9-25-2016

Dual Process Utility Theory: A Model of Decisions Under Risk and Over Time

Mark Schneider
Chapman University

Recommended Citation
Schneider, M. (2016). Dual-process utility theory: A model of decisions under risk and over time. ESI Working Paper 16-23. Retrieved from http://digitalcommons.chapman.edu/esi_working_papers/201

Dual Process Utility Theory: A Model of Decisions Under Risk and Over Time

Comments
Working Paper 16-23

DUAL PROCESS UTILITY THEORY:
A MODEL OF DECISIONS UNDER RISK AND OVER TIME*

BY MARK SCHNEIDER

This Version: January 25, 2018

ABSTRACT. The Discounted Expected Utility model has been a major workhorse for analyzing individual behavior for over half a century. However, it cannot account for evidence that risk interacts with time preference, that time interacts with risk preference, that many people are averse to timing risk and do not discount the future exponentially, that discounting depends on the magnitude of outcomes, that risk preferences are not time preferences, and that risk and time preferences are correlated with cognitive reflection. Here we address these issues in a decision model based on the interaction of an affective and a reflective valuation process. The resulting Dual Process Utility theory provides a unified approach to modeling risk preference, time preference, and interactions between risk and time preferences. It also provides a unification of models based on a rational economic agent, models based on prospect theory or rank-dependent utility, and dual system models of decision making. We conclude by demonstrating that a simple extension of the model provides a unified approach to explaining empirical violations of 'dimensional independence' across the domains of risk, time, and social preferences. Such violations of dimensional independence challenge the leading normative and behavioral models of decision making.

Keywords: Risk preferences; Time preferences; Social preferences; System 1; System 2

JEL Codes: D01, D03, D81, D90.

* Economic Science Institute, Chapman University. One University Drive, Orange, CA 92866. (Phone: (714) 744-7083; Fax: (714) 628-2881; e-mail: schneide@chapman.edu). Acknowledgments: I thank Manel Baucells, Cary Deck, Enrico Diecidue, Manuel Nunez, Jeff Schneider, Peter Wakker and Nat Wilcox for helpful comments on previous drafts and insightful discussions regarding this research as well as seminar participants at the University of Pennsylvania Wharton School, the University of Chicago Booth School of Business, the University of Texas at Dallas, the University of California, Riverside, the University of Alabama and Chapman University. The usual disclaimers apply.

I. INTRODUCTION

Many decisions in life involve uncertain outcomes that materialize at different points in time. For example, the struggle to kick an addiction involves a tradeoff between short term gratification and an increased risk of future health problems. Saving for retirement involves consideration of preferences for immediate consumption and uncertainty about future income. Whether to pursue a long-term project involves consideration of the time the project is expected to take and the likelihood of project success. The decision to purchase a warranty on a television or appliance involves a higher immediate cost, but reduced product breakdown risk. The decision to take a 'buy-it-now' option on eBay or wait until the auction ends for the chance of a better deal, the decision to purchase a laptop today or wait for a potential Black Friday sale, and the decision to take out a mortgage on a home or wait for a possibly lower interest rate each involve a tradeoff between a certain, immediate payoff and a risky, delayed payoff.

As these examples illustrate, decisions often involve both risk and time delays. Yet the domains of risk and time have traditionally been studied separately. In cases where risk and time preferences are both considered, the discounted expected utility (DEU) model remains a major workhorse for analyzing individual behavior. There are, however, a variety of important shortcomings of DEU. For instance, it implies that introducing risk has no effect on time preference, that introducing delays has no effect on risk preference, that risk and time preferences are generated by the same utility function, that people are risk-seeking toward lotteries over uncertain payment dates, that people discount the future exponentially, and that discounting does not depend on the magnitude of outcomes. All of these predictions have been contradicted by experimental evidence.

In this paper, we introduce a dual process model of choices under risk and over time that naturally resolves each of these limitations of the DEU model. The model generalizes both rank-dependent utility theory (Quiggin, 1982), and Mukherjee's (2010) dual system model of choices under risk to develop a unified model that accounts for attitudes toward risk, attitudes toward time, and a variety of interaction effects between risk and time preferences. The proposed model also provides a unification of three classes of decision models – rank-dependent utility theory, expected utility, and dual process (or dual selves) theories. We refer to the model developed here as Dual Process Utility (DPU) theory.

The DPU theory introduces a new parameter into the analysis of economic decision models which represents the decision maker's 'cognitive type' or 'cognitive skills'. Essentially, an agent's 'cognitive type' identifies whether a person naturally engages in more intuitive and feeling-based processing or relies on more analytical and calculation-based processing. While cognitive skills have been found to correlate with a wide variety of economic behaviors including risk and time preferences (e.g., Frederick, 2005; Burks et al., 2009), saving behavior (Ballinger et al., 2011), strategic sophistication (Carpenter et al., 2013), and efficiency in experimental asset markets (Corgnet et al., 2015), they appear nowhere in the conventional economic models of individual choice.

In addition to introducing the DPU model, we provide DPU with a strong theoretical foundation, motivated by plausible psychological assumptions regarding the properties of dual systems in decision making, as well as by a simple axiomatic approach in which the convex combination functional form of DPU and the existence and uniqueness of the parameter representing the decision maker's 'cognitive type' are implied by the axioms.

We subsequently show that DPU predicts a variety of the major empirical findings regarding risk and time preferences. After providing some background in §II, the model is introduced in §III. In §IV we demonstrate that DPU explains present bias and that DPU resolves a long-standing paradox in decision theory by simultaneously predicting both the magnitude effect for choice over time and peanuts effect for choice under risk. In §V, we show that DPU explains empirically observed interaction effects between risk and time preferences (time affects risk preference, risk affects time preference, subendurance). In §VI we demonstrate that DPU permits a separation between risk and time preferences, it predicts a preference for diversifying payoffs across time, it predicts risk aversion toward timing risk, and it predicts observed correlations between risk preference, time preference, and cognitive reflection. The behaviors implied by the model are summarized in §VII. Related literature is discussed in §VIII. A simple extension to social preferences is considered in §IX - §XI. Concluding remarks are provided in §XII. An application of the model to consumer behavior is illustrated in Appendix A. Proofs are provided in Appendix B.

II. BACKGROUND

The study of risk preferences and time preferences, both analytically and empirically, has been the primary focus of research on individual choice for over half a century. However, although expected utility theory was axiomatized by von Neumann and Morgenstern in 1947, and discounted utility theory was axiomatized by Koopmans in 1960, it was not until 1991 when researchers identified remarkable parallels between the major anomalous behaviors across both domains – such as a common ratio effect in choice under risk and a common difference effect in choice over time (Prelec and Loewenstein, 1991). However, even in pointing out parallel behaviors between risk and time, Prelec and Loewenstein also presented a kind of impossibility result, indicating that no model that simultaneously applied to risk and time could resolve both the peanuts effect in choice under risk (Markowitz, 1952) and the magnitude effect in choice over time (Prelec and Loewenstein, 1991). Prototypical examples of both of these effects are illustrated in Table I. In the example of the peanuts effect, preferences switch from risk-seeking at small stakes (e.g., preferring a 1% chance of winning \$100 to \$1 for certain) to risk-averse at larger stakes (preferring \$100 for certain over a 1% chance of winning \$10,000). In the example of the magnitude effect, behavior switches from impatient at small stakes (e.g., preferring \$7 now to \$10 in one year) to more patient at larger stakes (preferring \$1,000 in one year over \$700 now). Note that both effects involve scaling outcomes up by a common factor. The peanuts effect is not explained by the most widely used specification of cumulative prospect theory due to Tversky and Kahneman (1992) with a power value function, even when allowing for any probability weighting function. A more fundamental challenge when relating risk and time preferences is that the peanuts effect seems to reveal decreasing sensitivity to payoffs at larger stakes, while the magnitude effect seems to reveal increasing sensitivity to payoffs at large stakes. Thus, any conventional approach to explaining the peanuts effect should predict the *opposite* of the magnitude effect (and vice versa). Prelec and Loewenstein could not explain both effects, and this challenge has remained unresolved over the subsequent twenty-five years, posing an apparent impossibility result that no common approach to modeling risk and time preferences can capture both of these basic behaviors. Somewhat surprisingly, we will demonstrate that the model presented here simultaneously predicts both effects.

Since the ‘common approach’ to risk and time preferences pioneered by Prelec and Loewenstein (1991), other models have been developed to explain behaviors across both domains. For instance, models of similarity judgments apply the same cognitive process to explain anomalies under risk and anomalies over time (Rubinstein 1988; Leland 1994; Leland

2002; Rubinstein, 2003). However, this approach does not address another basic question of how risk and time preferences *interact*.

It has been only fairly recently that attention has shifted to explaining interactions between risk and time preferences. This research direction was partially spurred by experimental studies from Keren and Roelofsma (1995) and Baucells and Heukamp (2010) who each observed different and systematic interactions between risk and time preferences. For instance, Keren and Roelofsma (1995) observed that uncertainty induces more patient behavior. Baucells and Heukamp (2010) and Abdellaoui et al. (2011) both observed that time delays induce more risk-taking behavior. Andersen et al. (2011) and Miao and Zhong (2015) observed a preference for diversifying risks across time. Onay and Onculer (2007) and DeJarnette et al. (2015) observed risk aversion to lotteries over uncertain payment dates. These behaviors are illustrated in Table I.

Table I. Choices between Options A and B involving Risk and Time

Observation	Option A vs.	Option B
Peanuts Effect* (Markowitz, 1952)	**(100, with 1%, now)** (10,000, with 1%, now)	(1, for sure, now) **(100, for sure, now)**
Magnitude Effect* (Prelec and Loewenstein, 1991)	**(7, for sure, now)** (700, for sure, now)	(10, for sure, 1 year) **(1,000, for sure, 1 year)**
Common Ratio Effect* (Baucells and Heukamp, 2010)	**(9, for sure, now)** (9, with 10%, now)	(12, with 80%, now) **(12, with 8%, now)**
Common Difference Effect** (Keren and Roelofsma, 1995)	**(100, for sure, now)** (100, for sure, 26 weeks)	(110, for sure, 4 weeks) **(110, for sure, 30 weeks)**
Time affects Risk Preference* (Baucells and Heukamp, 2010)	**(9, for sure, now)** (9, for sure, 3 months)	(12, with 80%, now) **(12, with 80%, 3 months)**
Risk affects Time Preference** (Keren and Roelofsma, 1995)	**(100, for sure, now)** (100, with 50%, now)	(110, for sure, 4 weeks) **(110, with 50%, 4 weeks)**
Subendurance* (Baucells et al., 2009)	**(100, for sure, 1 month)** (5, for sure, 1 month)	(100, with 90%, now) **(5, with 90%, now)**
Diversification across Time*** (Miao and Zhong, 2015)	**If Heads: 100, now** **If Tails: 100, 1 week**	If Heads: 100 now, 100, 1 week If Tails: 0 now, 0, 1 week
Aversion to Timing Risk*** (Onay and Onculer, 2007)	If Heads: 100, 5 weeks If Tails: 100, 15 weeks	**If Heads: 100, 10 weeks** **If Tails: 100, 10 weeks**

Adapted from Baucells and Heukamp (2012). Complementary probabilities for all options correspond to payoffs of 0. Sources of experimental results are in parentheses. Majority responses of experimental subjects are in bold font. *Currency in Euros; ** Currency in Dutch Guilders; *** Prototypical examples.

An intuitive approach to modeling risky and intertemporal choices is to multiply a time discount function by a probability weighting function by a utility or value function. However, this general approach does not explain the finding that time affects risk preference (see Table I) since both alternatives in the example by Baucells and Heukamp (2010) are delayed by the same amount (e.g., three months) and so the discount weights cancel when comparing options A and B. This approach also does not explain the finding that risk affects time preference, since both payoffs in the example by Keren and Roelofsma (1995) have the same probability (e.g., 50%), and so the probability weights cancel when comparing options A and B. In addition, this approach does not explain the finding of subendurance in the example by Baucells et al. (2009), since both options have the same payoffs (e.g., €100) and so the utilities cancel when comparing options A and B. It is then not obvious how to model such interaction effects between time delays, probabilities, and payoffs. It may be even less clear how to derive behaviors in the direction observed in experimental studies, or whether the same approach that might explain interaction effects for time delays can also explain interaction effects for probabilities, and payoffs. We will show that such a unified approach to these interaction effects is not only possible, but has a simple and intuitive interpretation.

After the peanuts and magnitude effects, the next five examples in Table I are adapted from Baucells and Heukamp (2012). The last two examples are prototypical illustrations of a preference for diversification across time and aversion to timing risk with outcomes determined by the toss of a fair coin. Further research on relations between risk and time preferences was spurred by Halevy (2008) and Saito (2011) who investigated formal equivalences between the Allais paradox and hyperbolic discounting, and by Andreoni and Sprenger (2012).

II.A. Dual Processes in Decision Making

Recent literature in cognitive science argues that people do not have a single mental processing system, but rather have two families of cognitive processes. Qualitatively similar distinctions have been made by many authors. Stanovich and West (2000), and Kahneman and Frederick (2002) label these families neutrally as System 1 processes and System 2 processes where System 1 includes automatic, intuitive and affective processes and System 2 includes more deliberative, logical, and reflective processes. Kahneman (2011) simply distinguishes between processes that are ‘fast’ and ‘slow.’ Rubinstein (2007, 2013) distinguishes between “instinctive” and “cognitive” processes. Denes-Raj and Epstein (1994) distinguish between an ‘experiential system’ and a ‘rational system’. Hsee and Rottenstreich (2004) posit two qualitatively different types of valuation processes – valuation by feeling and valuation by calculation[1]. Following Stanovich and West (2000), we adopt the neutral System 1/System 2 distinction in our analysis.

