Comment on "Risk Preferences Are Not Time Preferences": Balancing on a Budget Line

Appendix: For Online Publication

Thomas Epper and Helga Fehr-Duda

October 27, 2014

A1. Aggregating Time and Risk

Being atemporal by construction, non-expected-utility preference models are silent on the way time and risk are to be aggregated when both risk and delay are present. Figure 1 illustrates three principle possibilities, the portfolio, the separable and the recursive cases.

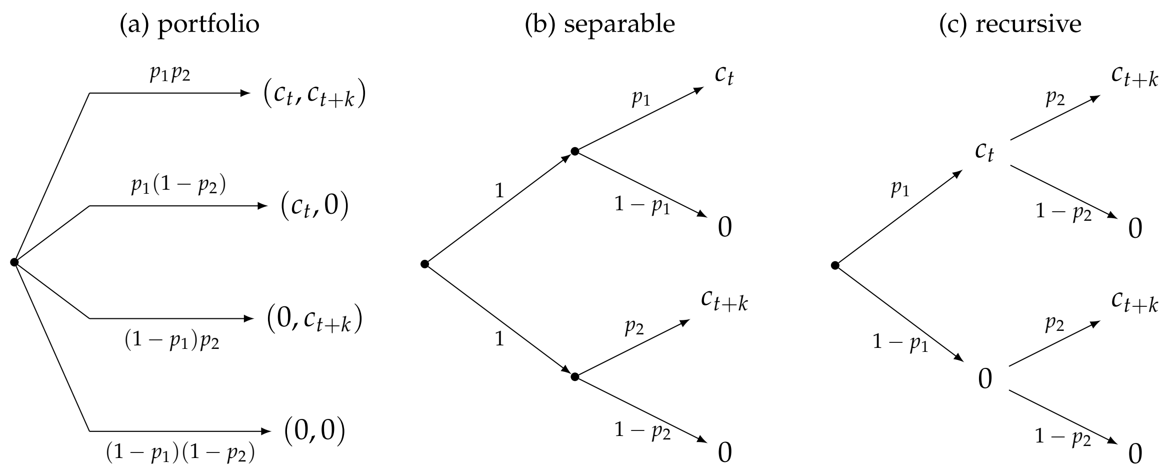
(a) The portfolio case

Panel (a) refers to the case where the prospect $P(p_1, p_2)$ is evaluated as a lottery over consumption streams, as commonly assumed in the literature on intertemporal risk preferences (see, for example, Chew and Epstein (1990)), i.e. the decision maker computes the probability distribution of the discounted utilities of the consumption streams and then evaluates this distribution according to her risk preferences. The portfolio case is discussed in detail in the manuscript where we show that all the key AS findings can be accommodated by rank-dependent probability weighting.

(b) The separable case

Panel (b) of Figure 1 depicts the case underlying the argument in AS. Here, subjects treat sooner and later consumption as two separate temporal prospects that are discounted for risk first and then aggregated over time. Recalling that $v(0) = 0$,

Figure 1: Different Representations of $P(p_1, p_2)$



Panel (a): $P(p_1, p_2) = ((c_t, c_{t+k}), p_1p_2; (c_t, 0), p_1(1-p_2); (0, c_{t+k}), (1-p_1)p_2; (0,0), (1-p_1)(1-p_2))$. The prospect is viewed as a lottery over consumption streams. The probability distribution corresponds to four states of nature generated by two independent random devices.
Panel (b): $P(p_1, p_2) = (c_t, p_1; 0, 1-p_1) + (c_{t+k}, p_2; 0, 1-p_2)$. The prospect is viewed as the sum of two separate temporal prospects that are devalued for risk first and subsequently for delay.
Panel (c): $P(p_1, p_2) = (c_t + (c_{t+k}, p_2; 0, 1-p_2), p_1; 0 + (c_{t+k}, p_2; 0, 1-p_2), 1-p_1)$. The prospect is viewed as a two-stage lottery where uncertainty resolves sequentially.

