Experienced Utility versus Decision Utility: Putting the 'S' in Satisfaction*

Steven Carter†
Department of Economics
University of California, Irvine

Michael McBride
Department of Economics
University of California, Irvine

This version: June 2009

Abstract

Recent research distinguishes an individual's decision utility, inferred from her observed choices, from her experienced utility, which more closely matches the notion of happiness. Using various estimation techniques with a unique experimental data set, we test whether post-choice satisfaction (experienced utility), like decision utility, is S-shaped with loss aversion around a given reference point. We also present a model which estimates the satisfaction function and reference point simultaneously. When pooling the data across individuals, we find an S-shaped satisfaction function in which the reference point depends on past payments, social comparisons, and subjective expectations. There is mixed evidence of loss aversion. At the individual level, there is substantial variation in satisfaction function shapes, although the S-shape is common. Though the two notions of utility are distinct, our findings imply that the two are related at a fundamental level.

JEL Classifications: C91, D70, I30.
Keywords: happiness, utility, experiment, value function, Prospect Theory.

*We thank the Center for the Study of Democracy at UC Irvine for financial support and the California Social Science Experimental Laboratory at UCLA for programming assistance and use of laboratory facilities. For insightful and encouraging comments, we thank many friends and colleagues, including Garth Halloway, Ivan Jeliazkov, Marianne Bitler, and participants of the UCI Junior Faculty Lunch seminar, the 2008 Public Choice Society Meetings, and the Mini-conference on Decision Analysis held at UC Irvine in July 2007. Special thanks to Bill Branch for suggesting the subtitle.

†Corresponding author. 3151 Social Science Plaza, Irvine, CA, 92697-5100, sjcarter@uci.edu.

"[The Prospect Theory value function] represents the *decision utility* of the gains and losses associated with possible outcomes of the decision at hand; it is silent about the post-choice *experienced utility* of the reference situation ... [I]nferences from decision utility to experienced utility should be made with great caution." Kahneman (1999, pp. 18-19, his italics)

1 Introduction

The modern economist's notion of utility differs dramatically from that of utilitarian Jeremy Bentham and his contemporaries. To these earlier thinkers, utility was the sum of experienced pleasures minus pains. This hedonic view of utility fell out of favor in the early twentieth century when critics argued that pleasures and pains could not be measured.[1] Economists instead redefined utility to be a representation of preferences revealed through observed behavior and commenced the reconstruction of economic theories.[2] It was concluded that theories of experienced pleasures and pains were not only scientifically problematic, they were also unnecessary because utility conceived as consistent choice or revealed preference was sufficient for constructing economic theories and evaluating policies.[3]

Though this behaviorist view of utility still dominates economics, Kahneman and others have recently sparked a renewed interest in a notion of utility corresponding to hedonic experience (e.g., see Kahneman, Wakker, and Sarin 1997). "Experienced utility" is not only measurable, they argue, but also of fundamental importance for both understanding behavior and selecting public policies because, as shown in various studies, experienced utility ("enjoying") differs from decision utility ("wanting") in significant ways. For example, Tversky and Griffin (2000 [1991]) report evidence that realized outcomes such as payments matter more for decisions while contextual factors such as comparisons matter more for judgements of hedonic experience, Loewenstein and Adler (1995) find that subjects incorrectly predict the extent to which their post-choice satisfaction adapts, and Schwarz and Strack (1999)

[1] This shift in economic thinking was initiated by Robbins (1932) and is tied to skepticism in the early twentieth century about the possibility of interpersonal comparisons of utility (see Fleurbaey and Hammond 1998).

[2] For modern treatments of utility, see Fishburn (1987) or Mas-Collel, Whinston, and Green (1995). Stigler (1950a, 1950b) summarizes some developments of the use of the word "utility" in economics.

[3] However, Mandler (1999) describes problems that persist in the new theories.

review evidence that judgements result from cognitive processes that depend on non-choice related factors (also see Loewenstein and Ubel 2008).

This paper confronts the caution expressed in the epigraph. We explore whether the S-shape property of empirically estimated decision utility functions is also found in experienced utility. This S-shape represents Kahneman and Tversky's (1979) surprising yet ubiquitous finding that individuals are generally risk averse when making decisions about gains but risk loving when making decisions about losses. Figure 1 depicts such a function that is concave in gains but convex in losses. Although not all individuals' choices over gambles exhibit the S-shape, the S-shape is modal and therefore provides a benchmark for understanding how outcomes are valued (Luce 2000). The S-shaped value function assumed in Prospect Theory (Kahneman and Tversky 1979; Tversky and Kahneman 1992) provides perhaps the best known example of a theory inspired by the empirically derived S-shape.

This paper asks two questions: Is experienced utility S-shaped? If so, what determines the reference point that separates gains from losses? The reference point is often taken to be the status quo wealth position, a natural assumption in the context of gambles with monetary gains and losses as the possible outcomes. Yet, Kahneman and Tversky (1979: 277, 286) initially acknowledged that the reference point will more generally depend on hedonic adaptation, expectations, social comparisons, or other contextual factors. Previous research has examined various candidates as factors that influence the reference point; e.g., the social utility function literature focuses on social comparisons (see Suls and Wheeler 2000), the goal literature focuses on subjective expectations (see Pervin 1989), and the adaptation literature focuses on past experience (see Frederick and Loewenstein 1999). Because these factors have generally been studied separately, it is not known which are relatively more important when all are present. Even though "[t]he most basic problem of hedonic psychophysics is the determination of the level of adaptation and aspirations that separates positive from negative outcomes" (Kahneman and Tversky 1984: 349), "[h]ow multiple reference points are integrated is an open question" (Camerer and Loewenstein 2004: 17). This open question applies both to decision utility and experienced utility.

We use unique experimental data to estimate the shape of the post-choice *satisfaction function* and its associated reference point. A key characteristic of our data is that they contain measures of three factors believed to affect an individual's reference point: past outcomes, expected outcomes, and outcomes of potential comparison groups. We conduct a series of non-parametric and parametric regressions to let the data reveal the underlying relationship between an outcome, reference point, and satisfaction. Many satisfaction functions, when estimated by individual, are S-shaped, though this is not the modal shape. Thus, similar to what Luce (2000) concludes with respect to decision theory value functions, the S-shape should be understood as one of the various possible shapes observed in a heterogeneous population. When pooling all individuals' observations, we find that the shape of the satisfaction function depends on the factor selected as the reference point: it is S-shaped when using the expectations and social comparison as the reference point but is not always S-shaped when using past outcomes as the reference point. When simultaneously estimating a reference point function, we find strong evidence of an S-shape, though the evidence on loss aversion is mixed. We also find that the reference point depends on all three factors. Overall, we find strong evidence in favor of an S-shaped satisfaction function, and, as we suggest later, this evidence suggests that experienced utility and decision utility are fundamentally related even though they are conceptually distinct.

For a comprehensive treatment of the resurgent interest in experienced utility, see Kahneman, Diener, and Schwarz's (1999) edited volume.[4] See Kahneman and Thaler (2006) for a brief and more recent discussion. Kahneman, Wakker, and Sarin (1997) provide a formal normative theory of experienced utility, which provides a useful distinction between it and decision utility. Because this work is relatively recent, it is not surprising that there have been very few studies that examine whether experienced utility is S-shaped. Galanter (1990) uses magnitude scaling to find evidence that reported feelings of hypothetical monetary losses and gains are S-shaped. Vendrik and Woltjer (2007) use reported life satisfaction as a proxy for experienced utility and the predicted income of one's social comparison group as one's reference income to find that satisfaction is concave (not S-shaped) in annual income. Layard, Mayraz, and Nickell (2008) also use reported happiness to proxy for experienced utility and find happiness to be concave in income.[5]

[4] See Clark, Frijters, and Shields (2008) for a more recent survey on the closely related research on relative income and happiness.

[5] Though of a different vein, interested readers may also consider Brandstätter (2000), Loewenstein,

Our study differs from these three last mentioned studies in important ways. While Galanter (1990) asks respondents to rate happiness associated with hypothetical monetary gains and losses, our experimental data have subjects rate an experience related to actual monetary payments. Thus, our subjects' responses more closely approximate the feeling of an actual experience, while Galanter's data are better described as representing predicted experienced utility, which is known to suffer a variety of biases (Loewenstein and Schkade 1999, Kahneman and Thaler 2006). Our data also differ in that they contains multiple reference point factors, thereby allowing us to estimate how subjects calculate reference points. The survey data used by Vendrik and Woltjer (2007) and Layard, Mayraz, and Nickell (2008) capture the subjectively perceived value of life experience, and their rejection of the S-shape in life satisfaction (happiness) and income survey data is an important finding in the happiness literature. However, there are reasons why we might not predict an S-shape in the life-as-a-whole context. For example, individuals with income well below that of their comparison group may be living near or below subsistence, and such individuals may be very risk averse in income instead of risk seeking. The experimental data studied here are from a context much closer to the gambles setting (short time-frames, lower payments, etc.) in which are found S-shaped decision utility. That we find evidence of S-shape satisfaction in the gambles setting implies that the life-as-a-whole context contains factors that work against those that generate the S-shape. Vendrik and Woltjer also assume that the predicted income of a comparison group is the reference point, while we consider how multiple factors into a reference point. Layard, Mayraz, and Nickell do not consider reference points.

2 Hypotheses

Although we emphasize the decision utility and experienced utility distinction, there is reason to believe that an individual's satisfaction (experienced utility or happiness) function, like her decision utility function, should be S-shaped in our study. Specifically, decision utility and experienced utility may both arise from the same underlying psychophysical processes.

Thompson, and Bazerman (1989), and Messick and Sentis (1985). Rablen (2008) defines a value function to be the difference between *ex ante* decision utility and *ex post* experienced utility, which he calibrates to an S-shape.