Concurrent with the dual process paradigm developing in the psychology literature, a plethora of dual process models have emerged recently in economics, each with similar distinctions between the types of processes involved. The relation between DPU and alternative dual system or dual selves models is discussed in §13. Despite their recent rise to theoretical prominence, two-system (dual process) theories date back to the early days of scholarly thought. The conflict between reason and passion, for instance, features prominently in Plato’s *Republic* and in Adam Smith’s *Theory of Moral Sentiments*.

III. DUAL PROCESS UTILITY THEORY

III.A. Setup and Assumptions

Our approach in this section builds on the variant of Harsanyi’s (1955) group decision theorem in Keeney and Nau (2011). Keeney and Nau adapted Harsanyi’s theorem for groups of agents with subjective expected utility preferences. We adapt the setting to decisions under risk and decisions over time in a model of individual choice.

Formally, we proceed as follows: There is a finite set, T, of time periods, a finite set $\mathcal{M}$ of outcomes, with $\mathcal{M} \subset \mathbb{R}$, and a finite set X of consumption sequences. We index consumption sequences by $j \in \{1,2,\dots,n\}$ and we index time periods by $t \in \{0,1,\dots,m\}$. A *consumption sequence* $x_j \coloneqq [x_{j0},\dots,x_{jm}]$ is a sequence of dated outcomes. A *stochastic consumption plan* is a probability distribution over *consumption sequences.* We denote a stochastic consumption plan by a function $f\colon \mathrm{X} \to [0,1]$, with $f(x_j)$ denoting the probability it assigns to consumption sequence x_j. Denote the set of stochastic consumption plans by Ω.

We first make assumptions about the risk and time preferences of Systems 1 and 2. We first assume that the preferences for each system are internally consistent. Viewing discounted expected utility (DEU) preferences as a model of consistent risk and time preferences, we

[1] We do not employ ‘calculation’ to mean the agent is necessarily calculating expected utilities consciously. Rather, ‘calculation’ is meant in a broad sense to refer to reliance on logic and reasoning to make choices.

impose this structure on the risk and time preferences for each system. We will see that imposing consistent preferences for each system can nevertheless generate a variety of decision anomalies as emergent phenomena that arise through the interactions between systems. As a further restriction, we let System 1 be more risk-averse and more delay-averse (i.e., less patient) than System 2. The assumption that System 1 is more risk-averse than System 2 is broadly consistent with a range of evidence. Frederick (2005), Burks et al., (2009), Dohmen et al., (2010), and Benjamin et al. (2013) find that individuals with higher cognitive skills are less risk-averse (in particular, closer to risk-neutrality) than individuals with low levels of cognitive skills or cognitive reflection. Deck and Jahedi (2015) review evidence suggesting that increasing cognitive load (an experimental manipulation designed to increase reliance on System 1) leads to greater small-stakes risk aversion. The assumption that System 1 is less patient than System 2 is consistent with findings by Frederick (2005), Burks et al. (2009), Dohmen et al. (2010), and Benjamin et al. (2013) that decision makers with higher cognitive skills or cognitive reflection are more patient than individuals with lower levels of cognitive skills. In addition, Tsukayama and Duckworth (2010) find that decision makers are less patient for affect-rich outcomes. Our Assumption 1 is thus consistent with empirical evidence regarding the relationship between the risk and time preferences of Systems 1 and 2.

Formally, let $\succsim_s$ and $\succ_s$ denote weak and strict preference, respectively, between pairs of stochastic consumption plans for system $s,\ s \in \{1,2\}$ that satisfy the non-triviality conditions $f \succ_1 g$ and $f' \succ_2 g'$ for some $f, g, f', g' \in \Omega$.

Assumption 1 (Preferences[2] of Systems 1 and 2): *System s, $s \in \{1,2\}$ has discounted expected utility preferences, with System 1 more risk-averse and more delay-averse than System 2. That is, there exist utility functions u_1, u_2, with u_1 more concave than u_2, and unique discount factors, δ_1, δ_2 with $0 < \delta_1 < \delta_2 \leq 1$, such that for all $f, g \in \Omega$,*

(1) $f \succsim_s g \Leftrightarrow V_s(f) \geq V_s(g)$, *where* $V_s(f) = \sum_t \sum_j \delta_s^t \cdot f(x_{jt}) \cdot u_s(x_{jt})$,

for $s \in \{1,2\}$. All of our results continue to hold even in the more restrictive case where System 2 is risk-neutral and delay-neutral. Mukherjee (2010) and Loewenstein et al. (2015) both argue that risk-neutrality is a plausible, and even natural, assumption for System 2. To the extent that

[2] We avoid explicit preference axioms for Systems 1 and 2 to simplify the exposition and to focus on the novel parameter θ in our model which is derived from our assumptions in Proposition 1. See Traeger (2013) for an axiomatization of discounted expected utility.

System 2 characterizes an idealized rational agent, it appears at least plausible that it does not have a pure rate of time preference which some authors have argued to be irrational (e.g., Harrod, 1948; Traeger, 2013), and that it maximizes expected value. Formally, let $(c, p, t) \in \Omega$ denote the stochastic consumption plan f defined as $f(\mathrm{x}) = p$ if x is the all zeros vector except at $x_t = c$, and $f(\mathrm{x}) = 1 - p$ if x is the all zeros vector, and $f(\mathrm{x}) = 0$ for all other $\mathrm{x} \in \mathrm{X}$. By 'delay neutrality' we mean $(c, p, t) \sim_2 (c, p, r)$ *for all* $t, r \in T$. Hence, in this special case, System 2 still discounts the future for reasons of uncertainty, given the future is often more uncertain than the present, but not for reasons of impatience.

Let $\gtrsim$ and $\succ$ represent, respectively, weak and strict preferences of the decision maker over stochastic consumption plans. We minimally constrain the agent's time preferences, and do not impose stationarity, nor do we impose that time preferences are multiplicatively separable into a discount function and a utility function, so that we do not rule out common behaviors such as present bias or the magnitude effect. But note that while not ruling out these behaviors, we do not assume them either. Present bias and the magnitude effect as well as some observed interactions between risk and time preferences will emerge as general properties *implied* by our representation. Formally, our Assumptions 2 and 3 can be viewed as special cases of Assumptions 2 and 3 in Harsanyi's theorem, as presented in Keeney and Nau (2011).

Assumption 2 (Preferences of the Decision Maker): *The decision maker has time-dependent expected utility preferences among stochastic consumption plans. That is, there exists a possibly time-dependent utility function,* $u_t(x)$*, such that for all* $f, g \in \Omega$*,*

(2) $$f \gtrsim g \Leftrightarrow V(f) \geq V(g), \text{ where } V(f) = \sum_t \sum_j f(x_{jt}) \cdot u_t(x_{jt}).$$

Assumption 3 (Pareto Efficiency): *If both systems weakly prefer* f *to* g*, then* $f \gtrsim g$*, and if, in addition, one system strictly prefers* f *to* g *then* $f \succ g$.

Proposition 1 (Dual Process Utility Theorem I): *Given Assumptions 1, 2, and 3, there exists a unique constant*[3] $\theta \in (0,1)$*, unique discount factors* δ_1, δ_2 *with* $0 < \delta_1 < \delta_2 \leq 1$ *and utility functions,* u_1 *and* u_2*, with* u_1 *more concave than* u_2*, such that for all* $f, g \in \Omega$*, the decision maker's preferences are given by* $f \gtrsim g \Leftrightarrow V(f) \geq V(g)$*, where*

[3] In Harsanyi's theorem, the weights on individual member utilities are positive and unique up to a common scale factor. Without loss of generality, the weights can be scaled to sum to 1 in which case $\theta \in (0,1)$ is uniquely determined. We will also continue to discuss $\theta = 0$ and $\theta = 1$ as special cases of DPU since they are limiting cases of the model and θ can be arbitrarily close to 0 or 1.

(3) $$V(f) = (1-\theta)V_1(f) + \theta V_2(f)$$
$$= (1-\theta)\left(\sum_t \sum_j \delta_1^t \cdot f(x_{jt}) \cdot u_1(x_{jt})\right) + \theta\left(\sum_t \sum_j \delta_2^t \cdot f(x_{jt}) \cdot u_2(x_{jt})\right).$$

One could imagine many ways of aggregating preferences of Systems 1 and 2. Proposition 1 provides a formal justification for the convex combination approach. Although including both risk and time, and applying Harsanyi's theorem from social choice theory to model *individual* choice behavior are new, the proof for Proposition 1 follows straightforwardly from the proof of Theorem 1 in Keeney and Nau (2011). A related finding for decisions involving time but not risk was obtained in the context of group decision making by Jackson and Yariv (2015).

III.B. Aggregation of Non-expected Utility Preferences

The assumption of discounted expected utility preferences in Assumption 1 seems particularly appropriate for System 2 which may be intuitively thought to resemble the rational economic agent. However, in addition to differences in the content of risk and time preferences between systems (i.e., that the systems differ in their degrees of risk aversion and impatience), one might further propose that the two systems differ in the structure of their risk and time preferences, with System 2 having normative DEU preferences, and with System 1 having behavioral preferences based on prospect theory (PT) or rank dependent utility (RDU) theory. Supporting this, Rottenstreich and Hsee (2001) find that inverse S-shaped probability weighting (as assumed in RDU theory (Quiggin 1982) and PT[4] (Tversky and Kahneman 1992)) is more pronounced for affect-rich outcomes. Support for assuming System 1 has PT preferences also comes from Barberis et al. (2013) who use PT to model "System 1 thinking" for initial reactions to changes in stock prices. Reflecting on PT three decades later, Kahneman (2011, pp. 281-282) remarks, "It's clear now that there are three cognitive features at the heart of prospect theory…They would be seen as operating characteristics of System 1."

Recent impossibility results (Mongin and Pivato, 2015; Zuber, 2016) have demonstrated difficulties in using axiomatic methods to aggregate non-expected utility preferences. For instance, Zuber (2016) considers a general class of non-expected utility preferences and concludes, "non-expected utility preferences cannot be aggregated consistently." However, this

[4]Although this version is commonly called cumulative prospect theory, Peter Wakker has noted (personal communication, July 2, 2016) that it was Amos Tversky's preference for this latter version to be called prospect theory. Following Tversky's preference, we refer to the 1992 version as prospect theory.

conclusion applies to the approach in which the social planner (or, in our case, the decision maker) has preferences over the same stochastic consumption plans as the two systems. We next show how to derive a representation that generalizes (1) by allowing for the possibility that System 1 engages in non-linear probability weighting, as in RDU and PT. Such analysis is possible by assuming the agent cares about the payoffs to each system and has preferences over distributions of payoffs to Systems 1 and 2. While we employ RDU theory, the behaviors we study involve only positive outcomes, in which case PT coincides with RDU.

Our approach is based on the group decision theorem in Keeney and Nau (2011). If System $s, s \in \{1,2\}$, has DEU preferences, the function that is defined by the product of discount factor δ_s^t and utility function u_s can be represented by a vector $\boldsymbol{v_s}$ of length $|\mathrm{T}| \times |\mathrm{X}|$ in which $v_s(t, x_{jt}) = \delta_s^t \cdot u_s(x_{jt})$. We denote stochastic consumption plans by f and g, and denote their corresponding vector representations by $\boldsymbol{f}$ and $\boldsymbol{g}$. Similarly, we denote the rank-dependent probability weighting function of f by $\pi(f)$ and the corresponding vector of probability weights by $\boldsymbol{\pi_f}$. The discounted expected utility of f to System s can then be expressed as the inner product $\boldsymbol{f} \cdot \boldsymbol{v_s} = \sum_t \sum_j f(x_{jt}) v_s(t, x_{jt})$ in which case $f \succsim_s g$ if and only if $\boldsymbol{f} \cdot \boldsymbol{v_s} \geq \boldsymbol{g} \cdot \boldsymbol{v_s}$. If instead System s has discounted RDU preferences, it distorts cumulative probabilities by $\pi: [0,1] \to [0,1]$, with $\pi(0) = 0$ and $\pi(1) = 1$, that takes the standard rank-dependent form with weighting function w:

$$\pi(f(x_{jt})) = w\big(f(x_{jt}) + \cdots + f(x_{1t})\big) - w\big(f(x_{j-1,t}) + \cdots + f(x_{1t})\big),$$

for $j \in \{1,2, \dots, n\}$, where consumption sequences are ranked according to the discounted utility for System s for each sequence such that $\sum_t \delta_s^t \cdot u_s(x_{nt}) \leq \cdots \leq \sum_t \delta_s^t \cdot u_s(x_{1t})$.

Note that the discounted rank-dependent utility of f for System s can be expressed as the product $\boldsymbol{\pi_f} \cdot \boldsymbol{v_s}$ in which case $f \succsim_s g$ if and only if $\boldsymbol{\pi_f} \cdot \boldsymbol{v_s} \geq \boldsymbol{\pi_g} \cdot \boldsymbol{v_s}$. The rank-dependent form for π avoids violations of stochastic dominance, and the typical inverse S-shaped form often assumed for w reflects the psychophysics of probability perception as it exhibits diminishing sensitivity from the endpoints of the probability scale. This approach for ranking consumption sequences essentially collapses each sequence into its discounted utility and then rank-dependent probability weighting is applied to these discounted utilities for System 1. This effectively reduces outcomes to a single dimension to which RDU can be applied as it would be applied to lotteries with static outcomes.

