# ONLINE APPENDIX TO "UNTIL THE BITTER END: ON PROSPECT THEORY IN A DYNAMIC CONTEXT"

SEBASTIAN EBERT AND PHILIPP STRACK

This supplement details additional results that are alluded to in the main paper. Section W.1 analyzes the important assumption that probability weighting is stronger than loss aversion in more detail and proposes a new weighting function. Section W.2 discusses the case of a power-S-Shape value function which allows for an additional result called skewness preference in the *large*. Section W.3 presents two examples on never stopping in discrete time, one with a finite and one with an infinite investment horizon.

## APPENDIX W.1: THE BENCHMARK WEIGHTING FUNCTION

In this section, we discuss the important Assumption 2 in more detail which says that probability weighting must be strong enough relative to loss aversion. Assumption 2 allows for an intuitive graphical interpretation that yields additional economic insights. We will see that our results do not rely on the distortion of extremely small probabilities. Moreover, even S-shaped weighting functions as observed in some empirical studies may satisfy Assumption 2. At the end of the section, we state a local version of skewness preference in the small (Theorem 1) which illustrates when Assumption 2 can be significantly relaxed.

For $\theta > 0$ we define the function family $b_\theta : [0,1] \longrightarrow [0,1]$ by

$$(1) \qquad b_\theta(p) = \frac{\theta p}{1 - p + \theta p}\,.$$

With this definition Assumption 2 can be restated as: There exists at least one $p \in (0,1)$ such that

1. $w^+(p) > \frac{\lambda p}{1-p+\lambda p} \equiv b_\lambda(p)$ and
2. $w^-(1-p) < \frac{1-p}{1-p+\lambda p} \equiv b_{1/\lambda}(1-p)$.[1]

The functions $b_\theta$ are strictly increasing and continuous with $b_\theta(0) = 0$ and $b_\theta(1) = 1$, and therefore are weighting functions themselves. We refer to $b_\lambda$ and $b_{1/\lambda}(p)$ as the *benchmark weighting function* for gains and losses, respectively. $b_\theta$ intersects with the 45-degree line at zero and one. For $\theta = 1$, $b_\theta$ coincides with the 45-degree line. For $\theta > 1$ ($\theta < 1$), $b_\theta$ is strictly concave (convex) and lies above (below) the 45-degree line. Figure W.1 plots the gain and loss benchmark weighting functions to illustrate Assumption 2. Assumption 2 requires that at least one probability $p$ is "sufficiently" overweighted by $w^+$ in the sense that $w^+(p) > b_\lambda(p) \geq p$ and the complementary probability $1-p$ is "sufficiently" underweighted by $w^-$ in the sense that $w^-(1-p) < b_{1/\lambda}(1-p) \leq 1-p$.

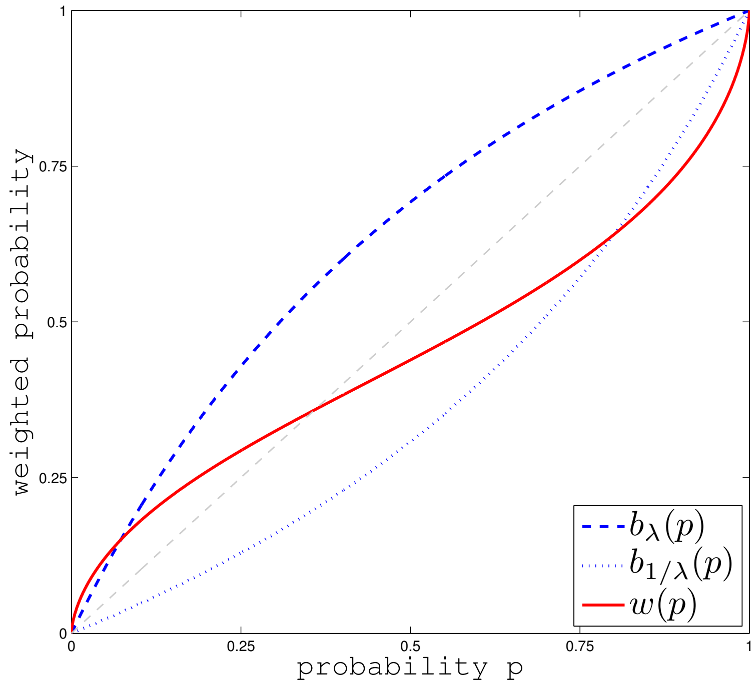
By considering the tangent lines of the gain (loss) benchmark weighting function at zero (at one) one obtains that Assumption 2 follows from the simpler but stronger condition that there exists a $p \in (0,1)$ such that