Psychophysics is a branch of psychology that studies the relationship between causal physical stimuli and psychological values or judgements. An example would be the study of how subjects rate the heaviness of various objects. Psychophysical research finds that the amount a stimulus must be increased for an individual to recognize the difference is proportional to the stimulus magnitude itself (Marks and Algom 1998). For example, an individual notices the difference in weight between ten and fifteen pound objects but not between fifty and fifty-five pound objects; to notice an increase in weight from fifty pounds, the increase must be much more than five pounds. Similar findings arise when studying the perception of brightness or loudness. The conclusion is that psychological intensity is generally found to be a concave function in physical intensity. Many years ago, Bernoulli used similar logic to propose utility functions with diminishing marginal returns (Zabell 1987).

The S-shaped decision utility function is believed to arise from similar psychophysical processes: an individual experiences "diminishing sensitivity" from her reference point so that gains relative to the reference point feel less and less good and losses feel less and less bad and that a loss is always more worse than an equal sized gain is good (Hastie and Dawes 2001). Diminishing marginal returns to gains yields a concave function in gains. Diminishing returns to losses implies that the "badness" of a loss is a concave function in losses, which translates into a convex function in goodness.

Notice that this claim about the origin of the S-shaped decision utility function implies that the mapping from psychological valuation to decision utility maintains the fundamental S-shape of the underlying psychological valuation. In effect, the claim is that "wanting" is directly related to underlying psychological valuation. Moreover, if one accepts this claim, one might also accept a claim that "enjoying" or "likeability" is also directly related to the underlying psychological valuation in a manner that also maintains the S-shape. According to this logic, decision utility and experienced utility are indirectly related: each is a mapping of the same underlying psychophysical processes.

Formally, let $z(y|r)$, with $y \in \mathbb{R}$, $r \in \mathbb{R}$, represent the underlying psychophysical sensation of stimulus y given reference point r such that $z(\cdot)$ is differentiable over $(-\infty, \infty)$ except at inflection point r, and with

$$\left.\frac{\partial z}{\partial y}\right|_{y\geq r} > 0, \left.\frac{\partial^2 z}{\partial y^2}\right|_{y\geq r} < 0,$$
$$\left.\frac{\partial z}{\partial y}\right|_{y<r} > 0, \left.\frac{\partial^2 z}{\partial y^2}\right|_{y<r} > 0.$$

Let $\widetilde{u}(z(y|r))$ be a real-valued *wantability* function that maps sensation to wantability ("u" for the standard decision utility), and let $\widetilde{h}(z(y|r))$ be a real-valued *likeability* function ("h" for happiness) that maps sensation to likeability or enjoyment. It is natural to assume that both wantability and likeability are increasing in sensation, $\frac{\partial \widetilde{u}}{\partial z} > 0$ and $\frac{\partial \widetilde{h}}{\partial z} > 0$. For convenience, we may also assume that each is also differentiable over $(-\infty, \infty)$.

When a choice based study reports a decision utility function, it would thus be reporting $u(y|r)$, where $u(y|r) \equiv (\widetilde{u} \circ z)(y|r)$. Although the distinction between the psychological sensation $z(\cdot)$ and the decision mapping $\widetilde{u}(\cdot)$ is conflated, as long as any curvature of $\widetilde{u}(\cdot)$ is not too strong, the S-shape in the underlying sensation is reflected in $u(\cdot)$. Specifically, if

$$\left.\frac{\partial^2 \widetilde{u}}{\partial z^2}\right|_{y<r} > \frac{-\frac{\partial \widetilde{u}}{\partial z}\left.\frac{\partial^2 z}{\partial y^2}\right|_{y<r}}{\frac{\partial z}{\partial y}\frac{\partial z}{\partial y}}, \left.\frac{\partial^2 \widetilde{u}}{\partial z^2}\right|_{y\geq r} < \frac{-\frac{\partial \widetilde{u}}{\partial z}\left.\frac{\partial^2 z}{\partial y^2}\right|_{y\geq r}}{\frac{\partial z}{\partial y}\frac{\partial z}{\partial y}},$$

then $u(\cdot)$ is S-shaped in y with inflection point at r (see Proof 1 in the appendix). With $\frac{\partial \widetilde{u}}{\partial z} > 0, \frac{\partial z}{\partial y} > 0$, $\left.\frac{\partial^2 z}{\partial y^2}\right|_{y\geq rr} < 0$, and $\left.\frac{\partial^2 z}{\partial y^2}\right|_{y<r} > 0$, this condition states that $\frac{\partial^2 \widetilde{u}}{\partial z^2}$ must be greater than a negative amount and less than a positive amount. In other words, the curvature of $\widetilde{u}$ must be sufficiently close to zero.

Observed happiness $h(y|r)$, with $h(y|r) \equiv \left(\widetilde{h} \circ z\right)(y|r)$, also conflates the enjoyment mapping and sensation, yet a similar condition on the second derivative of $\widetilde{h}(\cdot)$ yields a similar result. Given that we observe S-shaped $u(\cdot)$ in experimental data, which implies limited curvature in $\widetilde{u}(\cdot)$, we may expect there to be limited curvature in $\widetilde{h}(\cdot)$.

Hypothesis 1 *Experienced utility $h(y|r)$ is S-shaped.*

For loss aversion in $z(\cdot)$ to also be reflected in $u(\cdot)$ or $h(\cdot)$ requires other conditions on the wantability and likeability functions. With $z(x|x) = 0$, we say that sensation $z(\cdot)$ has (strict) loss aversion if, for all $x > 0$,

$$z(r + x|r) < |z(r - x|r)|.$$

For u to have loss aversion, assuming $\widetilde{u}(z(0|r)) = 0$, it must be true that for all $x > 0$

$$\widetilde{u}(z(r+x|r)) < |\widetilde{u}(z(r-x|r))| .$$

A sufficient condition for this to be true is that $\widetilde{u}(\cdot)$ is concave (weak or strict) or not too strictly convex (see Proof 2 in the appendix). That wantability exhibits loss aversion suggests that this is the case, and we can conjecture that likeability exhibits a similar property.

Hypothesis 2 *Experienced utility $h(y|r)$ exhibits loss aversion.*

As mentioned above, the literature on reference points has identified three particular factors at work in reference point determination. The first is that individuals adapt to past payments, i.e., that higher past payments increase one's current reference point. The second is that higher payments within one's comparison group lead to an increase in one's reference point. The third is that the higher an individual's expected payment, the higher the reference point.

Hypothesis 3

(a) *The reference point is increasing in past payment.*

(b) *The reference point is increasing in comparison payment.*

(c) *The reference point is increasing in expected payment.*

There currently is no unifying theory of how different reference factors aggregate into a single reference point. Without any particular guidance, we here propose a benchmark hypothesis.

Hypothesis 4 *The reference point is a weighted average of past payment, comparison payment, and expected payment.*

3 Experimental Data

This study uses data from the experiment presented in McBride (2009). We provide here a brief description of the experiment and refer the reader to his original paper for more details.

After entering the experimental computer lab,[6] each subject sits at a computer terminal, receives verbal instructions, participates in one practice round, and participates in 25 real rounds. In each round, each subject plays the Matching Pennies game against a randomly selected computer "partner-type." There are five possible partner-types:

20% heads – 80% tails

35% heads – 65% tails

50% heads – 50% tails

65% heads – 35% tails

80% heads – 20% tails.

The computer reports the partner-type to the subject which informs the subject of the probability distribution used by the computer to select the opponent's coins in that round. The choice of partner-type is i.i.d. across subjects and time so that in any given round some subset of the subjects will be matched with a 20-80 type, another subset will be matched with a 35-65 type, and so on. After being told the partner-type, the subject chooses heads or tails for each of five coins, and the computer randomly and independently selects heads or tails according to the partner-type distribution. For example, in a given round, the subject might be matched with a 35-65 partner-type, choose (heads, heads, tails, tails, tails) for its five coins, then learn that the computer randomly selected (tails, tails, heads, tails, tails) for the partner. If the subject's first coin and the computer's first coin match (either both are heads or both are tails), then the subject wins the coin, and so on for the other coins. Figure 2 depicts the subject's payoff matrix for a single coin choice. With five coin choices, a subject can win 0, 1, 2, 3, 4, or 5 coins in any given round.

After the computer partner's choices are made, the computer reports to the subject the coin choices made by the computer and the number of coins won by the subject. The subject is also told the average coins won by others (not including herself) by partner type; i.e., she is told the average of all those matched with a 20-80 partner-type, the average of all those matched with a 35-65 partner type, and so on.

Immediately after being told the outcome of a round (i.e., the number of coins won and,

[6] The experiment was conducted at the California Social Science Experimental Laboratory (CASSEL) located at UCLA.

depending on the treatment, information about others' coins won), the subject is asked, "How satisfied are you with the result of this round?" The subject then reports her satisfaction on a scale of 1 to 7, with 1 signifying "very dissatisfied," 4 signifying "satisfied," and 7 signifying "very satisfied."[7] The form of this question matches the convention used in happiness and life satisfaction surveys (Schwarz and Strack 1999). Although answers to these subjective questions suffer from various imperfections, it is widely held that these data meaningfully capture relevant aspects of happiness or satisfaction.[8]

After all subjects report their satisfaction levels, the next round begins. Subjects are randomly assigned a new, possibly different, partner-type. The selection of partner-types is independent across subjects. Coin choices are independent across subjects and across coins within a round for a given subject. Information on partner-types, coin choices, and payments from prior rounds remains on the computer screen. The experiment session ends after all 25 rounds have been completed. At the experiment's end, subjects were paid actual US dollars for their coins received according to an exchange rate of 8 coins for 1 dollar. The experiment used 36 subjects and lasted approximately one hour. The average total take home amount was roughly $17.