Assumption 1* (System 1 preferences): *System 1 has discounted rank-dependent utility preferences: There exist a unique discount factor* $\delta_1 \in (0,1)$, *a unique, strictly increasing continuous rank-dependent probability weighting function* π, *and utility function* $u_1: \mathcal{M} \to \mathbb{R}$, *such that for all* $f, g \in \Omega$, $f \succsim_1 g \Leftrightarrow V_1(f) \geq V_1(g)$, *where*

$$V_1(f) = \sum_t \sum_j \delta_1^t \cdot \pi(f(x_{jt})) \cdot u_1(x_{jt}),$$

and consumption sequences are ranked by System 1's discounted utility for each sequence such that $\sum_t \delta_1^t \cdot u_1(x_{nt}) \leq \cdots \leq \sum_t \delta_1^t \cdot u_1(x_{1t})$ *prior to applying the* π *transformation.*

Following the analogy between the rational agent of economic theory and System 2, we assume that System 2 has discounted expected utility preferences:

Assumption 2* (System 2 preferences): *System 2 has discounted expected utility preferences: There exist a unique discount factor* $\delta_2 \in (\delta_1, 1]$, *and a utility function* $u_2: \mathcal{M} \to \mathbb{R}$, *such that for all* $f, g \in \Omega$, $f \succsim_2 g \Leftrightarrow V_2(f) \geq V_2(g)$, *where:*

$$V_2(f) = \sum_t \sum_j \delta_2^t \cdot f(x_{jt}) \cdot u_2(x_{jt}).$$

The main differences between System 1 and System 2 preferences as formalized in Assumptions 1 and 2 are that System 1 has behavioral risk preferences (given by rank-dependent utility or prospect theory), System 2 has normative risk preferences (given by expected utility theory), and System 2 is more patient than System 1.

Since the set of outcomes, $\mathcal{M}$, is finite, the utility function for each system is bounded, and since V_1 and V_2 are each continuous and the set of stochastic consumption plans is compact, it follows that V_1 and V_2 are also bounded. Therefore, we can choose a constant $k_s > 0$ for each system such that $k_s u_s$ yields $V_s(f) \leq 1$ for all f for $s \in \{1,2\}$. One might also let u_1 and u_2 each be normalized to a 0-1 scale.

Note that every $f \in \Omega$ can be associated with the vector $(\boldsymbol{\pi_f} \cdot \boldsymbol{v_1}, \boldsymbol{f} \cdot \boldsymbol{v_2})$ whose elements are the value of f (the discounted rank-dependent utility of f) as evaluated by System 1 $(\boldsymbol{\pi_f} \cdot \boldsymbol{v_1})$**,** and the value of f (the discounted expected utility of f) as evaluated by System 2 $(\boldsymbol{f} \cdot \boldsymbol{v_2})$**.** Every such vector is a point in the square $[0,1] \times [0,1]$. In modeling the choices made by the decision maker, we let her consider the distribution of valuations that her decisions yield for each System. The set of all such vectors consists of the entire square, as if there are hypothetical alternatives available that yield all possible comparisons of the worst and best outcomes for System 1 and System 2 (all possible 0-1 profiles of discounted rank-dependent (System 1) utilities and discounted expected (System 2) utilities, and their convex combinations). The elements of the square will be referred to as 'signal vectors'. Denote a signal vector by $F = (F_1, F_2)$, where F_s is a 'reward value signal' or a 'reward prediction signal' representing the value of that alternative transmitted by System $s \in \{1,2\}$, measured on a scale of 0 to 1. If F corresponds to some f in the real choice faced by the agent, then $F_1 = V_1(f)$ and $F_2 = V_2(f)$. Let $\mathcal{F} = \{F^{(1)}, \dots, F^{(N)}\}$ denote a finite set of signal vectors and let $(\mathcal{F}, p) := \left(\left(p^{(1)}, F^{(1)}\right), \dots, \left(p^{(N)}, F^{(N)}\right)\right)$ denote a 'lottery over signals' with $p^{(1)} + \cdots + p^{(N)} = 1$ and $p^{(i)} > 0$ for all i, in which signal vector $F^{(i)}$ is chosen with probability $p^{(i)}$.

Let $\succsim$ represent the agent's preferences for lotteries over signals. To illustrate such a preference, let $\left(V_1(f), V_2(f)\right) = (0.5{,}0.5), \left(V_1(g), V_2(g)\right) = (1{,}0)$ and $\left(V_1(h), V_2(h)\right) = (0{,}0)$. Then if $\left(1, \left(V_1(f), V_2(f)\right)\right) \succsim \left(q, \left(V_1(g), V_2(g)\right); 1 - q, \left(V_1(h), V_2(h)\right)\right)$ for $q > 0.5$, the agent prefers the equitable assignment of valuations (0.5,0.5) to an assignment with a chance of the highest valuation for System 1 and the lowest valuation for System 2. In Assumption 3* we assume the agent has preferences over assignments of values to System 1 and System 2 that maximize the agent's expected utility with the agent's utility defined over vectors of System 1 and System 2 reward signals as if the utility function integrates the reward signals into a unifying utility value. Formally, Assumption 3* employs expected utility condition P2 from Fishburn (1965), and can be viewed as a special case of Assumption 4 in Keeney and Nau (2011).

An alternative (and simpler) interpretation of Assumption 3* is that the agent cares about the welfare of both systems and has preferences over distributions of payoffs to each system.

Assumption 3* (Preferences of the Decision Maker): *The decision maker has expected utility preferences for lotteries over signals. That is, there exists a von Neumann-Morgenstern utility function μ on the set of signal vectors such that for any two lotteries over signals $(\mathcal{F}, p)$ and $(\mathcal{F}, q)$, $(\mathcal{F}, p) \succsim (\mathcal{F}, q)$ if and only if $\mu(\mathcal{F}, p) \geq \mu(\mathcal{F}, q)$ where $\mu(\mathcal{F}, p) = \mu\left(\left(p^{(1)}, F^{(1)}\right), \dots, \left(p^{(N)}, F^{(N)}\right)\right) = p^{(1)}\mu\left(F^{(1)}\right) + \cdots + p^{(N)}\mu\left(F^{(N)}\right)$.*

Next, note that in a lottery over signals, it might be the case that System s receives the same discounted rank-dependent utility or discounted expected utility in two or more stochastic consumption plans. In that case, we can compute a marginal probability for the assignment of a given payoff to System s. Let $F_1^{(n)}$ and $F_1^{(n')}$ denote the discounted rank-dependent utilities assigned to System 1 by two signal vectors $F^{(n)}$ and $F^{(n')}$ that are among the signal vectors of a lottery $(\mathcal{F}, p)$. If $F_1^{(n)} = F_1^{(n')} = u$ and $F_1^{(n'')} \neq u$ for $n'' \neq n, n'$, then the marginal probability of discounted rank-dependent utility u for System 1 is $p^{(n)} + p^{(n')}$. In Assumption 4*, the decision maker chooses in accordance with one system for choices affecting only that system's marginal distribution of payoffs. Formally, this assumption is a special case of Assumption 5 in Keeney and Nau (2011):

Assumption 4* (Relationship between preferences in Assumptions 1*, 2*, and 3*):

(i) *If two lotteries over signals yield identical marginal probabilities for the discounted rank-dependent utilities of System 1, the decision maker is indifferent between these lotteries if and only if System 2 is indifferent between them.*

(ii) *If two lotteries over signals yield identical marginal probabilities for the discounted expected utilities of System 2, the decision maker is indifferent between these lotteries if and only if System 1 is indifferent between them.*

Assumption 4* relates the preferences from Assumptions 1* - 3*, and essentially says that if it does not matter to one system (e.g., intuition) which alternative is selected, the other system (e.g., logic) gets to make the decision. This assumption implies the decision maker's preferences satisfy Fishburn's (1965) mutual independence condition. Assumption 4* is weaker than the Pareto assumption in Harsanyi (1955) and, together with Assumptions 1* – 3*, it permits the aggregation of non-expected utility preferences.

Proposition 2 (Dual Process Utility Theorem II): *Given Assumptions 1* through 4*, there exist a constant $\theta \in (0,1)$, unique discount factors δ_1, δ_2 with $0 < \delta_1 < \delta_2 \leq 1$, a unique, strictly increasing continuous rank-dependent probability weighting function π, and utility functions, u_1 and u_2, each unique upon normalization to a 0-1 scale such that for all $f, g \in \Omega$, $f \succsim g \Leftrightarrow V(f) \geq V(g)$, where*

(4) $$V(f) = (1-\theta)V_1(f) + \theta V_2(f)$$
$$= (1-\theta)\left(\sum_t \sum_j \delta_1^t \cdot \pi\big(f(x_{jt})\big) \cdot u_1\big(x_{jt}\big)\right) + \theta \sum_t \sum_j \delta_2^t \cdot f(x_{jt}) \cdot u_2(x_{jt}).$$

Moreover, if there exists $f, g \in \Omega$ with $f \sim g$ and $V_1(f) \neq V_1(g)$ then θ is unique.

A proof of Proposition 2 is provided in the appendix. Proposition 2 extends (3) to the more general case of (4) where System 1 has behavioral preferences (given by prospect theory or RDU) and System 2 has normative preferences given by discounted expected utility theory. The aggregation of non-expected utility preferences in Proposition 2 is new, and perhaps surprising given the recent impossibility results by Mongin and Pivato (2015) and Zuber (2016) which seem to suggest that non-expected utility preferences cannot be aggregated consistently under fairly general conditions.

We say an agent with preferences given by (4) has *dual process utility* (DPU) preferences. To simplify notation and to illustrate the model with the delay-neutrality and risk-neutrality assumptions for System 2, we will employ the specification in (5) in our analysis, where we drop the subscripts on System 1's discount factor and utility function:

(5) $$V(f) = (1-\theta)\left(\sum_t \sum_j \delta^t \cdot \pi\big(f(x_{jt})\big) \cdot u\big(x_{jt}\big)\right) + \theta \sum_t \sum_j f(x_{jt}) \cdot x_{jt}.$$

Our subsequent results are robust to the risk-neutrality and delay-neutrality assumptions for System 2 and continue to hold provided u_2 is a monotonically increasing concave function with less curvature than u_1 and System 2 is more patient than System 1.

Under the premise that System 1 attends to subjective and affective ("hot") qualities of a stimulus and System 2 attends to the objective information ("cool" qualities) contained in a stimulus, one can interpret (5) as a weighted average of the subjective value of a stochastic consumption plan (the 'stimulus') as determined by System 1 and the objective value of the stochastic consumption plan as judged by System 2.

III.C. Interpretation of θ

The parameter θ may be interpreted as the degree to which an agent is 'hard-wired' to rely on System 2. Intuitive thinkers then have low values of θ, whereas more analytical thinkers have higher values of θ. We will refer to θ as the decision maker's "cognitive type," with one's cognitive type becoming less based on feeling and intuition and more reliant on logic and calculation as θ increases. From a neuro-economic perspective, there are tight neural connections between the prefrontal cortex, a brain region implicated in planning, analytical thinking, and executive function and the limbic system, an evolutionarily older brain region involved in the generation of emotions and the experience of pleasure. One might view θ as indexing the strength of neural connections in the prefrontal areas relative to the strength of neural connections in the limbic areas. As noted, $V_1(f)$ and $V_2(f)$ may be viewed as reward value signals or reward prediction signals transmitted by Systems 1 and 2. When these signals conflict, the agent's choice may be determined such that the agent chooses in favor of the system with the stronger signal, where the product $\theta V_2(f)$ can then be interpreted as the strength of the System 2 reward signal for any $f \in \Omega$, and $(1-\theta)V_1(f)$ is the strength of the System 1 reward signal. Ceteris paribus, an agent with a high value of θ will exhibit more reflection and more self-control (the agent is adapted to rely more on System 2), whereas an agent with a low value of θ will have greater weight on the reward signals from System 1 and will find it more difficult to exert self-control in the presence of temptation. For instance, when choosing between a tempting alternative, f (i.e., an alternative with a large System 1 reward signal $V_1(f)$), and a delayed alternative g such that $V_1(f) > V_1(g)$ and $V_2(f) < V_2(g)$, a sufficiently large value of θ is needed for an agent to resist temptation.

Note that DPU adopts the view of cognitive sophistication implicit in the interpretation of the cognitive reflection test (Frederick, 2005) and models of level-k thinking (Camerer et al., 2004), namely that there are reflective thinkers or those with high levels of cognitive sophistication and there are intuitive thinkers or those with lower levels of cognitive sophistication. It is in this sense in which we view θ as reflecting a decision maker's 'cognitive type' which allows for heterogeneity across agents. Within agents, DPU reflects a compromise between the System 1 preference for immediate gratification and the more patient preferences of System 2. This 'compromise' is consistent with the findings of Andersen et al. (2008) who "…observe what appears to be the outcome of a decision process where temptation and long-run considerations are simultaneously involved."

III.D. Basic Properties of DPU

Consider two stochastic consumption plans f and g, where $f(x_j)$ and $g(x_j)$ are the probabilities which f and g assign to consumption sequence x_j, respectively. Since a decision maker either receives one consumption sequence or another and so cannot interchange components of any arbitrary sequences, we first seek a means of objectively ranking different consumption sequences, analogous to how one would rank individual outcomes. We can then extend the standard definition of stochastic dominance from lotteries over outcomes to lotteries over consumption sequences. In particular, we say sequence x_j *dominates* sequence x_k if $x_{jt} \geq x_{kt}$ for all $t \in \{0,1,\dots,T\}$, with a strict inequality for at least one t. We say that consumption sequences $x_1, \dots, x_n$ are *monotonically ordered* if x_j dominates x_{j+1} for all $j \in \{1,\dots,n-1\}$. For any monotonically ordered consumption sequences $x_1, \dots, x_n$, we say f (first-order) *stochastically dominates* g if $F(x_{jt}) \leq G(x_{jt})$ for all $j \in \{1,\dots,n\}$, and all $t \in \{0,1,\dots,T\}$, where F and G are the cumulative distribution functions for f and g, respectively. Note that this reduces to the standard definition of stochastic dominance in an atemporal setting.