this procedure renders a total value of

$$\begin{aligned} V_{SEP}(P(p_1, p_2)) =& [g(p_1)v(c_t) + (1-g(p_1))v(0)]\delta(t) \\ &+ [g(p_2)v(c_{t+k}) + (1-g(p_2)v(0)]\delta(t+k) \\ =& g(p_1)v(c_t)\delta(t) + g(p_2)v(c_{t+k})\delta(t+k), \end{aligned} \tag{1}$$

which entails the first-order condition

$$FOC_{SEP}(P(p_1, p_2)) : \frac{v'(c_t)}{v'(c_{t+k})} = (1+r)\frac{\delta(t+k)}{\delta(t)}\frac{g(p_2)}{g(p_1)}. \tag{2}$$

If subjects are prone to atemporal common-ratio violations, i.e. if their probability weights are subproportional, $\frac{g(p_2)}{g(p_1)} < \frac{g(\lambda p_2)}{g(\lambda p_1)}$ holds for $0 < \lambda < 1$ and $p_2 < p_1$. Therefore, sooner consumption in the original conditions is higher or lower than in the scaled-down conditions depending on the relative magnitudes of p_1 and p_2. Hence, in the separable case, probability weighting can accommodate the AS findings in the two differential risk conditions, but not the cross-over result for $p_1 = p_2$, for which identical behavior is predicted.

(c) The recursive case

Panel (c) in Figure 1 interprets $P(p_1, p_2)$ as a two-stage prospect. If subjects evaluate the prospect recursively,

$$\begin{aligned} V_{REC}(P(p_1, p_2)) =& g(p_1)[v(c_t) + g(p_2)v(c_{t+k})\delta(k)]\delta(t) \\ &+(1 - g(p_1))[v(0) + g(p_2)v(c_{t+k})\delta(k)]\delta(t) \\ =& g(p_1)v(c_t)\delta(t) + g(p_2)v(c_{t+k})\delta(t)\delta(k), \end{aligned} \tag{3}$$

resulting in the first-order condition

$$FOC_{REC}(P(p_1, p_2)) : \frac{v'(c_t)}{v'(c_{t+k})} = (1 + r)\delta(k)\frac{g(p_2)}{g(p_1)}. \tag{4}$$

which is identical to Equation (1) if exponential discounting is assumed. Consequently, in this case the recursive procedure entails the same predictions as in the separable case which are, at least partially, at odds with the AS evidence.

To sum up: Contrary to approach (a), methods (b) and (c) predict behavior correctly only in the differential risk conditions for subproportional probability weights, but fail to predict the cross-over result in the baseline conditions.

A2. Time-Dependent Probability Weighting and Hyperbolic Discounting

In the following, we take a closer look at the predictions of time-dependent probability weighting (Halevy, 2008; Saito, 2011) in the context of the AS experiment. In Halevy's model time-dependence of probability weights arises from the inherent uncertainty of the future. If decision makers have doubts whether promised rewards will actually arrive they may perceive allegedly certain future payments as risky. Furthermore, a decision maker who is prone to probability weighting will distort the subjective probability of non-payment. Halevy shows that a specific characteristic of atemporal probability weights, increasing elasticity,[1] generates hyperbolic discounting. Epper and Fehr-Duda

[1]Increasing elasticity is equivalent to subproportionality.

(2012) extend this idea to risky prospects and demonstrate that the decision maker will appear more risk tolerant for future prospects than for present ones. However, this mechanism is effective only if the resolution of uncertainty coincides with the timing of payment. If uncertainty is resolved immediately the decision maker will know for sure which one of the outcomes is supposed to arrive and, consequently, time-dependent probability weighting solely affects allegedly certain amounts of money, i.e. it will generate hyperbolic discounting but not increasing risk tolerance.