$$w^+(p) > \lambda p \text{ and } w^-(1-p) < \lambda(1-p) + (1-\lambda).$$

Ebert: Tilburg University, Department of Finance, PO Box 90153, 5000 LE Tilburg, The Netherlands. Phone: +31(0)13 466 3126, Email: S.Ebert@uvt.nl; Strack: University of California at Berkeley, Department of Economics, 513 Evans Hall, Berkeley, 94720 CA, USA; E-mail: philipp.strack@gmail.com.

[1]Note that $b_{1/\lambda}(1-p) = 1 - b_\lambda(p)$ so that condition 2 could also be written as $1 - w^-(1-p) > b_\lambda(p)$.

FIGURE W.1.— Assumption 2 and the Benchmark Weighting Functions




This figure plots the benchmark functions for gains ($b_\lambda(p)$, dashed line) and losses ($b_{1/\lambda}(p)$, dotted line) for $\lambda = 2.25$ with the Tversky-Kahneman weighting function $w(p) = \frac{p^\delta}{(p^\delta+(1-p)^\delta)^{1/\delta}}$ (solid line) for $\delta = 0.65$. The dashed gray line indicates the 45-degree line. $w$ intersects with $b_\lambda(p)$ ($b_{1/\lambda}(p)$) at $p \approx 7.2\%$ ($p \approx 80,0\%$). Therefore, if $w^+ = w^- \equiv w$, then Assumption 2 is satisfied because—for example—$w^+(5\%) > b_\lambda(5\%)$ and $w^-(95\%) < b_{1/\lambda}(95\%)$.

As is apparent from this latter condition, a yet stronger sufficient condition for Assumption 2 is

$$\limsup_{p\to 0} \frac{w^+(p)}{p} > \lambda \text{ and } \limsup_{p\to 0} \frac{1-w^-(1-p)}{p} > \lambda,$$

i.e., the gain (loss) weighting function is steeper than the gain (loss) benchmark weighting function at zero (one). These *limit superior conditions* ensure that Assumption 2 is met for all $p$ close to zero. Intuitively, an arbitrarily small probability $p$ is overweighted by more than the index of loss aversion $\lambda$. If $w^+$ and $w^-$ are differentiable at zero and one, then the limit superior conditions simplify to

$$w^{+'}(0) > \lambda \text{ and } w^{-'}(1) > \lambda,$$

which is equation (3) discussed in the main text. If these derivatives do not exist because they approach infinity (or because $w^+$ has a jump at zero and $w^-$ has a jump at one), the limit superior is infinite. Also in this case, Assumption 2 is met.

As discussed, Assumption 2 is satisfied by the commonly used, inverse-S-shaped weighting functions proposed by Tversky and Kahneman (1992, see Figure W.1), Goldstein and Einhorn (1987), Prelec (1998), and the neo-additive weighting function (Wakker 2010, p. 208), for *all* parameter values. From Figure W.1 it is evident that also S-shaped weighting functions may well satisfy Assumption 2. These weighting functions would satisfy Assumption 2 for large probabilities. A weighting function that does not satisfy Assumption 2

TABLE W.1

PROBABILITIES $p$ FOR WHICH ASSUMPTION 2 IS SATISFIED

| | Tversky-Kahneman | Prelec | Goldstein-Einhorn |
|---|---|---|---|
| $\lambda = 2.25$ | $p \leq 7.2\%$ | $p \leq 6.2\%$ | $p \leq 3.5\%$ |
| $\lambda = 1.5$ | $p \leq 17.6\%$ | $p \leq 13.9\%$ | $p \leq 10.4\%$ |