This experiment design has many important features. First, it yields data on multiple reference point factors—past payments, expected payments, and others' payments—thought to affect an individual's assessment of well-being. The subject's past payments and that subject's information about others' payments are displayed on the subject's monitor. The subject's expected payment can be inferred from the partner-type. If the subject acts to maximize her expected payment, she will choose all heads when paired with an 80-20 or 65-35 type, all tails when paired with a 35-65 or 20-80 type, and anything when paired with a 50-50 type. These optimal actions yield expected payments for the round of 4 against the

[7] This type of question has been used to gather satisfaction or subjective well-being (i.e., happiness) data in experiments (e.g., Charness and Grosskopf 2001) and surveys (e.g., Ferrer-i-Carbonell and Frijters 2004). The term satisfaction is thought to entice a more cognitive response than the term happiness, which is thought to be more emotive. The income-happiness literature uses the terms interchangeably because they appear empirically equivalent (e.g., the title of Van Praag and Ferrer-i-Carbonell's (2004) book includes both terms). We suspect that the terms would yield similar results in this experiment, yet future experiments must verify that conjecture. We use satisfied in the question in the unlikely case that satisfy does prompt a more cognitive, thoughtful response.

[8] See Diener (1984) for an extended discussion of happiness and satisfaction questions. See Krueger and Schkade (2008) for a recent analysis of the statistical validity and interpretation of thse questions.

80-20 and 20-80 types, 3.25 against the 65-35 and 35-65 types, and 2.5 against the 50-50 type. If subjects act to maximize their expected payments, then these expected payments are valid proxies for subjects' subjectively perceived expected payments. Also note that the past payments, expected payments, and others' payments are all reported in the same units, so we can directly compare coefficients in my regressions to make specific statements about the relative magnitude of each factor's impact on satisfaction.

Second, the payment structure facilitates the interpretation of the results. The payment range is 0 to 5, thus yielding six possible payment outcomes, while the satisfaction rating is on a seven category scale. Because the scale and payment range are not the same, subjects will be less inclined to associate a particular monetary payment with a "natural" satisfaction report. For example, if the satisfaction scale was 0 to 5 like the payment range, and a subject received payment 3, she might automatically associate a payment 3 with satisfaction 3. Under this design, the satisfaction report requires more of a subjective assessment and any non-linearity in the responses can contain meaningful information.

Third, the optimal decision to maximize expected payments is a simple one. As described above, the expected payment maximizing choice depends on the partner-type, is easy for the subject to deduce, and is incentivized because outcomes correspond to take-home monetary amounts. As will be seen below, the large majority of individual coin choices do correspond to expected payment maximization, which supports our use of the expected payment to capture a subject's subjectively perceived expected payment.

Finally, the satisfaction data arising from the experiment are clearly post-choice. The satisfaction question is asked after the outcome of the choice. Moreover, it is asked in manner designed to get at the quality of the outcome.

4 Econometric Methodology

4.1 Econometric Issues and Estimation Strategy

There exist various econometric methods for estimating a satisfaction function, yet there is no single best method given the constraints of our data. Three particular issues stand out. First, the dependent variable is discretely ordinal and not cardinal. Technically speaking, concavity in gains and convexity in losses are psychological conditions that can exist whether or not they are cardinally measurable (Mandler 1999), yet we in this paper are concerned with empirically identifying such properties. As such, we must be interested in satisfaction as a cardinal property and thus must confront the limitation of having only ordinal data. Second, conducting formal statistical tests of the hypotheses requires that we impose a parametric structure on the satisfaction function, but imposing such structure limits the possible shapes that the estimated function can take. Third, previous studies find that subjective reports of satisfaction or happiness exhibit significant individual-level fixed effects due to personality or genetic traits (Ferrer-i-Carbonell and Frijters 2004), yet adequately controlling for fixed effects is not possible in all regression frameworks. It is possible for OLS but is problematic in discrete dependent variable contexts.

Economists typically use ordered probit or ordered logit regressions when working with ordinal data. Because these regressions condition on the ordinal properties of the dependent variable by assuming an underlying latent variable function, neither their estimates nor the estimates of the underlying latent variable function have a cardinal interpretation (see Vendrik and Woltjer 2007 and references therein). Thus, these regressions do not allow us to make statements about the shape of the satisfaction function. With this in mind, previous researchers in the economics of happiness have assumed that the ordinal response data have cardinal meaning, e.g., Vendrik and Woltjer (2007) and Layard, Mayraz, and Nickell (2008). We also make this assumption[9].

An argument can be made for assuming cardinality with our data. First, it makes interpretation of the results straightforward and can allow for formal statistical tests. Second, having many response categories—more than there are possible payment outcomes—suggests that any non-linearity in responses may reflect non-linearity in satisfaction. One might think that a player who receives the lowest or highest possible payment would give the lowest or highest possible satisfaction response, respectively. However, when there are more response

[9] It could be argued that Bayesian techniques can actually bridge the gap between assuming a cardinal interpretation of the dependent variable and treating it as strictly ordinal. To explore this possibility, we estimated ordered probit regressions using Bayesian methods. We discuss this methodology and results in a technical appendix that is available from the authors upon request. We note briefly here that we did find an S-shape using expected payment and comparison payment but not past payment. However, as explained in that appendix, giving the results a cardinal interpretation is problematic.

choices than payment outcomes, any non-linearity in the reported satisfaction can be more easily manifested. Third, there is evidence that ratings scales like those in our data provide what psychophysicists call linear judgement functions. The idea is that an underlying stimulus (e.g., payment) produces a psychological sensation (e.g., level of satisfaction) that the subject must then judge for placement on the reporting scale (e.g., satisfaction report). Some research indicates that the judgement process is linear for category scales (see Marks and Algom 1998, 137). For our study, this implies that any non-linearity in responses results from non-linearity in underlying psychological satisfaction and not the actual reporting of satisfaction.

With our cardinality assumption in place, our strategy to confront the other econometric issues mentioned above is to estimate satisfaction functions under a variety of methods and assumptions and look for an emergent pattern. The basic equation of interest is

$$\begin{aligned} h_{it} &= f(d_{it}) + \varepsilon_{it}, \\ d_{it} &= y_{it} - r_{it}, \end{aligned}$$

where, for subject i in round t, h_{it} is the reported happiness, d_{it} is the difference between the monetary payment y_{it} and the reference point r_{it}, and ε_{it} is a random disturbance term. We first estimate non-parametric spline regressions at both the individual level and with pooled data. These regressions allow for maximum flexibility in functional shape while assuming a particular reference point. We then estimate two types of parametric regressions: non-linear power form regressions and fixed effects OLS regressions. The power form regressions allow for formal statistical tests and for the explicit estimation of a reference point function, but they do not account for individual fixed effects. The fixed effects OLS regressions control for fixed effects and allow for higher order polynomials, but they do not allow for explicit reference point estimation.

4.2 Spline Regressions

One option available to allow for a flexible functional form is a cubic spline regression. In general, spline functions are functional forms fit onto a sub-segment of the data. These sub-segments are specified by k knot points $\xi_1, ..., \xi_k$, which can be chosen *a priori*, or determined analytically through the data.[10] Once knots are selected, the variable entering the spline operator (in our case, d_{it}) is transformed into a $1 \times k-1$ vector $D_{it} = [d_{i1}, d_{i2}, ..., d_{ik-1}]$ so as to estimate the parameters α and β of following linear regression

$$h_{it} = \alpha + D_{it}\beta + \varepsilon_{it}, \tag{1}$$

where

$$\begin{aligned} d_{i1} &= d_{it}, \\ d_{ij} &= \frac{\left(d_{it} - \xi_{j-1}\right)_+^3 - \frac{\left(d_{it}-\xi_{k-1}\right)_+^3\left(\xi_k-\xi_{j-1}\right)}{\left(\xi_k-\xi_{k-1}\right)} + \frac{\left(d_{it}-\xi_k\right)_+^3\left(\xi_{k-1}-\xi_{j-1}\right)}{\left(\xi_k-\xi_{k-1}\right)}}{\left(\xi_k - \xi_1\right)^2}. \end{aligned} \tag{2}$$

The subscript "+" denotes the inclusion of the term inside the parentheses if $d_{it} > \xi_{j-1}$, zero otherwise. For example, let $j = 2$ and $k = 3$. If $d_{it} = 20$, $\xi_1 = 19$, $\xi_2 = 25$, $\xi_3 = 36$, then $d_{i2} = \frac{(20-19)^3}{(36-19)^2}$. Since the observation $d_{it} = 20$ is less than the other knots, the other parts of equation (2) contain zeros in the numerator and drop out of the equation. We can estimate equation (1) by OLS and use linear predictions to calculate the expected satisfaction. These predictions are then plotted against the scaled payment variable to trace out an S-shape. We also calculate and plot confidence intervals for the spline functions using the standard error of prediction based on the estimation of equation (1). Simply, the interval is the prediction plus or minus the 1.96 times the estimated standard error.

As with prospect theory, the inflection point in the satisfaction function corresponds to the reference point. Choice of the reference point is a non-trivial matter because it is a central variable for theory. This data set provides us with three possible reference points taken exogenously. The first is the payment received by player i in the previous round $t-1$ (PREVPAY). The second is the expected payment that player i should receive based on the partner type (EPMAX). The third reference point is the average payment received by all other subjects (excluding player i) in the round who have the same partner type (TYPEAVG). The independent variable in the spline equation is the difference of the payment received and one of the reference points mentioned. We note that there exists the

[10] It is common to use percentiles or evenly spaced interval values as knot points. However, since knot point choice is not crucial to our analysis, we remain agnostic about our knot points.

possibility that the true reference point is some weighted average of PREVPAY, EPMAX and TYPEAVG,

$$r_{it} = g_0 + g_1 PREVPAY_{it} + g_2 EPMAX_{it} + g_3 OWNAVG_{it},$$

where the g coefficients are unknown *a priori* and can be estimated. For the spline analysis, we do not attempt to employ any unobserved weighted average as the reference point, but we do employ the weighted average for our power regressions.