In the AS experiment all uncertainty is indeed resolved during the experiment and, consequently, any additional risk from inherent future uncertainty only affects discounting but not risk tolerance. Therefore, time-dependent probability weights $g(p,t)$ are identical to the atemporal case. But since future payments are still inherently uncertain, they are discounted by an additional factor $g(1,t)$, which is independent of the probabilities p_1 and p_2. In general, $g(p,0)g(1,t) \neq g(p,t)$ holds for the type of probability weighting functions studied in Halevy (2008) and Epper and Fehr-Duda (2012). For example, the appropriate first-order conditions for $(p_1, p_2) \in \{(1,1),(0.5,0.5)\}$ are given by

$$FOC_{TEMP}(P(1,1)) : \frac{v'(c_t)}{v'(c_{t+k})} = (1+r)\frac{\delta(t+k)}{\delta(t)}\frac{g(1,t+k)}{g(1,t)}, \tag{5}$$

and

$$FOC_{TEMP}(P(0.5,0.5)) : \frac{v'(c_t)}{v'(c_{t+k})} = \begin{cases} (1+r)\frac{\delta(t+k)}{\delta(t)}\frac{g(1,t+k)}{g(1,t)}\frac{g(0.25,0)+g(0.75,0)-g(0.5,0)}{g(0.5,0)} & \text{if } c_t \succ c_{t+k} \\ (1+r)\frac{\delta(t+k)}{\delta(t)}\frac{g(1,t+k)}{g(1,t)}\frac{g(0.5,0)}{g(0.25,0)+g(0.75,0)-g(0.5,0)} & \text{if } c_t \prec c_{t+k} \end{cases} \tag{6}$$

The ratio of additional discount weights $\frac{g(1,t+k)}{g(1,t)}$ is generally smaller than one, which leads to a less pronounced decline of c_t^* than in the standard time-independent case. All the other predictions of the RDU model remain qualitatively unchanged.

A3. The Kihlstrom-Mirman Model and the CTB Findings

Besides RDU several other models can accommodate intertemporal risk aversion as encountered in the AS experiment. Bommier (2007) discusses two examples, the Kihlstrom and Mirman (1974) model (see also Epstein and Tanny (1980)) and the class of recursive preferences studied by Epstein (1983). Here we take a closer look at the Kihlstrom-Mirman approach. Assuming expected utility and positing a concave transformation of the discounted utility of consumption streams $\phi(\tilde{c}_t, \tilde{c}_{t+k}) = \phi(v(\tilde{c}_t)\delta(t) + v(\tilde{c}_{t+k})\delta(t + k))$, with $\phi(0) = 0, \phi' > 0, \phi'' < 0$, to model intertemporal risk aversion, an approach also taken by Cheung (forthcoming), renders the following general first-order condition for $P(p_1, p_2)$

$$FOC_{KM}(P(p_1,p_2)) : \frac{v'(c_t)}{v'(c_{t+k})} = (1+r)\frac{\delta(t+k)}{\delta(t)}\frac{\phi'(c_t,c_{t+k}) + \frac{1-p_1}{p_1}\phi'(0,c_{t+k})}{\phi'(c_t,c_{t+k}) + \frac{1-p_2}{p_2}\phi'(c_t,0)}.$$

It nests DEU if ϕ is linear. In the following, we discuss its predictions for the prototypical baseline and differential risk conditions of the AS experiment.

- **The baseline conditions**: $P(1, 1)$ versus $P(\lambda, \lambda)$.
  Inserting $p_1 = p_2 = \lambda$ into the equation above renders for the ratio of the marginal intertemporal utilities

  $$\frac{\phi'(c_t,c_{t+k}) + \frac{1-\lambda}{\lambda}\phi'(0,c_{t+k})}{\phi'(c_t,c_{t+k}) + \frac{1-\lambda}{\lambda}\phi'(c_t,0)} \gtreqless 1 \text{ iff } \phi'(0,c_{t+k}) \gtreqless \phi'(c_t,0). \tag{7}$$