Proposition 3: *Let* $\succsim$ *have a DPU representation as in* (4). *Then for any fixed* $\theta \in [0,1]$, $\succsim$ *satisfies the following properties over stochastic consumption plans:*

(i) *Weak order (*$\succsim$ *is transitive and complete)*

(ii) *Continuity*

(iii) *First order stochastic dominance.*

The proofs of properties (i) and (ii) in Proposition 3 are standard so we prove only (iii). Recall that $\pi(\cdot)$ ranks sequences such that $\sum_t \delta_1^t \cdot u_1(x_{1t}) \geq \cdots \geq \sum_t \delta_1^t \cdot u_1(x_{nt})$. Note that if consumption sequences $x_1, \dots, x_n$ are monotonically ordered, then $u_1(x_{1t}) \geq \cdots \geq u_1(x_{nt})$ for all $t \in \{0,1, \dots, T\}$, and for any increasing function u_1. Thus, $\pi(\cdot)$ preserves the monotonic ordering of the sequences. If f *stochastically dominates* g, then $\delta^t \sum_j \pi(f(x_j))u(x_{jt}) > \delta^t \sum_j \pi(g(x_j))u(x_{jt})$, for each period $t \in \{0,1, \dots, T\}$, which implies that $V_1(f) > V_1(g)$. Since $V_2(f) > V_2(g)$, the convex combination of V_1 and V_2 ranks f higher than g for all $\theta \in [0,1]$.

IV. Empirical Violations of Discounted Utility and Expected Utiility

In this and the following sections, all propositions assume the decision maker has dual process utility preferences (given by (5)). Proofs of Propositions in §4 and §5 are given in the appendix. Notably, each of these results (Propositions 4, 5, 6, 7, 8, and 9) do not hold when $\theta = 0$ or $\theta = 1$, indicating the need for a dual process paradigm in our setup. First, we show that DPU resolves two empirical violations of discounted utility theory - present bias and the magnitude effect.

IV.A. Present Bias

Systematic empirical violations of the stationarity axiom of discounted utility theory (Koopmans, 1960), such as present bias, have been well-documented in experiments (Frederick et al., 2002), and are thought to reveal time-inconsistent preferences (Laibson 1997; O'Donoghue and Rabin 1999). Formal accounts of present bias and hyperbolic discounting have often directly assumed such behavior in the functional form of the agent's preferences (e.g., Loewenstein and Prelec 1992; Laibson 1997). Surprisingly, present bias emerges as a general property of DPU without any explicit assumptions regarding hyperbolic discounting or diminishing sensitivity to delays. In fact, present bias is predicted by DPU even though each system has time consistent preferences.

As before, let (c, p, t) denote a stochastic consumption plan which has one non-zero outcome c, to be received with probability p at time t. We have the following definition:

Definition 1 (Present Bias): *Present bias* holds if for $y \in (0, c)$, and $t, \Delta > 0$, $(y, p, 0) \sim (c, p, \Delta) \Longrightarrow (y, p, t) \prec (c, p, t + \Delta)$

Proposition 4: *Under DPU, present bias holds if and only if* $\theta \in (0,1)$.

Proposition 4 implies that DPU explains the example of the common difference effect in Table I demonstrated by Keren and Roelofsma (1995). In particular, a decision maker indifferent between 100 Dutch guilders for sure now and 110 Dutch guilders for sure in 4 weeks will strictly prefer 110 Dutch guilders for sure in 30 weeks over 100 Dutch guilders for sure in 26 weeks. It is also clear from the proof of Proposition 4 that present bias does not hold if $\theta = 0$ or if $\theta = 1$. Thus, under DPU, present bias arises due to the interaction between System 1 and System 2.

IV.B. The Magnitude Effect

The DPU model also offers an explanation of the magnitude effect in intertemporal choice. The magnitude effect is the robust observation that behavior is more patient for larger rewards than for smaller rewards (Prelec and Loewenstein, 1991). Formally:

Definition 2 (Magnitude Effect): We say the *magnitude effect* holds if for $y \in (0, c)$, $s > t$, and $r > 1, (y, p, t) \sim (c, p, s) \Longrightarrow (ry, p, t) \prec (rc, p, s)$

Proposition 5: *For any concave power utility function u, the magnitude effect holds under DPU, if and only if* $\theta \in (0,1)$.

IV.C. The Peanuts Effect

While PT and RDU explain violations of expected utility theory (EU) such as the Allais paradoxes, standard specifications of PT or RDU do not explain the 'peanuts' effect. An example of this behavior is a willingness to pay \$1 for a one-in-ten million chance of \$1 million, but prefer a sure \$1000 over a one-in-ten million chance of \$1 billion. Under a power value function for PT, indifference in the former choice implies indifference in the latter for *any* probability weighting function and the peanuts effect does not hold. The problem is more challenging when incorporating both risk and time because, since Prelec and Loewenstein (1991), it has not been clear how the magnitude effect and the peanuts effect coexist. Yet DPU simultaneously predicts both effects. The peanuts effect holds since risk-seeking at small stakes is due to overweighting low probabilities (the domain where the peanuts effect is observed) while scaling payoffs up shifts more weight on the System 2 value function (if u_1 is more concave than u_2) which shifts preferences toward risk neutrality (if $u_2(x) = x$) or risk aversion (if u_2 is concave).

Definition 3 (Peanuts Effect): We say the *peanuts effect* holds if for $y \in (0, c),\ p > q$, and $r > 1, (c, q, t) \sim (y, p, t) \Longrightarrow (rc, q, t) \prec (ry, p, t)$.

Proposition 6: *Let* $\mathbb{E}[(y, p, t)] > \mathbb{E}[(c, q, t)]$. *Then for any concave power function* u, *the peanuts effect holds under DPU if and only if* $\theta \in (0,1)$.

For both the magnitude and peanuts effects, power utility is sufficient, but not necessary. Also, if System 2 is even slightly risk-averse, the peanuts effect holds when $\mathbb{E}[(y, p, t)] = \mathbb{E}[(c, q, t)]$.

In addition to resolving the peanuts and magnitude effects, DPU also explains the finding in Fehr-Duda et al. (2010) that probability weighting is stronger for low stakes than for high stakes. This observation holds naturally under DPU given the assumption that the System 2 value function is closer to risk-neutrality than the System 1 value function. However, this stake-size effect violates prospect theory which assumes probability weights and outcomes are independent.

V. INTERACTIONS BETWEEN RISK AND TIME PREFERENCE

In this section, we apply DPU to systematic interactions between risk and time preferences from Table I, identified in Baucells and Heukamp (2012).

V.A. Time Interacts with Risk Preference

We return now to the behaviors illustrated in Table I. As displayed in Table I, Baucells and Heukamp (2010) found most respondents in their study preferred a guaranteed 9 Euros immediately over an 80% chance of 12 Euros immediately, but chose the chance of receiving 12 Euros immediately when the probabilities of winning were scaled down by a factor of 10. This behavior is an instance of the Allais common ratio effect (Allais, 1953). Baucells and Heukamp further observed that when the receipt of payment is delayed 3 months, most respondents preferred an 80% chance of 12 Euros over a guaranteed 9 Euros. This finding that people are less risk-averse toward delayed lotteries was also observed by Abdellaoui et al. (2011).

The common ratio effect example from Baucells and Heukamp (2010) holds under DPU if the probability weighting function is sub-proportional. Here we confirm that DPU explains the finding that ‘time interacts with risk preference’ which holds even if System 1’s utility function is linear in probabilities. Let $\mathbb{E}[f]$ denote the (undiscounted) expected value of stochastic consumption plan f. We consider the case where the riskier lottery has the higher expectation as was the case in Baucells and Heukamp (2010).

Definition 4: We say *time interacts with risk preference* if for $y \in (0, c)$, $\alpha \in (0,1)$, and $s > t$, $(y, p, t) \sim (c, \alpha p, t) \Longrightarrow (y, p, s) \prec (c, \alpha p, s)$.

Proposition 7: *Let* $\mathbb{E}[(c, \alpha p, t)] > \mathbb{E}[(y, p, t)]$. *Then under DPU, time interacts with risk preference if and only if* $\theta \in (0,1)$

V.B. Risk Interacts with Time Preference

As displayed in Table I, Keren and Roelofsma (1995) found that most respondents in their study preferred a guaranteed 100 Dutch guilders immediately over a guaranteed 110 Dutch guilders in 4 weeks, but chose the guaranteed 110 when the receipt of both payments was delayed an additional 26 weeks. This behavior is an example of present bias. Keren and Roelofsma further observed that when the chance of receiving each payment was reduced, most respondents preferred a 50% chance of 110 Dutch guilders in 4 weeks over a 50% chance of 100 now. That is, making both options risky leads to more patient behavior, analogous to the effect of adding a constant delay to both options. This finding that people wait longer when the alternative is risky was also observed by Luckman et al. (2017).

Definition 5: We say *risk interacts with time preference* if for $y \in (0, c)$, $t, \Delta > 0$, and $q < p$, $(y, p, t) \sim (c, p, t + \Delta) \Longrightarrow (y, q, t) \prec (c, q, t + \Delta)$.

Proposition 8: *Under DPU, for any convex weighting function* w, *risk interacts with time preference if and only if* $\theta \in (0,1)$.

In the example by Keren and Roelofsma in Table I, convexity of the weighting function in Proposition 8 implies $w(0.5) < 0.5$, which implies $\pi(0.5) < 0.5$. This condition is a general feature of observed probability weighting functions (Starmer, 2000; Wakker, 2010) and represents a form of pessimism. Indeed this condition is implied by the assumption of pessimism in the rank-dependent utility framework. This condition holds for all convex probability weighting functions as well as for the familiar inverse-S-shaped weighting functions (such as those parameterized by Tversky and Kahneman (1992), Wu and Gonzalez (1996), Prelec (1998), and Gonzalez and Wu (1999)), Abdellaoui (2000), and Bleichrodt and Pinto (2000)). This condition is also a general property resulting from Prelec's (1998) axiomatic characterization of his one-parameter probability weighting function. This condition ($\pi(0.5) < 0.5$) will reappear in our analysis and is the only substantive property of the weighting function that is necessary for

DPU to explain the experimental observations studied here. The more general convexity condition is only necessary for the generalization of the behavior observed by Keren and Roelofsma to all $q < p$ as formalized in Definition 5.

V.C. Payoffs Interact with Risk and Time Preferences

Baucells et al. (2009) found that 81% of respondents preferred €100 for sure in one month to €100 with 90% probability immediately, but 57% preferred €5, with 90% probability immediately over €5 for sure, in one month. Baucells and Heukamp (2012) refer to this behavior as subendurance and they define it more generally as follows:

Definition 6: *Subendurance* holds if for $y \in (0, c)$, $t, \Delta > 0$ and $\lambda \in (0,1)$, $(c, p, t + \Delta) \sim (c, \lambda p, t) \implies (y, p, t + \Delta) \prec (y, \lambda p, t)$.

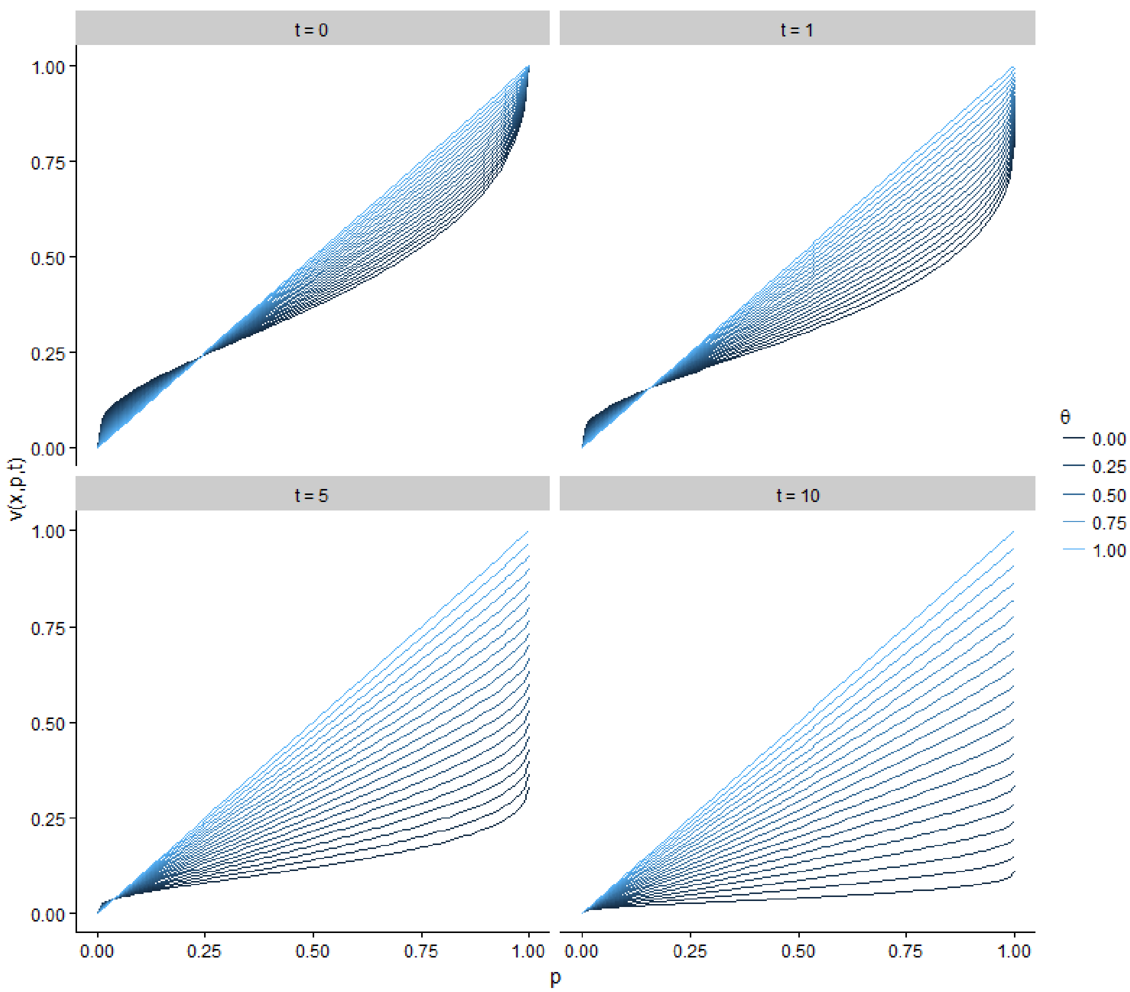
Proposition 9: *For any concave utility function u such that $u(0) = 0$, subendurance holds under DPU, if and only if $\theta \in (0,1)$.*