4.3 Power Form Non-linear Regressions

The power form is a functional form widely used in the economics and psychology literature because multiple studies find that it well represents subjects' choices (Luce 2000: 80). Assuming d_{it} maintains the definition given earlier, the power functional form is

$$h_{it}(d_{it}) = \begin{cases} \beta_1 d_{it}^{\alpha}, & d_{it} \geq 0, \\ \beta_2 (-d_{it})^{\alpha}, & d_{it} < 0, \end{cases} \tag{3}$$

where $h_{it}(d_{it}) \geq 0$ if $d_{it} \geq 0$ and $h_{it}(d_{it}) < 0$ if $d_{it} < 0$.

If the function is S-shaped, then we must have

$$\beta_1 > 0, \tag{4}$$

$$\beta_2 < 0, \tag{5}$$

$$0 < \alpha < 1. \tag{6}$$

Loss aversion requires

$$|\beta_1| < |\beta_2|. \tag{7}$$

One benefit of the power form regression is that it allows us to not only choose an exogenous reference point but to also model an unobserved reference point that is a function of PREVPAY, EPMAX and TYPEAVG. Let y_{it} be the current monetary payment define $\widehat{r}_{it}$ be the estimate of the unobserved reference point

$$\widehat{r}_{it} = g_0 + g_1 y_{it-1} + g_2 e_{it} + g_3 c_{it}. \tag{8}$$

Dropping the d_{it} notation temporarily so that the inequality constraints can be easily represented, we can thus write the satisfaction function for i in time t as

$$h_{it} = \begin{cases} \beta_1 (y_{it} - \widehat{r}_{it})^{\alpha} - 4, & y_{it} \geq \widehat{r}_{it} (PREVPAY_{it}, EPMAX_{it}, TYPEAVG_{it}), \\ \beta_2 (\widehat{r}_{it} - y_{it})^{\alpha} - 4, & y_{it} < \widehat{r}_{it} (PREVPAY_{it}, EPMAX_{it}, TYPEAVG_{it}). \end{cases} \tag{9}$$

The minus 4 normalizes the boundary between satisfaction and satisfaction to 0. For the S-shape and loss aversion, we need conditions (4)-(7) to hold as before.

We assume that any error comes in the satisfaction directly and not through the reference point calculation. These assumptions yield the following non-linear least squares regression:

$$\widehat{r}_{it} = g_0 + g_1 PREVPAY_{it} + g_2 EPMAX_{it} + g_3 TYPEAVG_{it} \tag{10}$$

$$h_{it} = \begin{cases} \beta_1 (y_{it} - \widehat{r}_{it})^{\alpha} - 4 + \varepsilon_{it}, & y_{it} \geq \widehat{r}_{it}, \\ \beta_2 (\widehat{r}_{it} - y_{it})^{\alpha} - 4 + \varepsilon_{it}, & y_{it} < \widehat{r}_{it}. \end{cases} \tag{11}$$

Notice that the power function shape (parameters β_1, β_2, and α) are estimated simultaneously with the reference point function parameters (g_0, g_1, g_2, and g_3). This fact is important because it allows the maximization procedure to identify the reference point that best matches the inflection point of the S-shape should it have an S-shape. The estimation procedure imposes some meaningful structure on the reference point by tying it to the three factors thought to affect judgements instead of just allowing the procedure to identify any reference point for each observation. Also note that, more in line with our understanding of reference points, this setup assumes that the reference point factors $(PREVPAY_{it}, EPMAX_{it}, TYPEAVG_{it})$ impact satisfaction via their effects on the reference point. The factors will thus shift the entire function horizontally instead of vertically. This regression also allows for direct testing of loss aversion because we can formally test conditions (4)-(7).

We can also conduct our estimation allowing the exponent term to differ for relative gains and losses, i.e.,

$$h_{it} = \begin{cases} \beta_1 (y_{it} - \widehat{r}_{it})^{\alpha_1} - 4 + \varepsilon_{it}, & y_{it} \geq \widehat{r}_{it}, \\ \beta_2 (\widehat{r}_{it} - y_{it})^{\alpha_2} - 4 + \varepsilon_{it}, & y_{it} < \widehat{r}_{it}. \end{cases}$$

The S-shape now arises when the following conditions hold: $\beta_1 > 0$, $\beta_2 < 0$, $0 < \alpha_1 < 1$ and $0 < \alpha_2 < 1$. Testing for loss aversion is not as simple as checking $|\beta_1| < |\beta_2|$ because differences in α_1 and α_2 are relevant. Instead, loss aversion requires checking the predicted happiness at each estimated point. That is, if $\widehat{h}_{it}(d)$ is the predicted happiness at point d, then loss aversion arises when

$$\left|\widehat{h}_{it}(d)\right| < \left|\widehat{h}_{it}(-d)\right| \text{ for } d > 0.$$

4.4 Fixed-effects OLS Regressions

Given the setup of the data, the observations include individuals who report satisfaction for multiple rounds. It is natural to suppose that there is some unobservable characteristic that affects the reported satisfaction with the outcome and that is heterogenous across individuals. For example, some individuals playing the game may have a higher level of baseline happiness. Controlling for these individual effects (these are the fixed effects because they are assumed to not change over the course of the game play) can be accomplished through differencing, or individual dummy variable identification, which strategy we use here. We also cluster the standard errors at the level of the individual. We model our estimation as

$$h_{it} = f(d_{it}) + c_i + \varepsilon_{it},$$

using the previously defined d_{it}. The term c_i is the unobserved individual level characteristic controlled for by dummy variables. To allow for the emergence of an S-shaped satisfaction function, we approximate the function $f(d_{it})$ in the fixed-effects OLS model with polynomial expansions of orders one, three, and five:[11]

$$\begin{array}{ll} \text{Order 1:} & f(d_{it}) = \beta_1 d_{it} \\ \text{Order 3:} & f(d_{it}) = \beta_1 d_{it} + \beta_2 d_{it}^2 + \beta_3 d_{it}^3 \\ \text{Order 5:} & f(d_{it}) = \beta_1 d_{it} + \beta_2 d_{it}^2 + \beta_3 d_{it}^3 + \beta_4 d_{it}^4 + \beta_5 d_{it}^5. \end{array}$$

The polynomial expansions capture any higher order curvature in reported satisfaction. The order one expansion is included in the analysis for comparison purposes with the order 3 and order 5 expansions.

[11] We estimated polynomial expansions of orders two and four as well, to check the sensitivity of the results. A polynomial of Order 2 constrains the functional form from exhibiting an S shape, and was left out. The Order 4 coefficient for each reference point was not statistically significant.

5 Results

5.1 Summary Statistics

Table 1 summarizes the payment and reported satisfaction data. Each cell in the table displays the number of subjects who reported the satisfaction level of the vertical axis when receiving the payment amount on the horizontal axis. The mean reported satisfaction, reported below each payment-satisfaction combination in the table, increases as the payment increases. The modal reported satisfaction for each payment amount (boxed) is also (weakly) increasing in payment. A cursory glance at both the mean and modal reported satisfaction suggests an S-shape because the reported satisfaction in either case increases at an accelerating rate at lower payments and that rate decreases at higher payment levels.

Table 2 presents evidence on the validity of the EPMAX measure. Roughly 90% of all coin choices are consistent with the assumption that subjects acted to maximize their expected payoffs. The percentage of times that a subject's five-coin decision in a round matches expected payment maximization is much lower, probably a reflection of a misunderstanding of independence. However, the large majority are consistent, thus suggesting that the expected payment proxy is a valid proxy of a subject's actual expected payment.

5.2 Non-parametric Regression Results

We first use non-parametric spline techniques to estimate three satisfaction functions for each individual, one for each of the three reference factors. Visual inspection then allows us to classify each function as either S-shaped, linear, convex, concave, with multiple inflection points, or other; formal statistical tests are not possible due to the flexible estimation structure. Table 3 summarizes what we observed. The S-shape is the modal shape when using the previous payment as the reference points: 13 out of 36 individual functions show the S-shape, 9 exhibit multiple inflections, 7 are linear, and so on. The S-shape is not modal when using the expected payment (only 7 out of 36) or comparison payment (only 9 out of 26) as the reference point; however, the S-shape does appear frequently. Moreover, we note that this classification is fairly strict. We have 25 observations per individual, and with more observations per individual, we suspect that some of those functions classified as multiple inflection, which comprise the largest category, would likely become S-shaped as the function's smoothness would increase.

Figures 3(a)-(c) plot the satisfaction functions and 95% confidence intervals estimated via non-parametric splines when pooling the data. Each estimated function used a different one of the three candidate reference points. When using expected payment or comparison payment as the reference point, we find clear evidence of the S-shape. It is not S-shaped when using previous payment; it is fairly linear but slightly concave. If indeed the reference point depends on all three factors, then the S-shape may best represent the data overall.

5.3 Power Form Results

Table 4 presents results from two non-linear least squares regressions that estimate coefficients of a power form. Regression 1 assumes that the exponent is the same regardless of whether or not the payment is larger than the reference point. Regression 2 allows for the exponent term to be different, depending on whether or not the payment is larger than the reference point. Figure 4 displays the two regressions graphically.

To test the first hypothesis (S-shape), we look at exponent coefficient α. An exponent greater between zero and one produces an S-shape, though values closer to one approximate linearity. Regression 1 yields $\alpha = 0.61$, which is statistically significant, confirming an S-shape. Regression 2 allows the exponent to differ for relative gains and losses. When the payment is greater than the reference point, the estimate α_1 is 0.73 and is statistically significant. When the payment is less than the reference point, the estimate of α_2 is 0.48 and is also statistically significant. All estimates confirm the S-shape. We reject the hypothesis that $\alpha_1 = \alpha_2$ at the 5% level; the concavity in relative gains is not as sharp as that for relative losses. Figure 4 plots the two regressions and clearly illustrates the S-shape.