  Recall from our analysis in Section 2 of the paper that optimal sooner consumption decreases in the ratio of the marginal utilities on the right side of the first-order condition. Comparing Equation (7) with the certain case where this ratio is equal to one, shows that the optimal sooner consumption c_t^* will be lower (higher) in the risky condition than in the certain condition if $c_t \succ c_{t+k}$ ($c_t \prec c_{t+k}$) due to the concavity of ϕ. Therefore, $c_t^*(\lambda, \lambda)$ is less sensitive to changes in r, intersecting the consumption curve for the certain condition, qualitatively similar to the RDU-case. The Kihlstrom-Mirman model differs from RDU in that it involves a smooth transition from sooner to later consumption levels whereas RDU predicts a sudden change in decision weights. Which one of the models fits better is an empirical question that still awaits investigation. However, modeling the preference for diversification in the spirit of Kihlstom-Mirman is based on the (clearly refuted) assumption that atemporal risk taking conforms to expected utility. In contrast, RDU can accommodate both, atemporal and intertemporal risk aversion. Moreover, in the atemporal context, RDU abandons the identity of diminishing marginal utility and risk aversion (Wakker, 1994).

- **The differential risk conditions**: $P(1, q)$ versus $P(\lambda, \lambda q)$ for some $0 < \lambda < 1$ and some probability $q < 1$.

Comparing the corresponding first-order conditions for $P(1, q)$ and $P(\lambda, \lambda q)$ yields the following relationships after some rearranging

$$c_t^*(\lambda, \lambda q) \gtreqqless c_t^*(1, q) \text{ iff } \frac{\phi'(c_t, c_{t+k})}{q\phi'(c_t, c_{t+k}) + (1-q)\phi'(c_t, 0)} \lesseqqgtr \frac{\phi'(0, c_{t+k})}{\phi'(c_t, 0)}.$$

It is interesting to note that these conditions do not depend on the scaling factor λ. Because of the concavity of ϕ, $\phi'(c_t, c_{t+k}) < q\phi'(c_t, c_{t+k}) + (1-q)\phi'(c_t, 0) < \phi'(c_t, 0)$ holds and, therefore, the left ratio of marginal utilities is always smaller than one. If $c_t \succ c_{t+k}$ holds, then $\phi'(c_t, 0) < \phi'(0, c_{t+k})$ is satisfied. In this case, the right ratio is greater than one, implying that $c_t^*(1, q) > c_t^*(\lambda, \lambda q)$. In the opposite case, $c_t \prec c_{t+k}$, both sides are less than one, and it depends on ϕ' as well as on q whether $c_t^*(1, q)$ may be smaller than $c_t^*(\lambda, \lambda q)$ and, hence, whether a cross-over of the consumption curves may occur. Due to the high interest rates offered in the AS experiment, it is likely that the discounted utility of later consumption is much greater than the discounted utility of sooner consumption, which implies that $\phi'(0, c_{t+k}) \ll \phi'(c_t, 0)$ and, therefore, their ratio is much smaller than one. For values of q in the vicinity of one, on the other hand, the left side of the equation is close to one. Consequently, for comparatively high values of q and high interest rates, it is likely that sooner consumption in the scaled-down condition is higher than in the original one.[2] In other words, under these specific circumstances that are likely to hold in the AS experiment, the model predicts a cross-over in the differential risk conditions as well and, hence, is unable to accommodate the intertemporal common-ratio effect evident in subjects' behavior.

A4. Probability Weighting and Intertemporal Risk Aversion

When decision makers evaluate risky consumption streams they often seem to have a preference for diversifying consumption across time, i.e. they prefer some good and some bad to all or nothing. Richard (1975) labeled such a preference "‘multivariate risk aversion", which is termed "intertemporal risk aversion" in the context of this paper.[3] Since discounted expected utility is additively separable it cannot accommodate

[2] The difference in the ratios is particularly pronounced when ϕ' is convex, i.e. if the decision maker exhibits intertemporal prudence.