The interaction effects in this section challenge a larger class of time preferences than DEU. Indeed, they cannot be explained by any model of discounting in which evaluation of payoffs, probabilities, or delays is multiplicatively separable. As Baucells and Heukamp (2012) note, when evaluating a stochastic consumption plan (x, p, t), "One may be tempted to propose $V(x, p, t) = w(p)f(t)v(x)$. Unfortunately, this form is not appropriate because…probability and time cannot be separated. One may then propose the more general form $V(x, p, t) = g(p, t)v(x)$, but this fails to accommodate subendurance." Moreover, Ericson and Noor (2015) reject the assumption that discounting and utility functions are separable for nearly 70% of their participants. Given the necessity of a seemingly complex form for evaluating (x, p, t) to explain the observations in Table I, the DPU functional form in (5) is surprisingly simple.

V.D. Variations in Risk and Time

Figure I graphs (5) for different values of θ (within panels) and for different delays (across panels), as probabilities increase from 0 to 1. The figure employs Prelec's (1998) probability weighting function and evaluates a stochastic consumption plan paying $x > 0$ with probability p at time t and 0 otherwise, under the simplified case where $u_1(x) = u_2(x) \coloneqq 1$ and $u_1(0) = u_2(0) \coloneqq 0$. This specification may be viewed as a time-dependent probability weighting function that becomes flatter as the time horizon increases. In general, DPU does not have a separable probability weighting function that is independent of outcomes or time, but we can see how time affects the shape of the weighting function in the special case when $u_1(.) = u_2(.)$. Figure I suggests people are less sensitive to variations in probability for longer time horizons. Also, relative to an event's probability p, the function over-weights low probability events occurring over short horizons, such as drawings for state lottery tickets (if $\delta^t w(p) > p$), but under-weights low probability events over long horizons, such as natural disasters and health risks (if $\delta^t w(p) < p$). Epper and Fehr-Duda (2016) also argue that accounting for time delays permits the coexistence of overweighting and underweighting tail events.

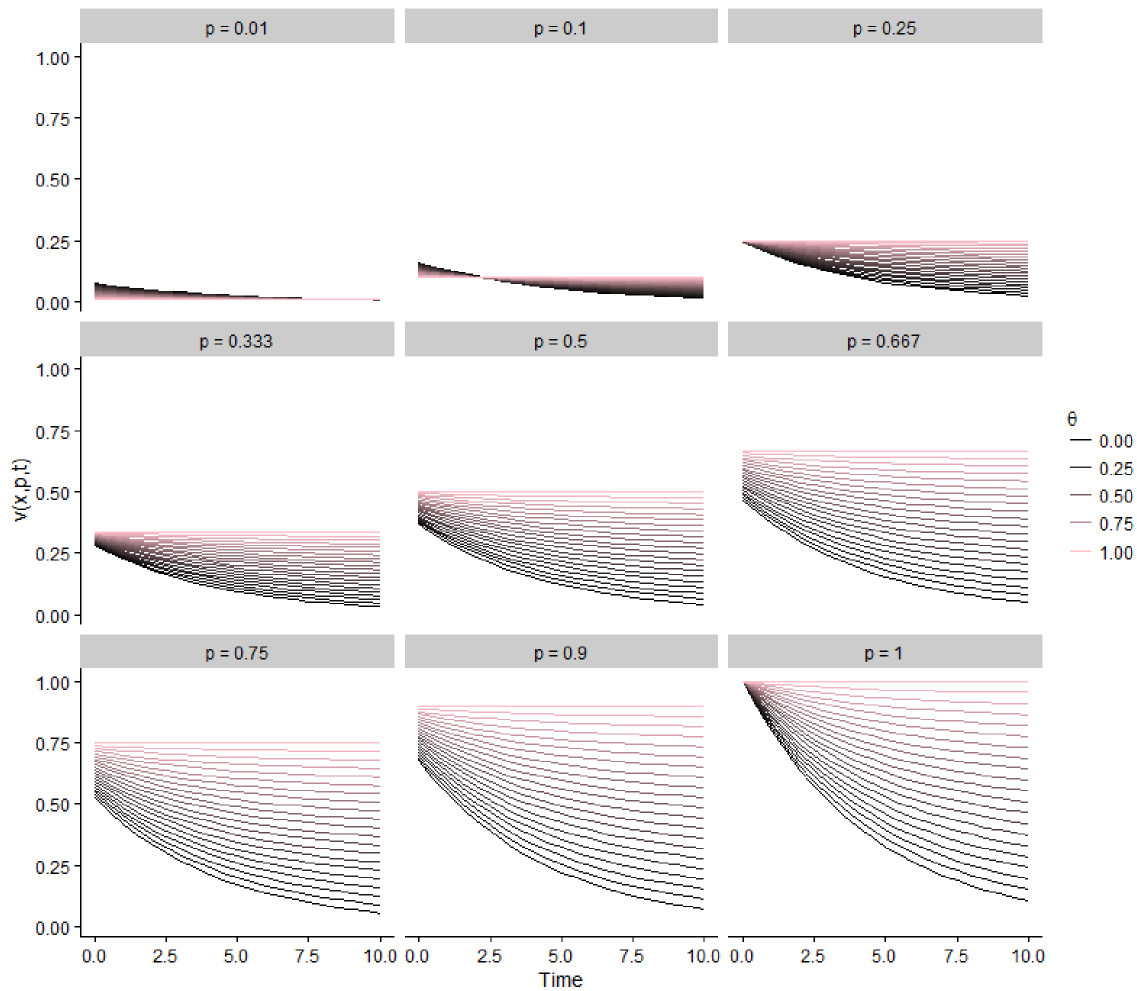
Figure I: Time Dependent Probability Weighting



For stochastic consumption plan (x, p, t) yielding \$$x$ with probability p at time t and 0 otherwise, Figure I plots (5) for different values of θ, different time delays (0 to 10 periods), and different probabilities. The parameters were set to $\delta = 0.8$, and $w(p) = \exp(-(-\ln(p))^{\alpha})$ (Prelec's one-parameter probability weighting function), with $\alpha = 0.5$. These parameter values are arbitrary and chosen for rough plausibility.

Figure II graphs the DPU function from (5) for different values of θ (within each panel) and for different probabilities (across panels), as the time horizon increases from 0 to 10 periods, using the same parametric specification as in Figure I. This specification may be viewed as a probability-dependent time discounting function for the special case where $u_1(.) = u_2(.)$. In Figure II, the function becomes steeper at higher probabilities, suggesting people are less patient as the outcome becomes more likely to be received, possibly reflecting anticipation prior to a reward.

Figure II. Probability Dependent Time Discounting



For a simple stochastic consumption plan (x, p, t) yielding outcome \$$x$ with probability p at time t and 0 otherwise, Figure II plots the function in (5) for different values of θ, different time delays (0 to 10 periods), and different probabilities for the parameters $\delta = 0.8,$ and $w(p) = \exp(-(-\ln(p))^{\alpha})$ with $\alpha = 0.5$.

VI. OTHER RELATIONSHIPS BETWEEN RISK AND TIME PREFERENCE

We next consider four other relationships between risk and time: risk preference and intertemporal substitution, a preference for diversifying risks across time, aversion to timing risk, and correlations between risk preference, time preference, and cognitive type.

VI.A. Risk Preference and Intertemporal Substitution

The discounted expected utility model uses the same utility function in both risky and temporal contexts. However, risk preference and inter-temporal substitution are often observed to be distinct (e.g., Miao and Zhong, 2015). Consider the simple stochastic consumption plan, f, below, also considered by Miao and Zhong (2015), subject to $(1+r)c_1 + c_2 = 100$ and $(1+r)c_1' + c_2' = 100$, where $r \in (0,1)$ is an interest rate.

Figure III. A Simple Stochastic Consumption Plan

	t_1	t_2
p	c_1	c_2
$1-p$	c_1'	c_2'

The present equivalents $\mathrm{PE}(c_1, c_2)$ and $\mathrm{PE}'(c_1', c_2')$ of consumption (c_1, c_2) and (c_1', c_2'), respectively, are determined such that PE/PE' at t_1 is indifferent under V to receiving $(c_1, c_2)/(c_1', c_2')$ on the time horizon. They are defined as

$$PE(c_1, c_2) := V^{-1}\big(\,(1-\theta)(u(c_1) + \delta u(c_2)) + \theta(c_1 + c_2)\big).$$

$$PE'(c_1', c_2') := V^{-1}\big((1-\theta)(u(c'_1) + \delta u(c'_2)) + \theta(c'_1 + c'_2)\big).$$

Employing rank-dependent probability weighting to aggregate the certainty equivalent as in the Chew-Epstein-Halevy approach (see Miao and Zhong, 2015), the certainty equivalent (CE) under DPU can be expressed as

$$CE(f) = V^{-1}\Big(w(p)V\big(PE(c_1, c_2)\big) + w(1-p)V\big(PE(c_1', c_2')\big)\Big)\, if\ PE \geq PE'$$

$$CE(f) = V^{-1}\Big(w(1-p)V\big(PE(c_1, c_2)\big) + w(p)V\big(PE(c_1', c_2')\big)\Big)\, if\ PE \leq PE'.$$

This approach permits a separation between risk attitude (which is partially determined by w) and inter-temporal substitution (which does not depend on w).

VI.B. Preference for Diversification across Time

Miao and Zhong (2015) provide a variant of the example shown below:

Figure IV. Preference for Diversification across Time (A Preferred to B)

Option A	$t=0$	$t=1$
$p=0.5$	100	0
$1-p=0.5$	0	100

Option B	$t=0$	$t=1$
$p=0.5$	100	100
$1-p=0.5$	0	0

We can think of the consumption sequences as being determined by the toss of a fair coin. Then Option A pays $100 in period 0 if the coin lands heads, and it pays $100 in period 1 if the coin lands tails. In contrast, Option B pays $100 in period 0 and $100 in period 1 if the coin lands heads, and it pays $0 in both periods if the coin lands tails. Miao and Zhong (2015) propose and find experimental support for the hypothesis that many people prefer Option A in which risks are diversified across time over Option B in which they are not. Such behavior has also been observed by Andersen et al. (2011) who refer to this preference pattern as 'correlation aversion' or 'intertemporal risk aversion.'

Correlation aversion is simply explained by DPU. Note that, for Option A, System 1 will rank consumption sequence $x \coloneqq (100, t=0; 0, t=1)$ higher than the sequence $y \coloneqq (0, t=0; 100, t=1)$ in order of preference since $\delta < 1$. Thus, DPU assigns weight $\pi(0.5)$ to x and $(1-\pi(0.5))$ to y, with weights assigned analogously for Option B. In most experimental studies of rank-dependent probability weighting functions (see references in §5.2), it has been found that $\pi(0.5) < 0.5$. Under DPU, with $u(0)=0$:

$$V(Option\ A) = (1-\theta)\Big(\big(\pi(0.5)\big)u(100) + \big(1-\pi(0.5)\big)\delta u(100)\Big) + \theta(100)$$

$$V(Option\ B) = (1-\theta)\Big(\big(\pi(0.5)\big)\big(u(100) + \delta u(100)\big)\Big) + \theta(100)$$

Since $\delta \in (0,1)$, $u(100) > \delta u(100)$. Hence, A is preferred to B if $\pi(0.5) < 0.5$.

VI.C. Aversion to Timing Risk

Onay et al. (2007) and DeJarnette et al. (2015) experimentally investigate preferences over lotteries that pay a fixed prize at an uncertain date. For instance, in choices such as receiving $100 in 10 weeks for sure (Option A), or receiving $100 in either 5 or 15 weeks with equal probability (Option B), they find that people are generally risk-averse toward timing risk, preferring Option A. However, DEU and the standard models of hyperbolic and quasi-hyperbolic discounting imply people will be *risk-seeking* toward timing risk.

Consider a choice between receiving \$100 at time t (Option A), or \$100 at either time $t-r$ or time $t+r$ with equal probability (Option B). Under DPU, the values are:

$V(A) = (1-\theta)\delta^t u(100) + \theta(100)$.

$V(B) = (1-\theta)\left(\delta^{t-r}\pi(0.5)u(100) + \delta^{t+r}\big(1-\pi(0.5)\big)u(100)\right) + \theta(100)$.

