable Face à l'Incertitude," in "Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40," Paris: Centre National de la Recherche Scientifique, 1953, pp. 29–33.*

_ , *The Foundations of Statistics, New York: J. Wiley, 1954.*

Segal, Uzi and Avia Spivak, *"First Order Versus Second Order Risk Aversion," Journal of Economic Theory, 1990, 51 (1), 111–125.*

Sprenger, Charles, *"An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," Journal of Political Economy, 2015, 123 (6), 1456–1499.*

Thakral, Neil and Linh T. To, *"Daily Labor Supply and Adaptive Reference Points," Working Paper, 2017.*

Thaler, Richard H., *"Toward a Positive Theory of Consumer Choice," Journal of Economic Behavior and Organization, 1980, 1, 39–60.*

Tversky, Amos and Daniel Kahneman, *"Loss Aversion in Riskless Choice: A Reference-Dependent Model," Quarterly Journal of Economics, 1991, 106, 1039–1061.*

_ **and** _ , *"Advances in Prospect Theory: Cumulative Representation of Uncertainty," Journal of Risk and Uncertainty, 1992, 5 (4), 297–323.*

Varian, Hal, *Microeconomic Analysis, third ed., Norton & Company, 1992.*

von Neumann, John and Oskar Morgenstern, *Theory of Games and Economic Behavior, Princeton University Press, 1944.*

Wakker, Peter, Ido Erev, and Elke U. Weber, *"Comonotonic Independence: The Critical Test Between Classical and Rank Dependent Utility Theories," Journal of Risk and Uncertainty, 1994, 9, 195–230.*

Wu, George, *"An Empirical Test of Ordinal Independence," Journal of Risk and Uncertainty, 1994, 9, 39–60.*