Table W.1 shows strictly positive probabilities $p$ for which Assumption 2 is satisfied for pronounced loss aversion ($\lambda = 2.25$) and milder loss aversion ($\lambda = 1.5$). Each column is for a different parametric choice of the weighting function, which is taken to be the same for gains and losses ($w^+ \equiv w^-$). The first (second, third) column shows this interval for the Tversky-Kahneman weighting function with parameter $\delta = 0.65$ (Prelec with parameters $b = 1.05$ and $a = 0.65$, Goldstein-Einhorn with parameters $a = 0.69$ and $b = 0.77$). For the functional forms as well as for a motivation of the parameter choices, see Wakker (2010, pp. 206-208).

would be a convex ("pessimistic") weighting function. Rieger and Wang (2006) made a farsighted proposal for a mildly inverse-S-shaped weighting function with finite derivatives at 0 and 1 in order to ensure non-occurrence of the St. Petersburg paradox under CPT. Whether this weighting function satisfies Assumption 2 depends on parameter estimates. Pfiffelmann (2011) argues that a small degree of probability weighting—as featured by the Rieger-Wang function—cannot explain the co-existence of gambling and insurance. That is, when dispensing with pronounced probability weighting as given by the more commonly used weighting functions, CPT may lose its predictive power in contexts where strong skewness preference has desirable effects.

Table W.1 shows for three weighting functions the non-zero probabilities $p$ for which conditions 1 and 2 of Assumption 2 are satisfied. While small, note that these probabilities need not be extremely small. The maximal values displayed in the table are easily obtained as the intersection of $w^+(p)$ with the benchmark weighting function $b_\lambda(p)$, or as the intersection of $w^-(1-p)$ with $b_{1/\lambda}(1-p)$, whichever is smaller. For the example depicted in Figure W.1, $w^+(p)$ intersects with $b(p)$ at approximately 7.2%, which yields the upper left entry of Table W.1.

Additional insight into the importance of Assumption 2 that probability weighting is stronger than loss aversion is obtained by looking at the following local version of skewness preference in the small (Theorem 1). When the value function is smooth except at the reference point as commonly assumed, "probability weighting stronger than loss aversion" is only needed to cover gambling at the reference point. At all other wealth levels any non-trivial amount of probability weighting is sufficient for gambling.

COROLLARY 2 (Local Result) *Under Assumption 1, at some given wealth level $\bar{x}$ there exists an attractive, arbitrarily small zero-mean binary lottery even if Assumption 2 is relaxed by replacing $\lambda = \sup_{x \in \mathbb{R}} \frac{\partial_- U(x)}{\partial_+ U(x)}$ with $\frac{\partial_- U(\bar{x})}{\partial_+ U(\bar{x})}$. If $U$ is differentiable at $\bar{x}$, then Assumption 2 may be further relaxed to: There exists at least one $p \in (0,1)$ such that $w^+(p) > p$ and $w^-(1-p) < 1-p$.*

Proof of Corollary 2. The first statement is evident from the proof of Theorem 1. The second statement follows because $\frac{\partial_- U(\bar{x})}{\partial_+ U(\bar{x})} = 1$ if $U$ is differentiable at $\bar{x}$, and because $b_\lambda(p) = b_{1/\lambda}(p) = p$ for $\lambda = 1$. □

## APPENDIX W.2: THE CASE OF AN S-SHAPED POWER VALUE FUNCTION

In this section, we study in detail the value function originally proposed by Kahneman and Tversky (1979).