[3] Bommier (2007) and Denuit, Eeckhoudt and Rey (2010), among others, use the term "correlation aversion" instead of intertemporal or multivariate risk aversion.

intertemporal risk aversion. One way of solving the problem, therefore, is to give up additive separability, a path followed for instance by Kihlstrom and Mirman (1974), Epstein and Tanny (1980) and Epstein (1983) (see also the discussion in A3.). In the following we show that rank-dependent utility implies intertemporal risk aversion if the decision maker is sufficiently pessimistic. To our knowledge, this link between intertemporal risk aversion and probability weighting has not been noted before. Hence, RDU is a descriptively valid model for atemporal decisions and, under appropriate assumptions on the aggregation of time and risk, for intertemporal risk taking behavior as well.[4]

The concept of *intertemporal risk aversion* indicates whether an individual dislikes consumption at different points in time to be positively correlated (Epstein and Tanny, 1980; Denuit, Eeckhoudt and Rey, 2010). Consider two different consumption levels at t, $\underline{c}_t, \overline{c}_t$ and two different consumption levels at $t+k$, $\underline{c}_{t+k}, \overline{c}_{t+k}$. Intertemporal risk aversion is defined as follows (Richard, 1975; Bommier, 2007):

An individual is *intertemporally risk averse* if and only if, for all $\underline{c}_t, \overline{c}_t, \underline{c}_{t+k}, \overline{c}_{t+k}$, such that $\overline{c}_t > \underline{c}_t$ and $\overline{c}_{t+k} > \underline{c}_{t+k}$,

$$P_{neg} = ((\overline{c}_t, \underline{c}_{t+k}), 0.5; (\underline{c}_t, \overline{c}_{t+k}), 0.5) \succ P_{pos} = ((\overline{c}_t, \overline{c}_{t+k}), 0.5; (\underline{c}_t, \underline{c}_{t+k}), 0.5).$$

P_{neg} represents a lottery over consumption streams that entail one high consumption level either sooner or later. Whatever the state of the world, one of the higher consumption levels, $\overline{c}_t$ or $\overline{c}_{t+k}$, will materialize. In P_{pos}, only the higher consumption levels or only the lower consumption levels will materialize in both periods. Therefore, there is an equal chance to end up with the worst consumption stream $(\underline{c}_t, \underline{c}_{t+k})$, an outcome many people are averse to. Consequently, an intertemporally risk averse individual prefers the prospect guaranteeing high consumption either sooner or later over the all-high/all-low prospect.

According to RDU the prospect P_{neg} is evaluated as

$$V_{RDU}(P_{neg}) = \begin{cases} g(0.5)[v(\overline{c}_t)\delta(t) + v(\underline{c}_{t+k})\delta(t+k)] + (1-g(0.5))[v(\underline{c}_t)\delta(t) + v(\overline{c}_{t+k})\delta(t+k)] & \text{or} \\ (1-g(0.5))[v(\overline{c}_t)\delta(t) + v(\underline{c}_{t+k})\delta(t+k)] + g(0.5)[v(\underline{c}_t)\delta(t) + v(\overline{c}_{t+k})\delta(t+k)], \end{cases}$$

depending on the relative magnitudes of $v(\overline{c}_t)\delta(t) + v(\underline{c}_{t+k})\delta(t+k)$ and $v(\underline{c}_t)\delta(t) + v(\overline{c}_{t+k})\delta(t+k)$. For P_{pos} we obtain

$$V_{RDU}(P_{pos}) = g(0.5)[v(\overline{c}_t)\delta(t) + v(\overline{c}_{t+k})\delta(t+k)] + (1-g(0.5))[v(\underline{c}_t)\delta(t) + v(\underline{c}_{t+k})\delta(t+k)].$$

[4]Chew and Epstein (1990) discuss the usefulness of rank-dependent utility for disentangling the coefficient of relative risk aversion and the elasticity of intertemporal substitution.