For the second hypothesis (loss aversion), we test the equivalence of β_1 and β_2 in regression 1. If $\beta_1 = \beta_2$ in regression 1, then a one unit loss relative to the reference point has the same magnitude change in utility as a one unit gain. Loss aversion requires that a change in satisfaction be larger for losses than for gains. We find, under the assumption of a common value of α, that β_1 differs statistically from β_2, yet, the difference is not in the direction as expected. Loss aversion would imply $|\beta_1| < |\beta_2|$ in regression 1, but we find $|\beta_1| > |\beta_2|$, though the level of statistical significance is not high. When estimating separate α's in regression 2, we do find $|\beta_1| < |\beta_2|$, however we must now consider the different exponent estimates α_1 and α_2 when checking for loss aversion. A visual inspection of the regression 2 in Figure 4 shows that there is not loss aversion for small gains or losses but that there is loss aversion for larger gains and losses. Thus, evidence of loss aversion is mixed.

For the third hypothesis (reference point effects), we consider the coefficients g_0, g_1, g_2, and g_3 estimated for the reference point function. In both regressions, we find significant evidence that the three coefficients on the reference point factors (g_1, g_2, and g_3) are jointly significant. Consistent with Hypothesis 3, the expected payment and comparison payment have statistically significant positive effects on the reference point in the first regression. Both work the same direction in the second regression, though the effect of an increase in comparison payment is now not statistically significant. Contrary to the hypothesis, the previous payment has a negative effect on the reference point. Larger values in the previous round reduce the perceived reference point. As explained by McBride (2008), this could be due to a short-term positive glow effect from receiving a high payment that lasts into the next round. In words, a subject who receives a high payment in round t might have a temporary happy feeling that lasts into the next round.

To test the fourth hypothesis (weighted reference point factors), we test if $g_0 = 0$ and $g_1 + g_2 + g_3 = 1$. The F-statistics in the last row of Table 4 strongly reject this hypothesis even at the 1% level in each regression. The reference point is not a simple weighted average of the three reference point factors. Instead, there is a large fixed component. Moreover, the expected payment and comparison payments have a much larger impact than the previous payment which, as mentioned above, works in an opposite direction.

5.4 Fixed-effect OLS Regression Results

Table 5 displays the results of various fixed-effects OLS regressions: three order 1 polynomial regressions, three order 3 polynomial regressions, and three order 5 polynomial regressions. For each order, we estimated a regression using each different reference point factor as the reference point. Various test statistics are reported at the bottom of the table. Consider the order 3 regressions. For the expected payment and comparison payment reference points, the coefficients for each polynomial order are significant, confirming the need for higher order polynomials. In contrast, the order 3 coefficient for the previous payment reference point is not significant at the 10% level. However, when tested jointly with the second order coefficient, the null hypothesis is rejected and the order 3 specification should be included. Under previous payment and comparison payment reference points, the order 5 regressions do not show significant order 4 coefficients, even when order 5 effects appear to be marginally significant, but there are such effects under previous payment. For the order 5 regressions, we conduct two joint tests: the first to test the restrictions on the order 4 and 5 coefficients, while the second is the joint significance test on all regression coefficients. We find that under previous payment and comparison payment reference points, the coefficients for orders 4 and 5 are jointly insignificant, indicating that an order 3 polynomial model would be sufficient to capture the curvature, though the order 5 polynomial specification is necessary when previous payment is used as the reference point.

Figure 6 plots the estimated satisfaction function for each of these regressions. We see the clear S-shape under expected payment and comparison payment. The higher order curvature estimated in the polynomial regressions produce S-shapes as predicted in these cases. The S-shape does not appear under previous payment; instead it is a mirrored S-shape with concavity in losses and convexity in gains.

To assess the robustness of this result, we partitioned the data into two subsamples. The first subsample included all subject-rounds in which the subject choose all five coins consistent with expected payoff maximization, while the second subsample included subject-rounds in which the subject did not act consistently with expected payoff maximization.[12] We find a much stronger S-shape in the first sample when using expected payment and comparison payment as the reference point (results not shown). We cannot be sure why this specific pattern is found, yet there is an interpretation consistent with our overall emerging picture. Subjects who act consistently with expected payoff maximzation have more accurate expectations about their payoffs, and the expected payoff maximization and comparison

[12] Any coin choice by a subject matched with a 50-50 partner-type is consistent with expected payoff maximization, and there is no way to separately identify the subjects matched with 50-50 partners who were actively trying to maximize expected payoffs from those who were not. For this reason, we took a conservative approach and included them in the second subsample.

payment reference point factors would more accurately reflect those subjects' self-perceived reference point. If true, then the second subsample underestimates the S-shape that exists in those subjects' actual satisfaction functions. This interpretation lends added credence to the overall conclusion that subjects have S-shaped satisfaction functions.

5.5 Discussion

Consistent with Hypothesis 1, we find an emerging pattern from our empirical analysis that satisfaction is well-represented by an S-shaped satisfaction function. This is true especially when using expected payment or comparison payment as reference point; it is less true when using previous payment as the reference point. When estimating a reference point function in our power regressions, we find that expected payment and comparison payment have the anticipated effects on reference points but previous payment does not. Thus, two out of the three parts of Hypothesis 3 are verified. Hypothesis 4 is rejected: the reference point is increasing faster in expected payment and comparison payment than it is decreasing in previous payment. How to assess Hypothesis 4 is less clear because the evidence of loss aversion is mixed. Overall, however, the S-shape fits the data well in most of our regressions.

Why do we find the S-shape? One possible two-part explanation goes as follows. First, decision utility and experienced utility are fundamentally related via underlying psychophysical processes as suggested in Section 2. Second, the differences that do exist between decision utility and experienced utility increase in prominence as the time between the decision and the reporting of the experience. Because the time between the making of a decision and the reporting of satisfaction in the experimental data is fairly short (usually a matter of seconds), any difference between decision utility and experienced utility would be minimized. This conjecture could be tested by running additional experiments that vary the length of time between the decision and satisfaction report.

This logic situates our paper somewhere between empirical estimations of the decision utility functions on one end of the spectrum and Vendrik and Woltjer's (2007) and Layard, Mayraz, and Nickell's (2008) results on the other end. Decision utility is derived from observed decisions. Life satisfaction, on the other hand, is a reflective assessment of one's overall life situation and reflects the outcomes of a large number of life decisions (e.g., whether or not to go to college, accept a job offer, marry, etc.). That the cited studies found a concave (not S-shaped) life satisfaction function would thus reflect a difference between decision utility and experienced utility that arises over time in the context of the reflective life satisfaction question. The satisfaction with an experiment payment in our data is much closer to the actual decision made. Any indirect linkage, psychophysical or not, between decision and experienced utility would be stronger in our setting.

Why is the evidence for loss aversion mixed? One possibility is that the distribution of realized payments is heavily skewed, with the mean payment 3.24 much closer to the highest possible payment 5 than the lowest possible payment 0. Suppose that 3.24 best represents the reference point, that a subject who receives a payment of 5 achieves her highest possible satisfaction and would report satisfaction 7, and that a subject who receives a payment of 0 achieves her lowest satisfaction and would report satisfaction 1. Then a relative gain of 1.76 (equals $5 - 3.24$) yields an increase in satisfaction of 3 (from $7 - 4$), while a relative loss of -3.24 (equals $0 - 3.24$) yields a change in satisfaction of -3 (from $1 - 4$). Having reference points closer to the highest possible category implies a smaller range of relative gains which, according to the interpretation of the reporting scale, are spread over a larger reporting range. This leads to a steeper function in gains than losses not due to psychophysical processes but due to distribution of payments.

6 Conclusion

This paper uses various econometric techniques to estimate experimental subjects' reported satisfaction functions. We find strong evidence that an S-shape best represents subjects' satisfaction responses, and we suggest that this shape could arise from psychophysical processes similar to those that generate S-shaped decision utility. We do not find compelling evidence of loss aversion, but we do find a role for various factors in the formation of subjects' reference points. Expected payments and comparison payments work to raise reference points as predicted but previous payments do not.

Future work has many avenues to pursue in building on this work. One direction is to explicitly account for time in a subjects' satisfaction function: How does the shape of experienced utility change over time? Do people become less sensitive to comparisons or expectations? Another direction is to further examine how reference points are determined: Which other reference point factors should be considered in a theory of reference point determination? A competing theory called range-frequency explicitly incorporates the properties of the distribution of stimuli into its account of how subjects rate those stimuli. Future work should compare the range-frequency theory with the theory presented here to determine which better fits the data. Work along these lines will yield a more complete understanding of how people subjectively experience the outcomes of their many economic decisions.

A Proofs

Proof 1: First apply the chain rule twice to obtain the second derivative $\frac{\partial^2 u}{\partial y^2}$ for $y \geq r$, and then solve for when it is less than 0 to obtain the condition for concavity for when $y \geq r$:

$$\left.\frac{\partial u}{\partial y}\right|_{y \geq r} = \frac{\partial \widetilde{u}}{\partial z}\frac{\partial z}{\partial y}$$

$$\left.\frac{\partial^2 u}{\partial y^2}\right|_{y \geq r} = \frac{\partial^2 \widetilde{u}}{\partial z^2}\frac{\partial z}{\partial y}\frac{\partial z}{\partial y} + \frac{\partial \widetilde{u}}{\partial z}\frac{\partial^2 z}{\partial y^2} < 0 \Rightarrow \frac{\partial^2 \widetilde{u}}{\partial z^2} < \frac{-\frac{\partial \widetilde{u}}{\partial z}\frac{\partial^2 z}{\partial y^2}}{\frac{\partial z}{\partial y}\frac{\partial z}{\partial y}}.$$

With $\frac{\partial \widetilde{u}}{\partial z} > 0, \left.\frac{\partial^2 z}{\partial y^2}\right|_{y \geq r} < 0$, and $\frac{\partial s}{\partial y} > 0$, this condition states that $\frac{\partial^2 \widetilde{u}}{\partial s^2}$ must be less than a positive amount.