For all $\theta \in [0,1)$, Option A is preferred to Option B if the following inequality holds:

(6) $$1 > [\delta^{-r}\pi(0.5) + \delta^{r}\big(1-\pi(0.5)\big)].$$

The above inequality can hold given $\pi(0.5) < 0.5$, a robust finding, noted in §5.2.

VI.D. Risk Preference, Time Preference, and Cognitive Type

The DPU model also captures observed relationships between risk preference, time preference, and cognitive reflection. An agent's 'cognitive type', as parameterized by θ can be interpreted as a measure of reliance on System 2 processing which may be correlated with cognitive reflection or cognitive skills. DPU accommodates a continuum of types - any $\theta \in [0,1]$. Note that the DPU specification in (5) predicts the following:

Proposition 10: *For a decision maker with preferences given by* (5)*:*

(i) *The decision maker approaches risk-neutrality as θ increases.*

(ii) *The decision maker becomes more patient as θ increases.*

(iii) *Expected value maximization is negatively correlated with impatience.*

Such correlations between risk neutrality, patience, and cognitive skills have been observed by Frederick (2005), Burks et al. (2009), Oechssler et al. (2009), Cokely and Kelley (2009), Dohmen et al. (2010), and Benjamin et al. (2013). Burks et al. (2009) report "those individuals making choices just shy of risk-neutrality have significantly higher CS [cognitive skills] than those making more either risk-averse or more risk-seeking choices" (p. 7747). However, Andersson et al. (2016) finds no correlation between risk preferences and cognitive skills.

The notion that System 2 is closer to risk-neutrality and is more patient than System 1 is also supported by studies which employ other means of manipulating System 1 versus System 2 processing. Placing people under a high working memory load is one approach to inducing greater reliance on System 1. Studies have found that increased cognitive load (Deck and Jahedi, 2015; Holger et al., 2016) increases deviations from risk-neutrality such as increased small-stakes risk aversion and produces less patient and more impulsive behavior (Shiv and Fedorikhin, 1999). Leigh (1986), Anderhub et al. (2001), and Andersen et al. (2008) also find that risk aversion is positively correlated with impatience. In a large study of response times to the common ratio effect choices of Kahneman and Tversky (1979), Rubinstein (2013) observed slow responders to be significantly more likely to choose the expected value maximizing alternatives in both decisions than fast responders.

Finally, the notion that the population of agents is comprised largely of RDU types and risk-neutral expected utility types is supported by Bruhin et al. (2010) who estimate a finite mixture model for three data sets and found that subjects could be classified with high probability as RDU decision makers or expected value maximizers. Similar heterogeneity was observed by Harrison and Rutstrom (2009) who also found the modal expected utility type to be risk-neutral. Since (5) reduces to risk-neutrality when $\theta = 1$, and (5) reduces to RDU when $\theta = 0$, it can capture the observed distribution of risk preferences. Rather than interpreting the mixture model as proportions of agents who are either RDU *or* EU types, DPU offers a unified perspective in which agents are a mixture of both RDU *and* EU types.

VII. SUMMARY OF RESULTS

Table II. Sufficient Conditions for DPU to Explain Observed Behaviors*

Property	**Parameter Values**
Present bias (Laibson, 1997)	$\theta \in (0,1)$, $\delta_1 < \delta_2$
Delay reduces risk aversion (Baucells & Heukamp, 2010)	$\theta \in (0,1)$, $\delta_1 < \delta_2$
Cognitive type and time preference (Frederick, 2005)	$\theta \in [0,1]$, $\delta_1 < \delta_2$
Cognitive type and risk preference (Frederick, 2005)	$\theta \in [0,1]$, u_1 more concave than u_2
Subendurance (Baucells & Heukamp, 2012)	$\theta \in (0,1), u_1$ more concave than u_2
Magnitude effect (Loewenstein & Prelec, 1991)	$\theta \in (0,1), u_1(z) = z^{\alpha}; u_2(z) = z^{\beta}$
Peanuts effect (Loewenstein & Prelec, 1991)	$\theta \in (0,1), u_1(z) = z^{\alpha}; u_2(z) = z^{\beta}$
Risk reduces impatience (Keren & Roelofsma, 1995)	$\theta \in (0,1), w(0.5) < 0.5$
Diversifying risks across time (Miao & Zhong, 2015)	$\theta \in [0,1), w(0.5) < 0.5$
Aversion to timing risk (DeJarnette et al., 2015)	$\theta \in [0,1), w(0.5) < 0.5$
Separation of risk and time preference (Epstein & Zin, 1989)	$\theta \in [0,1), w(p) \neq p$

*The sufficient conditions for the magnitude effect and the peanuts effect further include $\alpha < \beta \leq 1$, ensuring u_1 is more concave than u_2. Proposition 6 (the peanuts effect) also implicitly requires $w(p) > p$ for the domain of p in which it is observed. This condition holds naturally for small p under a standard inverse-S-shape weighting function (as in prospect theory or RDU) for System 1.

A summary of behaviors explained by DPU is provided in Table II which also displays sufficient conditions[5] for (4) to generate the shifts in preference in Propositions 4 – 10. As noted in §5.2, the condition $w(0.5) < 0.5$ is a standard finding in the literature. Note also that present bias, the magnitude effect, the peanuts effect, and the three interaction effects between risk, time, and money are only explained if θ is strictly between 0 and 1. All properties hold while preserving transitivity, continuity, and stochastic dominance.

VIII. RELATED LITERATURE

Many models for decisions under risk and for decisions over time have been developed in the past five decades and it is not feasible to review them all here. Since models developed for only decisions under risk or for only decisions over time cannot account for the majority of our results, we focus on models which consider both risk and time. The standard discounted expected utility model motivated our analysis and it provides a natural benchmark with which to compare our predictions. Table II may be viewed as a summary of properties of DPU that are not shared by DEU. Prelec and Loewenstein (1991) noted parallels between anomalies for decisions under risk and decisions over time, and Loewenstein and Prelec (1992) provided a general model which accounts for observed violations of DEU such as hyperbolic discounting and the Allais paradox. Rubinstein (1988, 2003) and Leland (1994, 2002) provided models of similarity judgments which explain many of the key anomalies for decisions under risk and over time such as the Allais paradox and hyperbolic discounting as arising from the same cognitive process. However, all of these approaches treat risk and time independently, and thus cannot explain interaction effects between risk and time preferences.

Classic approaches to studying interactions between risk and time preferences can be found in Kreps and Porteus (1978), Epstein and Zin (1989), and Chew and Epstein (1990). Kreps and Porteus consider preferences for the timing of the resolution of uncertainty, an issue not studied here. Epstein and Zin (1989) and Chew and Epstein (1990) provide models which can disentangle risk preferences from the degree of intertemporal substitution. Traeger (2013) introduces a model of intertemporal risk aversion in which a rational agent does not discount the future for reasons of impatience.

[5] For aversion to timing risk, $w(0.5) < 0.5$ in Table II is necessary rather than sufficient. In the other cases, $w(0.5) < 0.5$ is sufficient to explain the shifts in preference in the corresponding examples from Table I.

Recent models by Halevy (2008), Walther (2010) and Epper and Fehr-Duda (2015) focus on implications of rank-dependent utility theory when extended to an intertemporal framework. Halevy (2008) and Walther (2010) focus primarily on relationships between hyperbolic discounting over time and non-linear probability weighting under risk. Halevy notes that his model is also consistent with the experimental evidence of Keren and Roelofsma (1995). The observations of Keren and Roelofsma and Baucells and Heukamp (2010) are both explained by the probability-time tradeoff model of Baucells and Heukamp (2012). However, this model applies only to a restrictive class of prospects offering a single non-zero outcome to be received with probability p at time t.

Aside from extensions of RDU to intertemporal choice, one other major literature stream which has grown rapidly in recent years is the class of dual-selves models motivated to explain temptation and self-control as well as more general choices under risk and over time. In these models, the two families of processes have been characterized as controlled and automatic (Benhabib and Bisin, 2005), long-run and short-run (Fudenberg and Levine, 2006; 2011, 2012), hot and cold (Bernheim and Rangel, 2005), affective and deliberative (Mukherjee, 2010; Loewenstein et al., 2015), and rational and emotional (Bracha and Brown, 2012). In addition, Gul and Pesendorfer, (2001; 2004) model agents who have temptation preferences and commitment preferences.

A leading example in the class of dual-selves models is that of Fudenberg and Levine (2006, 2011, 2012) and Fudenberg et al. (2014) which can explain the Allais paradox as well as the interactions between risk and time preferences identified by Keren and Roelofsma and Baucells and Heukamp. However, Fudenberg et al. (2014) comment "Unfortunately the model of Fudenberg and Levine (2011) is fairly complex, which may obscure some of the key insights and make it difficult for others to apply the model." (p. 56). In addition, a drawback of the model from both a normative and a descriptive viewpoint is that it violates transitivity (Fudenberg et al., 2014), even though transitivity is rarely violated in experiments (Baillon et al., 2014; Regenwetter et al., 2011).

Aside from the work of Fudenberg and Levine, most dual-selves models in economics are restricted to either risk or time. For decisions involving only risk, (5) reduces to a variant of the dual system model (DSM) of Mukherjee (2010). The DPU model in (5) modifies the DSM by employing a rank-dependent probability weighting function[6] for System 1, and extends the model to encompass both risk and time preferences. Rank-dependent weighting for System 1 eliminates the undesirable property that the DSM violates first order stochastic dominance. McClure et al. (2007) and van den Bos and McClure (2013) employ a two-system model of time preference with two discount factors but with a single utility function. Their approach can explain present bias, but not the magnitude effect or the interaction effects involving risk and time. Our results also relate to the finding in the social choice literature that group discount functions are present-biased (Jackson and Yariv, 2015). We show a similar phenomenon in a dual system model of individual choice. However, it should be clear that DPU does not capture all important behaviors for decisions over time. For instance, DPU is additively separable across time periods and so does not account for complementarities in consumption across time which is a hallmark of the classic model of habit formation (e.g., Constantinides, 1990).

IX. EXTENSION TO SOCIAL PREFERENCES

Recent experimental work has identified systematic interaction effects across the domains of risk, time, and social preferences. Yet there is no unifying framework which simultaneously operates across all three domains. Jullien (2016) provides a survey of work demonstrating these interactions and proposes to 'see rationality in 3D'. Jullien distinguishes behaviors 'within' dimensions from behaviors 'across' dimensions, and notes:

> *"'Within' dimensions means that decision problems are of the form, e.g., 'a consequence for sure vs. a bigger consequence with uncertainty' or 'a consequence now vs. a bigger consequence later'", whereas decisions across dimensions include choices such as "'a consequence for sure but later versus another consequence now but with uncertainty.'"*

Jullien argues:

> *"'The proposed distinction between challenges within and across dimensions is more than conceptual, it also delimits a historical rupture between two periods that are nontrivial regarding the debates between behavioral and standard economics. The classical challenges posed by Kahneman, Tversky, Thaler and others focused on interactions within dimensions, posing problems to standard models. The more recent challenges from interactions across dimensions are posing problems to both standard and behavioral economists' models.''*

[6]In the DSM of Mukherjee (2010), the affective system weights all outcomes in the support of a lottery equally which can produce violations of first order stochastic dominance.

Here we show that a natural extension of DPU predicts systematic interaction effects across the dimensions of risk, time, and social preferences. Given the preceding comments, we reach a surprising conclusion: A simple way to model the observed interaction effects across the three decision domains is to *combine* the standard normative and behavioral models. In particular, we propose a parametric dual system model in which System 1 is assumed to have prospect theory risk preferences[7], and to be delay-averse and inequity-averse, whereas System 2 is assumed to have expected utility risk preferences and to be delay-neutral and inequity-neutral[8]. To the extent that issues of fairness are often emotionally charged, it seems plausible that System 1 cares about fairness. The social heuristics hypothesis (Rand et al., 2014, Rand, 2016) also provides a basis for predicting that intuitive responses are often more cooperative than deliberative responses. We can view the model developed here as representing *non-separable rational-behavioral preferences,* since choice alternatives are evaluated by the convex combination of a rational (System 2) value function and a behavioral (System 1) value function.

Formally, we update our notation as follows: There is a finite set, T, of time periods, a finite set $\mathcal{M}$ of outcomes with $\mathcal{M} \subset \mathbb{R}$, a finite set, I, of individuals, and a finite set X of consumption allocations. Consumption allocations are indexed by $j \in \{1,2, \dots, n\}$, time periods are indexed by $t \in \{0,1, \dots, m\}$ and individuals are indexed by $i \in \{1,2, \dots, k\}$. A *consumption allocation* consists of an outcome for each individual $i \in I$ at each time period $t \in \mathrm{T}$. It can be written:

$$x_j \coloneqq \{(x_{j11}, x_{j21}, \dots, x_{jk1}), 1; \dots; (x_{j1m}, x_{j2m}, \dots, x_{jkm}), m\}$$

where x_{jit} is the outcome assigned by consumption allocation x_j to individual i in period t. The decision maker is denoted $i = 1$. A *stochastic consumption allocation* is a lottery over consumption allocations. It is a function $p: \mathrm{X} \to [0,1]$, with p_j the probability it assigns to consumption allocation x_j. Denote the set of stochastic consumption allocations by $\Delta(\mathrm{X})$.