Examining the difference between these two values renders

$$V_{RDU}(P_{neg}) - V_{RDU}(P_{pos}) = \begin{cases} (1-2g(0.5))[v(\bar{c}_{t+k}) - v(\underline{c}_{t+k})]\delta(t+k) & \text{or} \\ (1-2g(0.5))[v(\bar{c}_t) - v(\underline{c}_t)]\delta(t). \end{cases}$$

Because $\bar{c}_t > \underline{c}_t$ and $\bar{c}_{t+k} > \underline{c}_{t+k}$ hold, the difference in prospect values is positive iff $1 - 2g(0.5) > 0$. In other words, the individual is intertemporally risk averse if $g(0.5) < 0.5$, which is satisfied for representative probability weights (Fehr-Duda and Epper, 2012). If $g(0.5) > 0.5$, the decision maker is intertemporally risk seeking, and if $g(0.5) = 0.5$, she is intertemporally risk neutral. Therefore, RDU preferences imply intertemporal risk aversion if the individual is sufficiently pessimistic at the probability $p = 0.5$.

In the context of the AS baseline conditions, intertemporal risk aversion manifests itself as a preference of $P_{neg} = ((c_t, 0), 0.5; (0, c_{t+k}), 0.5)$ over $P_{pos} = ((c_t, c_{t+k}), 0.5; (0, 0), 0.5)$. However, in the AS experiment, subjects do not have a choice between these prospects over consumption streams directly but evaluate

$$P = ((c_t, c_{t+k}), 0.25; (c_t, 0), 0.25; (0, c_{t+k}), 0.25; (0, 0), 0.25)\,,$$

generated by the independent risks of receiving the sooner payment and later payments. Richard (1975) shows that, in the realm of expected utility, intertemporal risk aversion can be defined in an entirely equivalent way by positing a preference of P over P_{pos}: "*To emphasize the risk averse aspect of this behavior just consider an extension [...] for the decision maker's entire lifetime. It seems reasonable to define risk averse behavior as preferring an independent lottery each year to one lottery for lifelong misery or happiness*" (Richard (1975), p. 14). In RDU, $P \succ P_{pos}$ requires $g(0.25) + g(0.75) > 2g(0.5)$ to hold, whereas $P_{neg} \succ P_{pos}$ requires $2g(0.5) < 1$. Both conditions are likely to hold for the average decision maker and, hence, RDU implies intertemporal risk aversion in this case as well.

References

Bommier, A. 2007. "Risk aversion, intertemporal elasticity of substitution and correlation aversion." *Economics Bulletin*, 4(29): 1–8.

Cheung, Stephen. forthcoming. "On the Elicitation of Time Preference under Conditions of Risk." *American Economic Review*.

Chew, SH, and Larry Epstein. 1990. "Nonexpected Utility Preferences in a Temporal Framework with an Application to Consumption-Savings Behaviour." *Journal of Economic Theory*, 50: 54–81.

Denuit, Michel, Louis Eeckhoudt, and Beatrice Rey. 2010. "Some Consequences of Correlation Aversion in Decision Science." *Annals of Operations Research*, 176: 259–269.

Epper, Thomas, and Helga Fehr-Duda. 2012. "The Missing Link: Unifying Risk Taking and Time Discounting." *Working Paper, Department of Economics, University of Zurich*, , (ECON96).

Epstein, Larry, and S Tanny. 1980. "Increasing generalized correlation: a definition and some economic consequences." *Canadian Journal of Economics*, 13(1): 16–34.

Epstein, Larry G. 1983. "Stationary Cardinal Utility and Optimal Growth under Uncertainty." *Journal of Economic Theory*, 31: 133–152.

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

Halevy, Yoram. 2008. "Strotz Meets Allais: Diminishing Impatience and Certainty Effect." *American Economic Review*, 98(3): 1145–1162.

Kihlstrom, Richard E, and Leonard J Mirman. 1974. "Risk Aversion with Many Commodities." *Journal of Economic Theory*, 8: 361–388.

Richard, Scott F. 1975. "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions." *Management Science*, 22(1): 12–21.

Saito, K. 2011. "A Relationship between Risk and Time Preferences." *American Economic Review*, 101(5): 2271–2275.

Wakker, Peter. 1994. "Separating Marginal Utility and Probabilistic Risk Aversion." *Theory and Decision*, 36: 1–44.