Applying the chain rule twice again and solving for when the second derivative is greater than 0 to find convexity for $y < r$, we get:

$$\left.\frac{\partial u}{\partial y}\right|_{y < r} = \frac{\partial \widetilde{u}}{\partial z}\frac{\partial z}{\partial y}$$

$$\left.\frac{\partial^2 u}{\partial y^2}\right|_{y < r} = \frac{\partial^2 \widetilde{u}}{\partial z^2}\frac{\partial z}{\partial y}\frac{\partial z}{\partial y} + \frac{\partial \widetilde{u}}{\partial z}\frac{\partial^2 z}{\partial y^2} > 0 \Rightarrow \frac{\partial^2 \widetilde{u}}{\partial z^2} > \frac{-\frac{\partial \widetilde{u}}{\partial z}\frac{\partial^2 z}{\partial y^2}}{\frac{\partial z}{\partial y}\frac{\partial z}{\partial y}}.$$

With $\frac{\partial \widetilde{u}}{\partial s} > 0, \left.\frac{\partial^2 s}{\partial y^2}\right|_{y < r} > 0$, and $\frac{\partial s}{\partial y} > 0$, this condition states that $\frac{\partial^2 \widetilde{u}}{\partial s^2}$ must be greater than a negative amount. $\square$

Proof 2: Need to show when $\widetilde{u}(z(r+x|r)) < |\widetilde{u}(z(r-x|r))|$.

<u>Case I:</u> $\widetilde{u}(\cdot)$ is weakly concave. Define $\Delta z(x,r) \equiv -z(r-x|r) - z(r+x|r)$. With $z(0|r) = 0$, we see that $-z(r-x|r) > 0$ and $z(r+x|r) > 0$. With loss aversion in $z(\cdot)$,

$\Delta z(x,r) > 0$ for all $x > 0$. Notice that $\frac{\partial \widetilde{u}}{\partial z} > 0$ and $\frac{\partial^2 \widetilde{u}}{\partial z^2} \leq 0$ implies $\widetilde{u}(z') - \widetilde{u}(0) \leq \widetilde{u}(0) - \widetilde{u}(-z')$ for all $z' > 0$. With $\widetilde{u}(0|r) = 0$, this in turn implies $\widetilde{u}(z') \leq |\widetilde{u}(-z')|$ for all $z > 0$ and $|\widetilde{u}(-z')| < |\widetilde{u}(-z' - z'')|$ with $z'' > 0$. If we let $z' = z(r + x|r)$ and $z'' = \Delta z(x,r)$, then it follows for all $x > 0$ that

$$\begin{aligned}
\widetilde{u}(z(r+x|r)) &< |\widetilde{u}(-v - v')| \\
&= |\widetilde{u}(-(z(r+x|r)) - \Delta z(x,r))| \\
&= |\widetilde{u}(z(r-x|r))|.
\end{aligned}$$

<u>Case II:</u> $\widetilde{u}(\cdot)$ is strictly convex. Contrary to Case I, convexity implies $\widetilde{u}(z') > |-\widetilde{u}(-z')|$. However, with $|-\widetilde{u}(-z')|$ increasing in z', it must be true that there exists some $z'' > 0$ such that $\widetilde{u}(z') < |-\widetilde{u}(-z' - z''')|$ for all $z''' > z''$. The more convex is $\widetilde{u}(\cdot)$, the higher is z''. Hence if $\widetilde{u}(\cdot)$'s convexity is sufficiently small, then z'' is sufficiently small so that $z'' < \Delta z(x,r)$. If true, then

$$\begin{aligned}
\widetilde{u}(z(r+x|r)) &< |\widetilde{u}(-v - v')| \\
&= |\widetilde{u}(-(z(r+x|r)) - \Delta z(x,r))| \\
&= |\widetilde{u}(z(r-x|r))|. \quad \square
\end{aligned}$$

References

[1] Brandstätter, Eduard. 2000. "Comparison Based Satisfaction: Contrast and Empathy." *European Journal of Social Psychology* 30: 673-703.

[2] Camerer, Colin and George Loewenstein. 2004. "Behavioral Economics: Past, Present, and Future," in C. Camerer, G. Loewenstein, and M. Rabin, eds., *Advances in Behavioral Economics*, New York, NY: Russell Sage Foundation.

[3] Charness, Gary and Brit Grosskopf. 2001. "Relative Payoffs and Happiness: An Experimental Study." *Journal of Economic Behavior and Organization* 45: 301-328.

[4] Clark, Andrew, Paul Frijters, and Michael Shields. 2008. "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature* 46: 95-144.

[5] Diener, Ed. 1984. "Subjective Well-being." *Psychological Bulletin* 95: 542-575.

[6] Ferrer-i-Carbonell, Ada and Paul Frijters. 2004. "How Important is Methodology for the Estimates of the Determinants of Happiness?" *Economic Journal* 114: 641-659.

[7] Fishburn, Peter. 1987. "Representation of Preferences," in John Eatwell, Murray Milgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics.* New York: Macmillan Press Limited.

[8] Frederick, Shane and George Loewenstein. 1999. "Hedonic Adaptation," pp. 302-329 in D. Kahneman, E. Diener, and N. Schwarz, eds., *Well-Being: The Foundations of Hedonic Psychology.* New York: Russell Sage Foundation.

[9] Hastie, Reid, and Robin Dawes. 2001. *Rational Choice in an Uncertain World: The Psychology of Decision Making.* Thousand Oaks, CA: Sage Publications.

[10] Galanter, Eugene. 1990. "Utility Functions for Nonmonetary Events." *American Journal of Psychology* 103(4): 449-470.

[11] Kahneman, Daniel. 1999. "Objective Happiness," in D. Kahneman, E. Diener, and N. Schwarz, eds., *Well-Being: The Foundations of Hedonic Psychology.* New York, NY: Russell Sage Foundation.

[12] Kahneman, Daniel, Ed Diener, and Norbert Schwarz. 1999. *Well-being: The Foundations of Hedonic Psychology.* New York, NY: Russell Sage Foundation.

[13] Kahneman, Daniel and Richard Thaler. 2006. "Utility Maximization and Experienced Utility." *Journal of Economic Perspectives* 20(1): 221-234.

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

[15] Kahneman, Daniel and Amos Tversky. 1984. "Choices, Values, and Frames." *American Psychologist* 39(4): 341-350.

[16] Kahneman, Daniel, Peter Wakker, and Rakesh Sarin. 1997. "Back to Bentham? Explorations of Experienced Utility." *Quarterly Journal of Economics* 112(2): 375-405.

[17] Krueger, Alan and David Schkade. 2008. "The Reliability of Subjective Well-being Measures." *Journal of Public Economics* 92(8-9): 1833-1845.

[18] Layard, Richard, Guy Mayraz, and Stephen Nickell. 2008. "The Marginal Utility of Income." *Journal of Public Economics* 92(8-9): 1846-1857.

[19] Loewenstein, George and Daniel Adler. 1995. "A Bias in the Prediction of Tastes." *Economic Journal* 105(431): 929-937.

[20] Loewenstein, George and David Schkade. 1999. "Wouldn't it be Nice? Predicting Future Feelings," in D. Kahneman, E. Diener, and N. Schwarz, eds., *Well-Being: The Foundations of Hedonic Psychology.* New York, NY: Russell Sage Foundation.

[21] Loewenstein, George, Leigh Thompson, and Max Bazerman. 1989. "Social Utility and Decision Making in Interpersonal Comparisons." *Journal of Personality and Social Psychology* 57 (3): 426-441.

[22] Loewenstein, George and Peter Ubel. 2008. "Hedonic Adaptation and the Role of Decision and Experience in Public Policy." *Journal of Public Economics* 92(8-9): 1795-1810.

[23] Luce, R. Duncan. 2000. *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches.* Mahwah, NJ: Lawrence Erlbaum Associates.

[24] Mandler, Michael. 1999. *Dilemmas in Economic Theory: Persisting Foundational Problems of Microeconomics.* New York: Oxford University Press.

[25] Marks, Lawrence and Daniel Algom. 1998. "Psychophysical Scaling," pp. 81-178 in Michael Birnbaum, ed., *Measurement, Judgement, and Decision Making.* New York, NY: Academic Press.

[26] Mas-Collel, Andreu, Michael Whinston, and Jerry Green. 1995. *Microeconomic Theory.* New York: Oxford University Press.

[27] McBride, Michael. 2009. "Money, Happiness, and Aspirations: An Experimental Study." Department of Economics, University of California, Irvine.

[28] Messick, David and Keith Sentis. 1985. "Estimating Social and Nonsocial Utility Functions from Ordinal Data." *European Journal of Social Psychology* 15: 389-399.

[29] Pervin, Lawrence. 1989. "Goal Concepts in Personality and Social Psychology: A Historical Perspective," pp. 1-17 in L. Pervin, ed., *Goal Concepts in Personality and Social Psychology.* Hillsdale, New Jersey: Erlbaum Associates.

[30] Rablen, Matthew. 2008. "Relativity, Rank, and the Utility of Income." *Economic Journal* 118: 801-821.

[31] Robbins, Lionel. 1932. *An Essay into the Nature and Significance of Economic Science.* London: Macmillan and Co., Limited.

[32] Schwarz, Norbert and Fritz Strack. 1999. "Reports of Subjective Well-being: Judgemental Processes and Their Methodological Implications," in D. Kahneman, E. Diener, and N. Schwarz, eds., *Well-Being: The Foundations of Hedonic Psychology.* New York, NY: Russell Sage Foundation.

[33] Stigler, George. 1950a. "The Development of Utility Theory, I" *Journal of Political Economy* 58: 307-327.

[34] Stigler, George. 1950b. "The Development of Utility Theory, II" *Journal of Political Economy* 58: 373-396.

[35] Suls, Jerry and Ladd Wheeler, eds. 2000. *Handbook of Social Comparison: Theory and Research.* New York: Kluwer Academic/ Plenum Publishers.

[36] Tversky, Amos and Dale Griffin. 2000 [1991]. "Endowments and Contrast in Judgements of Well-being," pp. 709-725 in D. Kahneman and A. Tversky, eds., *Choices, Values, and Frames*, New York, NY: Russell Sage Foundation. Originally published in R. J. Zeckhauser, ed., *Strategy and Choice*, 1991, Cambridge, MA: MIT Press.