Let $\mathbb{E}[p]$ denote the undiscounted expected value of $p \in \Delta(\mathrm{X})$ to the decision maker. We consider a simple extension of the DPU specification in (5) given by the parametric form in (6):

$$(6) \qquad V(p) = (1-\theta)\left(\textstyle\sum_t \sum_j \delta^t \cdot \pi(p_{jt}) \cdot u(x_{j1t} - \frac{\alpha}{k}\sum_i |x_{jit} - x_{j1t}|)\right) + \theta\mathbb{E}[p].$$

Note that (6) imposes a duality between Systems 1 and 2: System 1 preferences are non-linear in probabilities and payoffs, delay-averse, and inequity-averse, whereas System 2 preferences are

[7] To avoid proliferation of parameters, we continue to model System 1 risk preferences with rank-dependent utility.
[8] Our results in Propositions 11 - 16 also hold if System 2 cares about efficiency rather than being purely selfish.

risk-neutral, delay-neutral and inequity-neutral (None of these conditions on System 2 is necessary for our results; We only require that System 2 is closer to risk-neutrality, more patient, and less inequity-averse than System 1).

In (6), $\alpha \geq 0$ represents the degree of inequity aversion for System 1. One could instead use the two-parameter specification by Fehr and Schmidt (1999) or another model of social preferences. One might further simplify (6) by letting $u(x) = x$, such that both systems have linear utility for choices involving only the decision maker. If $u(x) = x$, then (6) has three domain-specific parameters (one each for the risk, time, and social preferences of System 1), plus the parameter θ representing the agent's 'cognitive type' that operates across domains.

X. Application to Violations of Dimensional Independence

We apply (6) to further study interactions between risk, time, and social preferences. One might consider six pairwise interactions across these domains: (i) risk affects time preference; (ii) time affects risk preference; (iii) risk affects social preferences; (iv) social context affects risk preferences; (v) time affects social preferences; (vi) social context affects time preferences. Additional interaction effects arise when one also considers changes in payoff magnitude as was illustrated with subendurance in Table I. Each of these interaction effects provides a test of the same general principle. This principle, called *dimensional independence* (Keeney & Raiffa, 1993; Bhatia, 2016) states that two attribute dimensions x and y are independent if for all x, y, x', y', an alternative (x, y) is chosen over (x', y) if and only if (x, y') is chosen over (x', y'). This principle reflects the intuition that identical attribute values in a dimension across alternatives will cancel in the evaluation process and not affect decisions. This principle is so basic that it is a general feature of the leading normative and behavioral decision models.

Table III reveals seven violations of dimensional independence. For each example, the table makes the common dimension shared within each choice explicit. We have already seen that DPU predicts the violations that risk affects time preference, that time effects risk preference, and that payoffs interact with risk and time preferences (subendurance), each in the direction observed in prior experiments. We now observe that the extended DPU model in (6) predicts four additional violations of dimensional independence. The examples in the table are used to illustrate the effects predicted by DPU. Similar behaviors have been observed in experiments for three of these four cases. The fourth case – that allocations affect time preferences, has to our knowledge not yet been empirically tested, but we believe it has an intuitive appeal.

Table III. Violations of Dimensional Independence

Behavior	Choice 1	Choice 2
Time affects Risk Preference (Baucells and Heukamp, 2010)	**A. (9, now, 100%)** B. (12, now, 80%)	A. (9, 3 months, 100%) **B. (12, 3 months, 80%)**
Risk affects Time Preference (Keren and Roelofsma, 1995)	**A. (100, now, 100%)** B. (110, 4 weeks, 100%)	A. (100, now, 50%) **B. (110, 4 weeks, 50%)**
Allocation affects Risk Preference (Bolton and Ockenfels, 2010)	**A. (9 self, 100%)** B. (16 self, 50%)	A. (9 self, 16 other, 100%) **B. (16 self, 16 other, 50%)**
Risk affects Social Preferences (Krawczyk & Le Lec, 2010)*	**A. (5 self, 5 other, 100%)** B. (10 self, 0 other, 100%)	A. (5 self, 5 other, 50%) **B. (10 self, 0 other, 50%)**
Allocation affects Time Preferences*	**A. (9 self, now)** B. (12 self, 3 months)	A. (9 self, 12 other, now) **B. (12 self, 12 other, 3 months)**
Time affects Social Preferences (Kovarik, 2009)*	**A. (5 self, 5 other, now)** B. (10 self, 0 other, now)	A. (5 self, 5 other, 1 year) **B. (10 self, 0 other, 1 year)**
Subendurance (Baucells et al., 2009)	**A. (100, for sure, 1 month)** B. (100, with 90%, now)	A. (5, for sure, 1 month) **B. (5, with 90%, now)**

Modal choice patterns from experiments in bold font; * denotes a prototypical example.

The new violations of dimensional independence predicted by (6) imply that both uncertainty and time reduce the propensity for giving in dictator games, and that distributional concerns can shift both risk and time preferences. Of the seven systematic violations of dimensional independence in Table III, five of them *cannot* be explained when either $\theta = 0$ or $\theta = 1$, requiring the interaction between systems within the model presented here. Surprisingly, we observe that the *same* model explains each of the preference patterns in Table I and makes strong directional predictions as it does not predict the reverse preference patterns.

We note that our modeling approach was merely intended to provide a formal representation of decision making that accounts for both System 1 and System 2 processes. We observe that a simple and even natural specification of this model has the additional property of providing a unified approach to predicting empirical violations of dimensional independence and to modeling interactions between risk, time, and social preferences.

To analyze the examples from Table III that involve distributional concerns more generally, we consider choices over stochastic consumption allocations of the form $p \coloneqq \{(x_1, y_1), t_1, p_1; (x_2, y_2), t_2, 1 - p_1\}$ when payoffs of another person are also involved (who receives y_1 or y_2). For choices involving only the decision maker, we will use the simpler notation (x, t, p) for a stochastic consumption allocation yielding x to the decision maker in period t with probability p. In both cases, individuals receive 0 in all other periods or events.

X.A. Allocations Interact with Risk Preferences

Bolton and Ockenfels (2010) observed a modal preference for 9 Euros with certainty over a 50% chance of 16 Euros. However, they also observed a modal preference for a 50% chance that the decision maker and a passive recipient each receive 16 Euros (and a 50% chance they each receive nothing) over the decision maker receiving 9 Euros and the recipient receiving 16 Euros with certainty. This preference pattern can hold under (6), due to System 1's inequity aversion even though neither system is risk-seeking toward gains of moderate or high probabilities. Similar behavior in which the social context shifts risk preferences toward less inequity has been observed by Leder and Betsch (2016).

Definition 7 (Effect of allocation on risk preference): We say *allocation interacts with risk preference* if for $x > y > 0, (x, t, p) \sim (y, t, 1) \Longrightarrow \big((x, x), t, p; (0,0), t, p\big) \succ ((y, x), t, 1)$

Proposition 11 (Effect of allocation on risk preference): *Let* $\mathbb{E}[(x, t, p)] > \mathbb{E}[(y, t, 1)$. *Then allocation interacts with risk preference for all* $\theta \in [0,1)$.

X.B. Risk Interacts with Social Preferences

A novel implication of DPU is that risk will interact with social preferences. Under DPU, a decision maker indifferent between splitting $10 evenly with a recipient or keeping all $10 will prefer a 50-50 chance of allocation ($10, $0) or ($0, $0) over a 50-50 chance of ($5, $5) or ($0, $0). That is, introducing risk into a dictator game is predicted to reduce inequity aversion. While we have not seen this precise example, 'probabilistic' dictator games have been conducted by Krawczyk and Le Lec (2010) and Brock et al. (2013). Both studies find that introducing risk into a dictator game decreases giving by the dictator. Exley (2016) also found less charitable giving under risk. Definition 8 provides a simple formalization of reduced dictator giving under risk.

Definition 8 (Risk affects social preferences): *Risk interacts with social preferences* if for $0 < y < x, \big((x, 0), t, 1\big) \sim \big((x - y, y), t, 1\big) \Longrightarrow \big((x, 0), t, 0.5\big) \succ \big((x - y, y), t, 0.5\big)$.

Proposition 12 (Risk affects social preferences): *Let* $w(0.5) < 0.5$. *Then risk interacts with social preferences if and only if* $\theta \in (0,1)$.

X.C. Allocations Interact with Time Preferences

The DPU model further predicts that the social context will affect time preference. An illustrative example from Table III is that a decision maker indifferent between receiving 100 guilders today and 110 guilders in 4 weeks is predicted to strictly prefer an allocation in which he and another person each receive 110 guilders in 4 weeks over an allocation in which he receives 100 guilders today and the other person receives 110 guilders today. That is, changes in allocations shift preferences toward consumption sequences with less inequity.

Definition 9 (Effect of allocation on time preference): We say *allocations interact with time preference* if for $> y > 0$, $(x, t, p) \sim (y, 0, p) \Rightarrow \big((x, x), t, p\big) \succ \big((y, x), 0, p\big)$.

Proposition 13 (Effect of allocation on time preference): *Allocations interact with time preference for all* $\theta \in [0,1)$.

X.D. Time Interacts with Social Preferences

The DPU model also predicts that time will interact with social preferences. For instance, under DPU, a person indifferent between splitting \$10 evenly today with a recipient or keeping all \$10 for himself will strictly prefer to keep all \$10 when the money is to be received after one year. That is, introducing delays into a dictator game reduces inequity aversion. In an experimental study on a 'temporal' dictator game, Kovarik et al. (2009) found that longer delays decrease giving by the dictator. This finding was also observed by Dreber et al. (2014).

Definition 10 (Time affects social preferences): *Time affects social preferences* if for all $0 < t < r;\ 0 < y < x, \big((x, 0), t, p\big) \sim \big((x - y, y), t, p\big) \Rightarrow \big((x, 0), r, p\big) \succ \big((x - y, y), r, p\big)$.

Proposition 14 (Time affects social preferences): *Time interacts with social preferences if and only if* $\theta \in (0,1)$.

Thus, DPU predicts all seven violations of dimensional independence illustrated in Table III. Note that in Propositions 4, 5, 6, 7, 8, 9, 12, and 14, it is necessary that $\theta \in (0,1)$ for the results to hold. Hence, these effects are not explained by rational or behavioral preferences alone.

XI. Cognitive Type and Social Preferences

The DPU model also makes novel predictions regarding the social preferences of agents with different levels of System 2 processing as parameterized by θ. Consider two agents of type θ_1 and θ_2, where $\theta_1 < \theta_2$, with preferences given by (6) and with the same System 1 preferences that have $\alpha > 0$. Let $\succsim_{\theta_1}$ and $\succsim_{\theta_2}$ denote the preferences of these decision makers with analogous notation for strict preference and indifference. Since agents with higher values of θ are less inequity-averse, DPU generates novel predictions for both the dictator game and the ultimatum game. In the dictator game, one participant (the dictator) decides how to allocate a fixed amount of money between himself and a passive recipient. In the dictator game, DPU predicts that agents with high cognitive types (higher θ) give less than agents with lower values of θ. Formally:

Proposition 15 (Cognitive Type and Giving in the Dictator Game): *As the dictator in a dictator game, the propensity to give money decreases with θ: For any $0 < y < x$, if $(x, 0) \sim_{\theta_1} (x - y, y)$ then $(x, 0) \succ_{\theta_2} (x - y, y)$.*

Empirical support for Proposition 15 comes from an experimental study by Ponti and Rodriguez-Lara (2015) who administered the cognitive reflection test (CRT) due to Frederick (2005) to players that also participated in a dictator game. The CRT is a three-item measure where each question has an intuitive but incorrect answer, and a correct answer which requires a moment of reflection. The test is designed to identify decision makers who rely more on System 1 versus System 2 processing, with higher scores relying more on reflective thinking (System 2). Consistent with Proposition 15, subjects with higher scores on the CRT gave less in the dictator game. Ponti and Rodriguz-Lara comment "Impulsive Dictators show a marked inequity aversion attitude," and that "Reflective Dictators show lower distributional concerns, except for the situations in which the Dictators' payoff is held constant." Cueva et al. (2016) and Capraro et al. (2017) also found low CRT subjects to be more inequity-averse than high CRT subjects in dictator game experiments. Schulz et al. (2014) likewise found that subjects under high cognitive load (a means to increase reliance on System 1) also gave more in a dictator game experiment.

The ultimatum game, a close relative of the dictator game, involves two players who make sequential decisions. The first mover decides how much of a fixed sum to offer the other player. If the other player accepts, the proposed offer is implemented. If the other player rejects the offer, both players receive nothing. DPU makes the following prediction for the ultimatum game:

Proposition 16: (Cognitive Type and Accepting in the Ultimatum Game): *As a responder in the ultimatum game, acceptance of unfair offers increases with θ: For any $0 < y < x$, if $(y, x - y) \sim_{\theta_1} (0,0)$, then $(y, x - y) \succ_{\theta_2} (0,0)$.*

That is, DPU predicts people who rely more on System 2 to be more likely to accept unfair offers (i.e., offers with a larger amount for the proposer) in the ultimatum game than those who rely more on System 1. Empirical support for Proposition 16 comes from experiments by Neys et al. (2011) and Calvillo and Burgeno (2015) who each found that higher scoring participants on the cognitive reflection test were more likely to accept unfair ultimatum game offers.

The extended DPU model in (6) is not a complete model of social preferences since it does not account for procedural fairness or preferences for efficiency[9]. Still, it offers novel predictions of how distributional preferences relate to risk and time preferences, and to the decision maker's thinking style.