Assumption 3 (S-Shaped Power Value Function) *The value function is given by*

$$U(x) = \begin{cases} (x-r)^{\alpha}, & \text{if } x \geq r \\ -\hat{\lambda}\left(-(x-r)\right)^{\alpha}, & \text{if } x < r \end{cases} \tag{2}$$

*with* $\alpha \in (0,1)$ *and* $\hat{\lambda} > 1$.[2]

For this very value function, the Köbberling-Wakker index of loss aversion $\frac{\partial_- U(r)}{\partial_+ U(r)}$ is not well-defined (in particular, it is not equal to $\hat{\lambda}$) because the power function has infinite slope at 0. Therefore, Assumption 1 is not fulfilled, and thus Theorem 1 does not apply. However, we can state a similar result under a modified assumption on the weighting functions.

Assumption 4 *There exists at least one* $p \in (0,1)$ *such that*

1. $w^{+}(p) > \frac{p^{\alpha}\hat{\lambda}}{(1-p)^{\alpha}+\hat{\lambda}p^{\alpha}}$ *and*
2. $w^{-}(1-p) < \frac{(1-p)^{\alpha}}{(1-p)^{\alpha}+\hat{\lambda}p^{\alpha}}$.

The functions $p \mapsto \hat{\lambda}p^{\alpha}/((1-p)^{\alpha}+\hat{\lambda}p^{\alpha})$ and $p \mapsto \hat{\lambda}^{-1}p^{\alpha}/((1-p)^{\alpha}+\hat{\lambda}^{-1}p^{\alpha})$ serve as benchmark weighting functions for the particular case of an S-shaped power value function. These functions are, respectively, similar in shape to the benchmark weighting function $b_{\hat{\lambda}}(p)$ ($b_{1/\hat{\lambda}}(p)$), but lie above $b_{\hat{\lambda}}(p)$ (below $b_{1/\hat{\lambda}}(p)$) for $\alpha \in (0,1)$. Thus Assumption 4 is stronger than Assumption 2. Nevertheless, Assumption 4 is met by the weighting functions of Tversky and Kahneman (1992) and Goldstein and Einhorn (1987) under parameter restrictions that are fulfilled according to most empirical studies; see Azevedo and Gottlieb (2012) for an elaboration.[3] For the weighting function of Prelec (1998), Assumption 4 is always met.

Table W.2 shows, analogously to Table W.1, probabilities for which the conditions in Assumption 4 are fulfilled. We see that these probabilities are smaller than those in Table W.1, which evidences that Assumption 2 is tighter than Assumption 4. However, in all but one case these probabilities may still be larger than two percent.

Theorem 3 (Skewness Preference in the Small for the S-Shaped Power Value Function) *Assume Assumptions 3 and 4. For every wealth level there exists an attractive, zero-mean binary lottery that is arbitrarily small.*

Proof of Theorem 3. Since $U$ is differentiable everywhere except at $r$, the result for $x \neq r$ follows from Corollary 2. Thus suppose $x = r$ and let $a_n = x - p/n$ and $b_n = x + (1-p)/n$ (as in the proof of Theorem 1). For the power-S-shaped value function it is easily seen that

$$\frac{U(x)-U(a_n)}{x-a_n} = \frac{0+\hat{\lambda}\left(\frac{p}{n}\right)^{\alpha}}{\frac{p}{n}} = \hat{\lambda}n^{1-\alpha}p^{\alpha-1}$$

and

$$\frac{U(b_n)-U(a_n)}{b_n-a_n} = \frac{\left(\frac{1-p}{n}\right)^{\alpha}+\hat{\lambda}\left(\frac{p}{n}\right)^{\alpha}}{\frac{1-p}{n}+\frac{p}{n}} = n^{1-\alpha}\left((1-p)^{\alpha}+\hat{\lambda}p^{\alpha}\right)$$

[2] We do not consider the ill-posed specification of utility with different power parameters for gains and losses. In that case, $\hat{\lambda}$ does not capture loss aversion in a meaningful way as is illustrated by Wakker (2010, pp. 267-270).

[3] Sufficient conditions for Assumption 4 are—similarly to the general case—$w^{+}(p) > p^{\alpha}\hat{\lambda}$ and $w^{+}(1-p) > p^{\alpha}\hat{\lambda}$, and also $\limsup_{p\to 0}\frac{w^{+}(p)}{p^{\alpha}} > \hat{\lambda}$ and $\limsup_{p\to 0}\frac{1-w^{-}(1-p)}{p^{\alpha}} > \hat{\lambda}$. Azevedo and Gottlieb (2012) discuss a condition similar to these limit superior conditions.