[37] Tversky, Amos and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cummulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5(4): 297-323.

[38] Van Praag, B.M.S and Ada Ferrer-i-Carbonell. 2004. *Happiness Quantified: A Satisfaction Calculus Approach.* Oxford: Oxford University Press.

[39] Vendrik, Maarten and Geert Woltjer. 2007. "Happiness and Loss Aversion: Is Utility Concave or Convex in Relative Income?" *Journal of Public Economics* 91(7-8): 1423-1448.

[40] Zabell, S.L. 1987. "Daniel Bernoulli," in John Eatwell, Murray Milgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics.* New York: Macmillan Press Limited.

Figure 1: An S-shaped Function with Loss Aversion

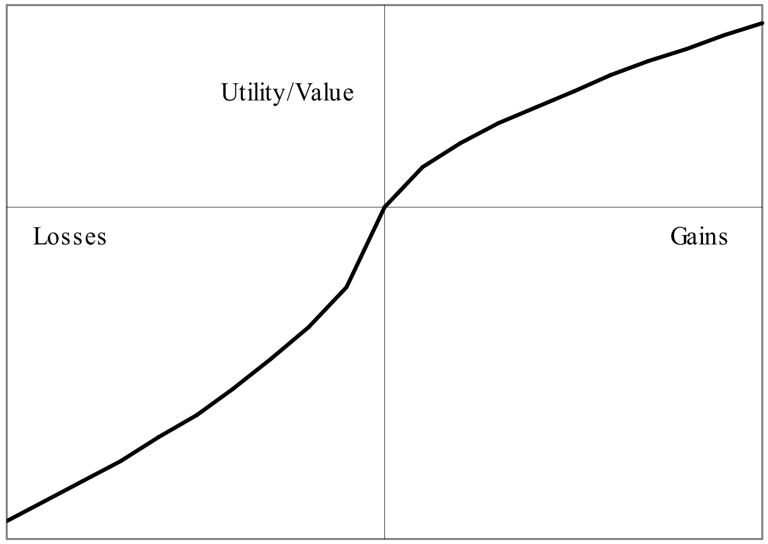

Figure 2: The Matching Pennies Game

		Computer: Heads	Computer: Tails
Subject	Heads	1	0
	Tails	0	1

Table 1: Observations by Payment and Reported Satisfaction

		Payment						
		0	1	2	3	4	5	Total
	7	0	0	1	10	47	**120**	178
	6	0	0	0	4	**117**	1	122
Reported	5	0	0	1	38	51	0	90
Satisfaction	4	1	3	9	**96**	39	7	155
	3	0	4	33	58	4	0	99
	2	0	5	**65**	28	1	0	99
	1	**11**	**50**	44	15	0	1	121
	Total	12	62	153	249	259	129	864
	Mean	1.25	1.35	2.10	3.67	5.62	6.78	4.24
	Increase in Mean		0.10	0.74	1.57	1.95	1.16	

Notes: 900 observations used for Treatment C with 36 total subjects.

Table 2: Percent of Decisions Consistent with Expected Payoff Maximization

	Percent of Coin Choices Consistent with Expected Payoff Maximization	Percent of Subject-rounds in which all Five Coin Choices Consistent with Expected Payoff Maximization	Percent of Subjects who always Chose Consistently with Expected Payoff Maximization
Pooled	93%	73%	25%
Expected Payment 2.5	100%	100%	100%
Expected Payment 3.25	86%	58%	25%
Expected Payment 4.0	94%	78%	53%

Notes: 900 observations used for Treatment C with 36 total subjects.

Table 3: Number of Subjects by Estimated Nonparametric Function Shape

Reference Point	S	Multiple Inflections	Linear	Concave	Convex	Other	Total
Previous Payment							
Number of Subjects	13	9	7	3	1	3	36
Percent of Subjects	36%	25%	19%	8%	3%	8%	
Expected Payment							
Number of Subjects	7	14	0	9	6	0	36
Percent of Subjects	19%	39%	0%	25%	17%	0%	
Comparison Payment							
Number of Subjects	9	13	9	1	2	2	36
Percent of Subjects	25%	36%	25%	3%	6%	6%	

Notes: Results obtained from cubic spline regressions with five knots.

Figure 3(a): Spline Satisfaction Function with Expected Payment as Reference Point and 95% Confidence Intervals

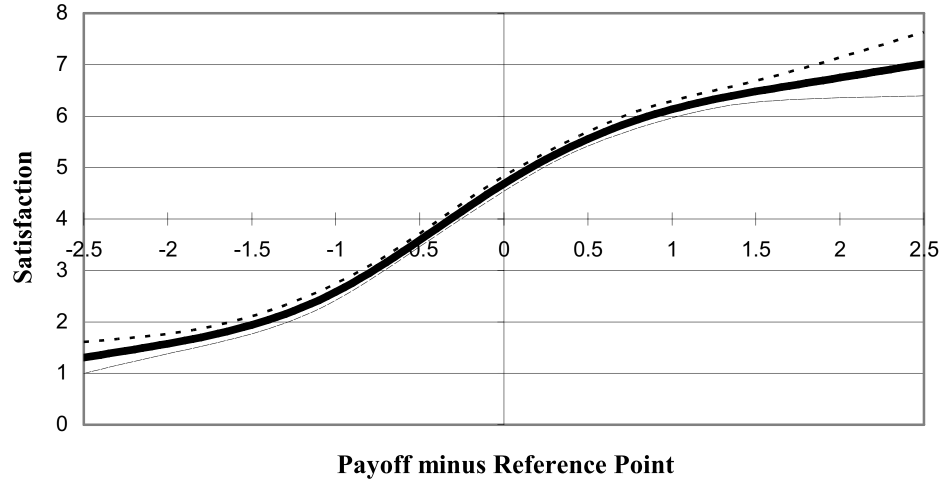

Figure 3(b): Spline Satisfaction Function with Previous Payment as Reference Point and 95% Confidence Intervals

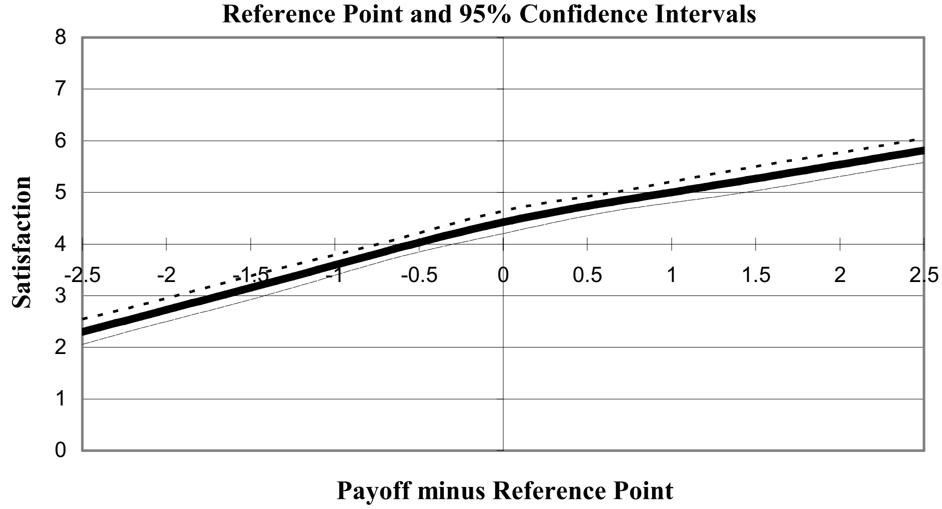

Figure 3(c): Spline Satisfaction Function with Comparison Payment as Reference Point and 95% Confidence Intervals

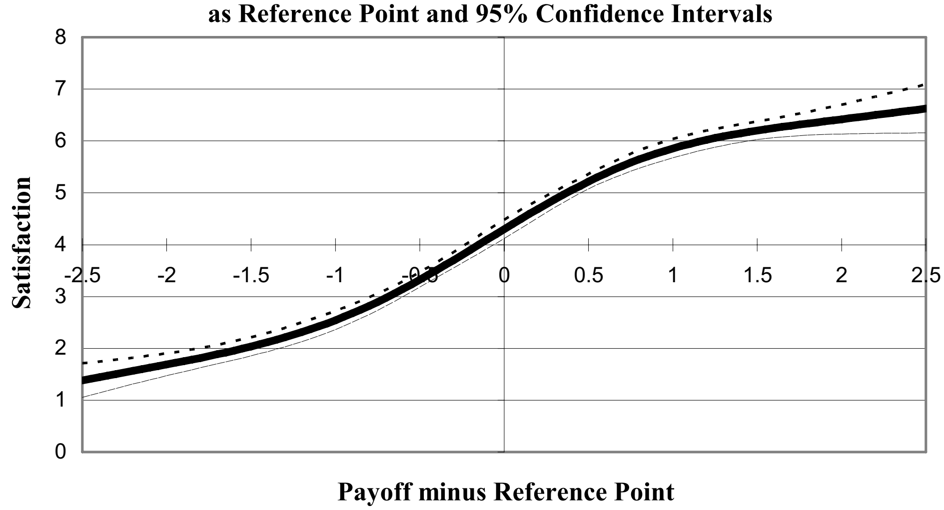

Table 4: Power Form Non-linear Regression Results

	(1)	(2)
Reference Point Function		
g_0 (Constant)	2.66**	2.81**
	(0.11)	(0.11)
g_1 (Previous Payment)	-0.04**	-0.03**
	(0.01)	(0.01)
g_2 (Expected Payment)	0.07*	0.04
	(0.03)	(0.03)
g_3 (Comparison Payment[1])	0.10**	0.07**
	(0.03)	(0.03)
Power Form Function		
α (Exponent)	0.61**	--
	(0.04)	
α_1 (Gain Exponent)	--	0.73**
		(0.07)
α_2 (Loss Exponent)	--	0.48**
		(0.07)
β_1 (Gain Coefficient)	1.87**	1.74**
	(0.05)	(0.07)
β_2 (Loss Coefficient)	-1.71**	-1.84**
	(0.06)	(0.08)
Observations	864	864
Adjusted R^2	0.75	0.75
W-test: $\beta_1 = -\beta_2$	2.87*	0.52
Test: $\alpha_1 = \alpha_2$	--	5.20*
Joint test: $\beta_1 = -\beta_2$, $\alpha_1 = \alpha_2$	--	7.81*
Test: $g_0 = 0$	514.85**	665.31**
Test: g_1 = g2 = g3 = 0	25.42**	9.56*
Test: $g_0 = 0$, g_1 + g2 + g3 = 1	525.35**	665.36**

Notes: The dependent variable in each regression was reported satisfaction. Each regression used the last 24 rounds of the experiment session. Standard errors are listed in parentheses. * and ** denote significance at 5% and 1% levels.

Figure 4: Power Regression Function

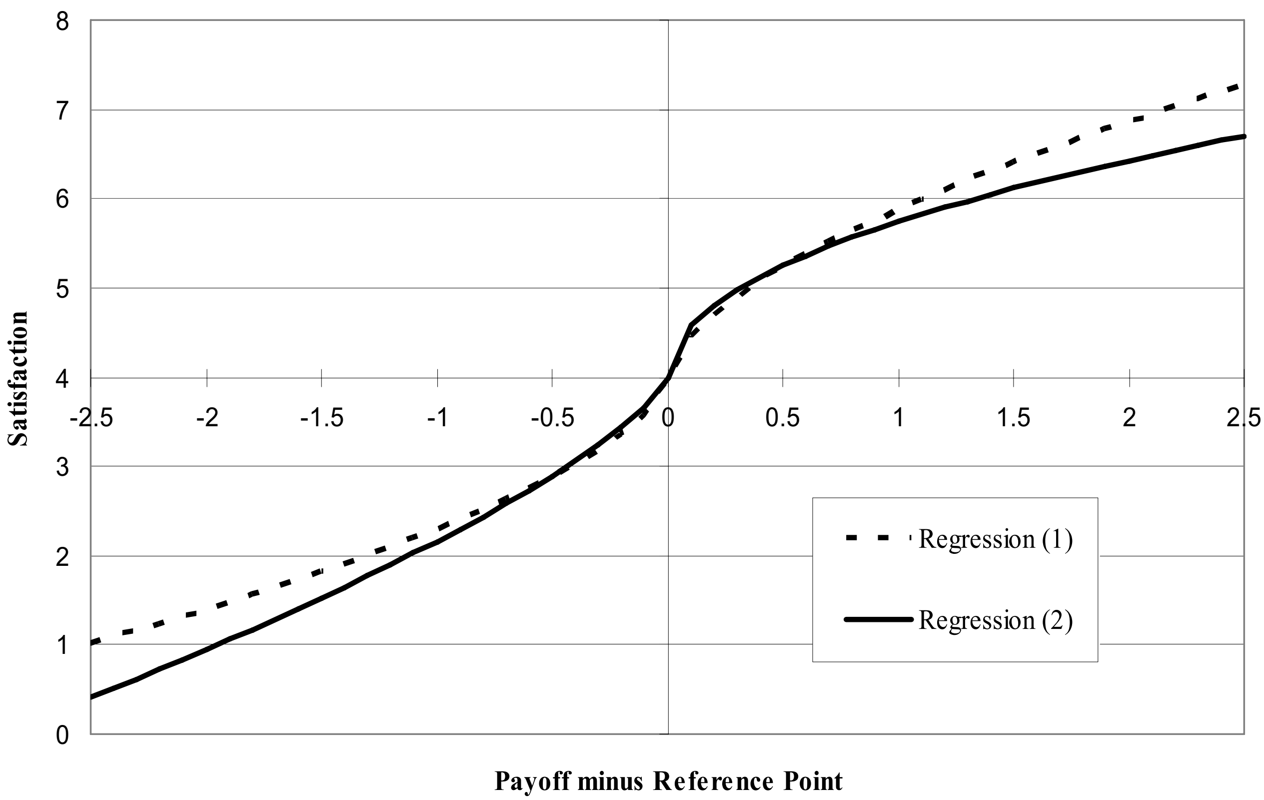

Table 5: Fixed-effects OLS Polynomial Regression Results by Reference Point

	Order 1 Polynomial			Order 3 Polynomial			Order 5 Polynomial		
Reference Point	Exp. Payment	Comp. Payment	Prev. Payment	Exp. Payment	Comp. Payment	Prev. Payment	Exp. Payment	Comp. Payment	Prev. Payment
Constant	4.49***	4.24***	4.23***	4.59***	4.28***	4.33***	4.62***	4.29***	4.43***
	(0.013)	(0.0000)	(0.00004)	(0.042)	(0.036)	(0.047)	(0.071)	(0.065)	(0.069)
Payment - Reference Point	1.49***	1.33***	0.70***	1.87***	1.63***	0.72***	1.97***	1.77***	0.67***
	(0.007)	(0.066)	(0.034)	(0.095)	(0.099)	(0.055)	(0.140)	(0.146)	(0.063)
$(\text{Payment - Reference Point})^2$	--	--	--	-0.13***	-0.06**	-0.032*	-0.21*	-0.08	-0.11***
				(0.042)	(0.029)	(0.017)	(0.013)	(0.082)	(0.039)
$(\text{Payment - Reference Point})^3$	--	--	--	-0.13***	-0.08***	-0.001	-0.20***	-0.17***	0.01
				(0.021)	(0.015)	(0.004)	(0.053)	(0.058)	(0.014)
$(\text{Payment - Reference Point})^4$	--	--	--	--	--	--	0.02	0.01	0.005***
							(0.016)	(0.010)	(0.002)
$(\text{Payment - Reference Point})^5$	--	--	--	--	--	--	0.01*	0.01*	-0.006
							(0.005)	(0.005)	(0.0006)
Observations	864	864	864	864	864	864	864	864	864
R^2	0.61	0.56	0.33	0.63	0.57	0.33	0.63	0.58	0.33
Test: all coefficients = 0	370***	405***	422***	182***	220***	169***	162***	144***	237***
Test: $\beta_2 = \beta_3 = 0$	--	--	--	21.01***	15.98***	2.54*	--	--	--
Test: $\beta_4 = \beta_5 = 0$	--	--	--	--	--	--	1.71	2.10	5.92***

Notes: The dependent variable in each regression was reported satisfaction. Robust standard errors clusted by individual are listed in parentheses. Subscripts *, **, *** and denote significance at 10%, 5%, and 1% levels.

Figure 6: Fixed-effects OLS Polynomial Satisfaction Functions

Figure 6(a): Order 1 Satisfaction Functions

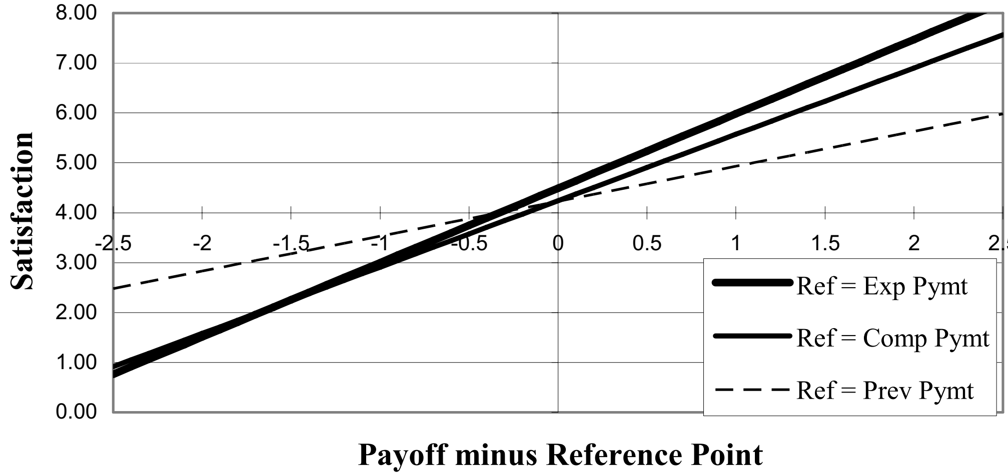

Figure 6(b): Order 3 Satisfaction Functions

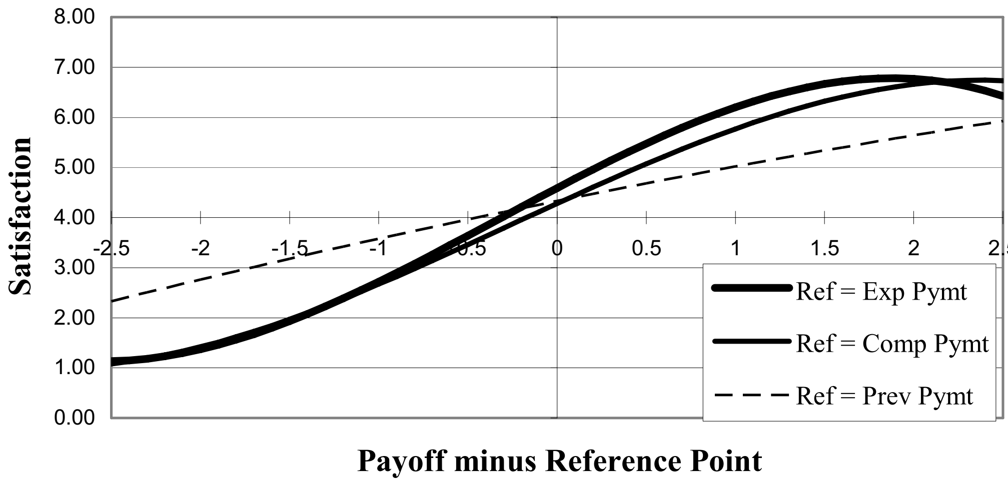

Figure 6(c): Order 5 Satisfaction Functions