TABLE W.2

PROBABILITIES $p$ FOR WHICH ASSUMPTION 4 IS SATISFIED

| | Tversky-Kahneman | Prelec | Goldstein-Einhorn |
|---|---|---|---|
| $\lambda = 2.25$ | $p \leq 2.4\%$ | $p \leq 2.8\%$ | $p \leq 0.35\%$ |
| $\lambda = 1.5$ | $p \leq 10.1\%$ | $p \leq 7.6\%$ | $p \leq 2.9\%$ |

*Notes.* Table W.2 shows strictly positive probabilities $p$ for which Assumption 4 is satisfied for pronounced loss aversion ($\lambda = 2.25$) and milder loss aversion ($\lambda = 1.5$) when $\alpha = 0.88$. Each column is for a different parametric choice of the weighting function, which is taken to be the same for gains and losses ($w^+ \equiv w^-$). The first (second, third) column is for the Tversky-Kahneman weighting function with parameter $\delta = 0.65$ (Prelec with parameters $b = 1.05$ and $a = 0.65$, Goldstein-Einhorn with parameters $a = 0.69$ and $b = 0.77$). For the functional forms as well as for a motivation of the parameter choices, see Wakker (2012, pp. 206-208).

such that

$$\frac{\frac{U(x)-U(a)}{x-a}}{\frac{U(b)-U(a)}{b-a}} = \hat{\lambda}\frac{p^{\alpha-1}}{\left((1-p)^\alpha + \hat{\lambda}p^\alpha\right)}.$$

Therefore, according to equation (A4) in the main text, $L(p, b_n, a_n)$ is attractive for large enough $n$ if

$$0 < \left(\frac{w^+(p)}{p^\alpha} - \hat{\lambda}\frac{1}{\left((1-p)^\alpha + \hat{\lambda}p^\alpha\right)}\left(w^-(1-p) + w^+(p)\right)\right) \Longleftrightarrow w^+(p) > \hat{\lambda}p^\alpha\frac{w^-(1-p)}{(1-p)^\alpha}.$$

The proof concludes analogously to that of case 3 of Theorem 1 by employing Assumption 4. □

With this result at hand, for the model outlined in Section III we obtain the never-stopping result for CPT with a power-S-shaped value function.

THEOREM 4 *Under Assumptions 3 and 4, the naïve CPT agent never stops.*

Proof of Theorem 4. The proof is analogous to that of Theorem 2 except for applying Theorem 3 instead of Theorem 1. □

Power utility is differentiable everywhere except at the reference point. Therefore, note that Corollary 2, which assumes just minimal probability weighting, also applies to power utility whenever we are not at the reference point. Therefore, we need Assumption 4 exclusively to cover gambling at the reference point.

We called Theorems 1 and 3 skewness preference in the *small*. However, attractive risks may also be large. Recall Assumptions 1 and 2 and note that these are global assumptions that ensure gambling at any wealth level. In particular, they ensure gambling at wealth levels where utility is very concave, which supports risk-aversion. In other words, Assumption 2 on probability weighting is a sufficient condition for gambling at any wealth level, and for *whatever* value function with loss aversion parameter $\lambda$. The maximal size and skewness of attractive risks obviously depends on the specific choice of the value function and its shape at current wealth.

To illustrate this point, let us mention a result on *large* skewed attractive risks for the S-shaped power function. It is obtained as a combination of our result with an observation made recently by Azevedo and Gottlieb (2012). The authors consider a prospect theory agent with S-shaped power value function and whose reference point equals current wealth. Azevedo and Gottlieb (2012) show that for any attractive zero-mean binary gamble $L$, the multiple $cL$ $(c > 1)$ is also attractive. We obtain the following corollary.

Corollary 3 (Skewness Preference in the Small and in the Large for the S-Shaped Power Value Function) *Assume Assumptions 3 and 4. If the decision maker's reference point equals current wealth, then there exists an attractive, zero-mean binary lottery of arbitrary (large or small) size.*

Proof of Corollary 3. According to Theorem 3 there exists an attractive, arbitrarily small binary risk with mean zero. According to Azevedo and Gottlieb (2012, p. 1294) it can be scaled up to any size. □

## APPENDIX W.3: TWO EXAMPLES FOR NEVER STOPPING IN DISCRETE TIME

Our main result was formulated as a stopping theorem in continuous time. In this section, we illustrate through two examples that the continuous time setup is not crucial to our paper as discussed in Section V, but that the richness of the investment opportunity is. The first example is a never-stopping result in discrete, infinite time with symmetric basic gambles. However, a flexible stake size and an infinite investment horizon allow for generating small and skewed gambles, and thus the agent never stops. The second example is the casino model in Barberis (2012) where the casino offers French roulette so that the gambler faces a skewed basic gamble. Even though time is discrete and finite and the stake size is not small, the naïve CPT investor never stops.

### W.3.1. *Never stopping in infinite discrete time*

Consider a binomial random walk $(X_t)_{t\in\mathbb{N}}$ with jump size one and equal probability for up- and down movements. At every point in time $t$ the agent can choose the stakesize $s_t \in [0,1]$ (as a fraction of his wealth $y_t$) to bet. The evolution of his wealth is then given by

$$y_{t+1} = y_t + s_t y_t (X_{t+1} - X_t)$$

with initial wealth $y_0 > 0$. The following strategy (of choosing $s_t$) results in any given fair binary lottery $L(p,b,a)$. Choose $s_t$ maximal such that $y_{t+1} \in [a,b]$, i.e.,

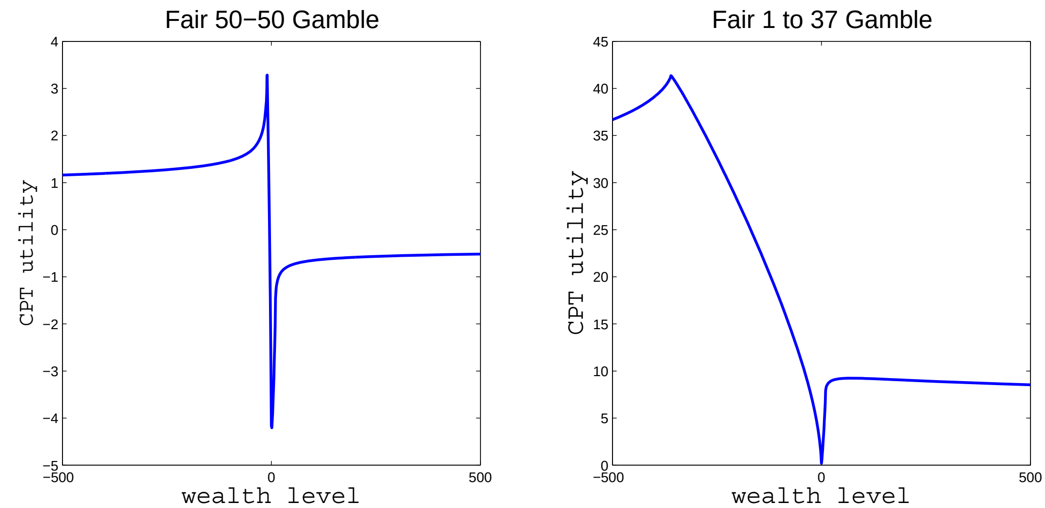
$$s_t = \max\{\tilde{s} \in [0,1] : (1-\tilde{s})y_t \geq a \text{ and } (1+\tilde{s})y_t \leq b\} = \min\{1 - \frac{a}{y_t}, \frac{b}{y_t} - 1\}\,.$$

Due to the martingale property it follows from Doob's optional sampling theorem that the probabilities of hitting $b$ and $a$ that are induced by this strategy are fair, i.e., are $p$ and $1-p$, respectively. Therefore, if $L(p,b,a)$ is attractive according to either Theorem 1 or 3, then the agent will gamble. Since an attractive lottery exists at any wealth level (i.e., at any time $t$) the agent never stops.

### W.3.2. *Never stopping in a casino that offers skewed gambles*

Consider a casino that offers a fair version of French Roulette. We assume a fair casino to be close to the model of Barberis (2012). Then the basic gamble considered by Barberis is the fair analogue to a bet on Red or Black, which occur with equal probability. Now suppose the agent can also bet on a single number, which occurs with probability $\frac{1}{37}$. Consider an agent who only considers betting 10 units of money on a single number. He is not even able to form a gambling strategy over several periods. This implies a rather coarse strategy space, a feature which is working against our never-stopping result. However, the basic gamble is skewed whereas the basic gamble in Barberis (2012) is symmetric. Let $(X_t)_{t\in\mathbb{R}_+}$ be the binomial random

FIGURE W.2.— Gambling utility for a symmetric and a skewed gamble




This figure shows the excess utility an agent gains from gambling (over not gambling) for different wealth levels at which the gamble is evaluated. The left panel shows the utility from gambling a fair 50-50 bet, while the right panel shows the utility from gambling a fair 1 to 37 bet. The agent is a CPT maximizer with the parametrization of Tversky and Kahneman (1992) with parameters given by $\alpha = 0.88$, $\delta = 0.65$, and $\lambda = 2.25$. The agent's reference point is 0.

walk that represents the gambler's wealth. It increases by 360 with probability $\frac{1}{37}$ and decreases by 10 with probability $\frac{36}{37}$, starting at some level $X_0 \in \mathbb{R}$, i.e.,

$$\mathbb{P}(X_{t+1} = X_t + 360) = \frac{1}{37} \text{ and } \mathbb{P}(X_{t+1} = X_t - 10) = \frac{36}{37}.$$

The agent is forced to stop in the final period $T$, which is exogenous, or if the random walk reaches zero.

Suppose the agent has CPT preferences given by the original parametrization of Tversky and Kahneman (1992) with parameters as estimated by the authors. Figure W.2 plots the excess utility from gambling for the two basic gambles described above, as a function of current wealth. For the 50-50 gamble (left panel), gambling is attractive over the area of losses, and unattractive at the reference point and thereafter. This fits with the common intuition of risk-seeking over losses and risk-aversion over gains, which is induced by the S-shaped value function.

Note that the probability weighting component has no grip when evaluating 50-50 gambles. However, the right panel shows that gambling the skewed basic gamble is attractive *everywhere*. The lowest utility from gambling is at the reference point, but this utility is still positive (approximately +0.56). Therefore, at any node of the binomial tree, the agent will want to gamble. That is, the agent never stops even though we have finite time with an *arbitrary* number of gambling periods and a rather limited strategy space. Only one basic gamble is available, but this gamble is sufficiently skewed and small to be attractive to this very CPT agent. A stop-loss plan would grant even higher utility to the agent, but the one-shot gamble is attractive in itself already.

## REFERENCES


AZEVEDO, E. AND D. GOTTLIEB (2012): "Risk-Neutral Firms Can Extract Unbounded Profits from Consumers with Prospect Theory Preferences," *Journal of Economic Theory*, 147, 1291–1299.

Barberis, N. (2012): "A Model of Casino Gambling," *Management Science*, 58, 35–51.

Goldstein, W. and H. Einhorn (1987): "Expression Theory and the Preference Reversal Phenomena," *Psychological Review*, 94, 236–254.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–292.

Pfiffelmann, M. (2011): "Solving the St. Petersburg Paradox in Cumulative Prospect Theory: the Right Amount of Probability Weighting," *Theory and Decision*, 71, 325–341.

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

Rieger, M. and M. Wang (2006): "Prospect Theory and the St. Petersburg Paradox," *Economic Theory*, 28, 665–679.

Tversky, A. and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P. (2010): *Prospect Theory for Risk and Ambiguity*, Cambridge University Press.