XII. CONCLUSION

Economic models of choice are great simplifications of the complexity of actual decisions, compressing all information about risk and time preferences into a utility function and a discount factor. In the present paper, an additional parameter is given to the decision maker: Everything that affects a person's thinking style (whether the agent relies more on feeling and intuition or logic and calculation) is summarized by the parameter θ representing one's 'cognitive type', which is derived from our assumptions in Propositions 1 and 2. We saw that any $\theta \in (0,1)$ explains present bias, the magnitude effect, and the peanuts effect, as well as the observed interaction effects between risk and time preferences (risk reduces impatience, delay reduces risk aversion, and payoffs interact with risk and time preferences). Moreover, none of these effects is explained under DPU when either $\theta = 0$ or $\theta = 1$, motivating the need for a dual process model that involves both systems. We demonstrated that DPU also predicts a preference for diversification across time, aversion to timing risk, a separation between risk preference and intertemporal substitution and observed correlations between risk preferences, time preferences, and cognitive types. In addition, DPU makes strong predictions as it rules out the opposite preference patterns. Moreover, these observations hold while preserving transitivity, continuity, and stochastic dominance. As DPU is linear in θ, the model may be analytically convenient, for instance, when allowing for heterogeneity in cognitive types. We provide a simple illustration of such an analysis in Appendix A.

The DPU model was developed for the purpose of formalizing behaviors based on System 1 and System 2 processes which are often discussed qualitatively. We have shown that one natural approach to constructing such a model (in which System 1 has behavioral preferences and

[9] For promising approaches to integrating procedural and distributional preferences, see for example Saito (2013).

System 2 has rational preferences) also predicts empirical violations of the dimensional independence axiom, as well as systematic interaction effects between risk, time, and social preferences and observed correlations between cognitive types and social preferences. Moreover, in Propositions 4, 5, 6, 7, 8, 9, 12, and 14, it is necessary to have the interaction between systems ($\theta \in (0,1)$) for the results to hold. Hence, these effects are not explained by rational or behavioral preferences alone. In addition to providing a unified approach to risk and time (and social) preferences, DPU provides a unification of models based on the rational economic agent, models based on prospect theory or rank-dependent utility and dual system or dual selves models of behavior.

Three prominent types of tradeoffs in decision making are between risk and expected return, between immediate and delayed rewards, and between other-regarding and selfish behavior. Traditionally, these three tradeoffs have been studied separately with studies investigating risk preferences, or time preferences, or social preferences. The approach proposed here provides a unifying framework that accounts for these specific tradeoffs as well as how they interact. As the volume of research in these areas is fast expanding and it may be too early to identify 'canonical' effects, we view DPU as more of a theoretical framework for generating novel predictions and guiding new experiments than as a definitive theory. Theoretical approaches to integrating risk, time, and social preferences must start somewhere, and DPU provides a simple unifying approach to portray rationality in three dimensions.

Economic Science Institute, Chapman University

APPENDIX A: APPLICATION TO CONSUMER BEHAVIOR

We provide a simple application of DPU from (5) which exploits the fact that DPU is linear in θ. We consider a setting studied in Baucells et al. (2016), in which consumers strategically determine their decision to buy a product today, or wait for a sale with the possibility that the product will be sold out. Baucells et al. formulate the problem as a Stackelberg-Nash game with a continuum of consumers in which a retailer, anticipating the best responses of consumers is the 'leader', setting its equilibrium discount percentage, and the consumers, taking into account the actions of the other consumers and the resulting product availability risk are the 'followers.'

As in Baucells et al. (2016), we consider a game with one retail seller and a continuum of consumers with total mass $\lambda > 0$. Consumers have identical DPU parameters given by (4) except for θ which is private information and is drawn independently from a distribution with cdf $F(\theta)$ that is continuous with support $[0,1]$. That is, we allow for heterogeneity in cognitive types. Both λ and F are common knowledge. The retailer has an initial inventory Q of a homogeneous, perishable, and infinitely divisible product that cannot be replenished and must be depleted over two periods, where time 0 is the 'retail price period' (Period 1) and some $t > 0$ is the 'sale period' (Period 2) in which all products unsold at the tag price, $p \in [0,1]$ before Period 2 are marked down by the retailer according to a discount percentage $d \in [0,1]$. The quality[10] of the product is $x \in [0,1]$ and the probability the product will be available[11] in Period 2 is $q \in [0.5,1]$. Let λ_1 and λ_2 denote the mass of consumers who prefer to buy in Period 1 or in Period 2, respectively. We consider the case where $\lambda_1 < Q < \lambda_1 + \lambda_2$ (there is enough inventory to meet the demand of all consumers who prefer to buy now, but not enough to additionally satisfy the demand of all consumers who prefer to wait). We also let $x > p$ so that both buy now and wait are profitable strategies for consumers. If $\lambda_1, \lambda_2 > 0$, then the probability of obtaining the item in Period 2 is $q = \min(Q - \lambda_1)/\lambda_2, 1)$. Under DPU, the values of 'buy now' and 'wait' are:

$$V(buy\ now) = (1-\theta)u_1(x-p) + \theta u_2(x-p)$$

$$V(wait) = (1-\theta)u_1\big(x-p(1-d)\big)w(q)\delta_1^t + \theta u_2\big(x-p(1-d)\big)q\delta_2^t$$

where $w(q) < q$ for all $q \in [0.5,1]$.

[10] As in Bordalo et al. (2013), we let quality and price be measured in the same units (e.g., money), but in our analysis, they are normalized to a 0-1 scale. We also let u_1 and u_2 each be normalized to a 0-1 scale.

[11] Setting $q \geq 0.5$ enables System 1 to be more risk-averse and more delay-averse than System 2 while being consistent with typical estimates of prospect theory weighting functions ($w(q) < q$ for all $q \geq 0.5$).

Excluding indifferences, there are four cases to consider: (i) Both systems prefer to buy now; (ii) Both systems prefer to wait; (iii) System 1 prefers to buy now and System 2 prefers to wait; (iv) System 1 prefers to wait and System 2 prefers to buy now. For all $q \geq 0.5$, case (iv) never arises given the correlations between risk and time preferences and cognitive type, since System 2 is more patient and less risk-averse than System 1. In this case, for any level of delay and product availability risk such that System 1 prefers to wait, System 2 also prefers to wait. We will focus on case (iii) as it involves an interesting conflict between System 1 and System 2.

In case (iii), $V(buy\,now)$ and $V(wait)$ are linear functions of θ with $V(buy\,now)$ having smaller slope and higher intercept. This implies that there is a unique threshold $r \in [0,1]$ such that consumers will wait if $r \leq \theta \leq 1$, and buy now if $0 \leq \theta < r$.

Suppose there exists a symmetric equilibrium in pure strategies, in which all consumers use the same threshold, $r \in [0,1]$. To identify this equilibrium we assume all consumers use this threshold, determine each consumer's best response according to a threshold $b(r) \in [0,1]$, and employ the equilibrium condition $b(r) = r$. Denote any solution by r^*.

Proposition 17 (Nash equilibrium strategy for consumers): *Let System 1 prefer to buy now and System 2 prefer to wait and assume all consumers use threshold r. Then, $\lambda_1 = \lambda F(r)$, and $\lambda_2 = \lambda\big(1 - F(r)\big)$. The best response of each consumer is to wait if $b(r) \leq \theta \leq 1$ and buy now if $0 \leq \theta < b(r)$, where $b(r) = \gamma_1/(\gamma_1 + \gamma_2)$, and*

$$\gamma_1 = u_1(x-p) - u_1\big(x - p(1-d)\big)w(q)\delta_1^t,\ \gamma_2 = u_2\big(x - p(1-d)\big)q\delta_2^t - u_2(x-p).$$

In Proposition 17, $q = \min(Q - \lambda_1)/\lambda_2, 1)$, γ_1 is the net surplus to System 1 from buying now, γ_2 is the net surplus to System 2 from waiting, and $\gamma_1, \gamma_2 > 0$ since we are analyzing case (iii). A symmetric equilibrium in pure strategies exists since $b(r)\colon [0,1] \to [0,1]$ is a continuous mapping from a closed and convex set into itself and therefore admits at least one fixed point, $b(r^*) = r^*$. The equilibrium is not unique, but r^* values can be Pareto-ranked since a higher q benefits consumers who wait, without affecting those who buy now. Thus, the equilibrium with the highest q is Pareto dominant.

As the Stackelberg leader, the retailer chooses the discount percentage, d, that maximizes expected revenue in equilibrium. The retailer's objective function is given by:

$$\max_d R(d) = \max_d\{p\lambda F(\theta^*) + p(1-d)\beta^t \min\{(\lambda\big(1 - F(\theta^*)\big), Q - \lambda F(\theta^*)\}\}.$$

where the discount factor β reflects the retailer's time value of money.

Given the consumer equilibrium strategies and the retailer's objective function in the Stackelberg-Nash game, for a fixed mass of consumers, $\lambda > 0$, we briefly consider two cases:

(i) The case where the retailer attracts more 'System 1' consumers than 'System 2' consumers for the product (a larger mass of consumers have lower values of θ).

(ii) The case of advertisement framing where a firm decides between using an affectively appealing ad or a more cognitively appealing ad for the product.

System 1 versus System 2 Customers

Proposition 18: *For a given (λ, x, Q, p, d), if the distribution of θ is positively skewed (more mass at low θ values) then in equilibrium, relative to a symmetric distribution:*

(i) *At least as many or more customers buy in Period 1.*

(ii) *The firm receives weakly higher expected revenue.*

Observation (ii) indicates for a fixed (λ, x, Q, p, d), System 1 customers (low θ) are more profitable than System 2 customers (high θ). This holds since the firm earns more revenue per purchase in Period 1 than in Period 2.

Advertisement Framing

A product can be advertised or framed by appealing more to System 1 (affective ads which promote positive feelings about the product), or by appealing more to System 2 (cognitively appealing ads which provide facts and justifications to promote the product).

Proposition 19: *For a given (λ, x, Q, p, d), if the distribution of θ has greater mass at lower values for affective ads relative to cognitively appealing ads, then in equilibrium:*

(i) *At least as many or more customers buy the product in Period 1 if the firm employs the affectively appealing ad than if it uses the cognitively appealing ad.*

(ii) *Affective ads generate weakly higher expected revenue than cognitive ads.*

Implications (i) and (ii) can hold either if there is a framing effect in which affective ads reduce θ (by increasing reliance on feelings) for a given consumer, *or* if ad framing affects the distribution of θ such that affective ads attract a larger proportion of consumers with low θ values and cognitive ads attract a larger proportion with high θ values. This latter case holds even if each consumer has a fixed θ and is consistent with experimental evidence that messages with affective appeals are more persuasive when the message recipient is affectively oriented, and that messages with cognitive appeals are more persuasive when the recipient is cognitively oriented (Mayer and Tormala, 2010).

Appendix B: Proofs of Propositions

Proof of Proposition 2: The proof follows straightforwardly from extending the argument in Keeney and Nau (2011), Theorem 2 (although they work in a context of group decisions under uncertainty). We provide a proof following their approach: Assumption 4* implies the decision maker is indifferent between any two lotteries over signals that yield, for each system, the same marginal distribution of expected payoffs. This means the decision maker's preferences among lotteries over signals satisfy the assumption of mutual independence in the utility sense (Fishburn 1965). Then Fishburn's Theorem 2 implies the von Neumann-Morgenstern utility function μ that represents the agent's preferences has the additive representation $\mu(F) = \mu_1(F_1) + \mu_2(F_2)$ for functions μ_1 and μ_2. Applying Assumption 4* again to choices among lotteries over signals which vary only the marginal distribution of payoffs to System s, it follows that $\mu_s(F_s)$ is a von Neumann-Morgenstern utility function that represents the preferences of System s among stochastic consumption plans. This means, up to the addition of an arbitrary constant, μ_1 is an increasing linear function of $(\boldsymbol{\pi_f} \cdot \boldsymbol{v_1})$ and analogously, μ_2 is an increasing linear function of $(\boldsymbol{f} \cdot \boldsymbol{v_2})$**.** We can thus write $\mu_s(F_s) = \theta_s(F_s)$ for some scalar θ_s, and without loss of generality, given the non-triviality conditions $f \succ_1 g$ and $f' \succ_2 g'$ for some $f, g, f', g' \in \Omega$, we can scale the $\theta_s{}'s$ so that $\theta_1 + \theta_2 = 1$ where $\theta_s > 0$ for $s \in \{1,2\}$. Next, write $\theta_2 \equiv \theta$. Then for any F that corresponds to the choice of some $f \in \Omega$, we have $\mu(F) = (1-\theta)F_1 + \theta F_2 = (1-\theta)V_1(f) + \theta V_2(f) = V(f)$. ■

Our last step is to prove the uniqueness of θ in Proposition 2. Write $V(f)$ from (5) as $V(f,\theta)$. Then $f \sim g$ implies $V(f,\theta) = V(g,\theta) = (1-\theta)V_1(f) + \theta V_2(f) = (1-\theta)V_1(g) + \theta V_2(g)$ which implies $(1-\theta)[V_1(f) - V_1(g)] = \theta[V_2(g) - V_2(f)]$. Given $V_1(f) \neq V_1(g)$, we can without loss of generality let $V_1(f) > V_1(g)$. *Then* $V(f,\theta) = V(g,\theta)$ *implies* $V_2(g) > V_2(f)$. If there is another parameter θ', such that $V(\cdot,\theta')$ represents the same preferences as $V(\cdot,\theta)$, then $V(f,\theta') = V(g,\theta')$ implies $(1-\theta')[V_1(f) - V_1(g)] = (\theta')[V_2(g) - V_2(f)]$. Indifference between f and g as represented by $V(\cdot,\theta)$ is broken in the direction of strict preference by $V(\cdot,\theta')$ if $\theta \neq \theta'$. Thus $V(\cdot,\theta)$ and $V(\cdot,\theta')$ represent the same preferences only if $\theta = \theta'$. ■

It is worth pointing out that the aggregation of non-expected utility preferences is possible here because the preferences are mapped to points in a square. Hence, it does not matter whether the individual value functions are linear in probabilities, provided their numerical values are normalized to a 0-1 scale. Consequently, this approach is general. It extends to arbitrary (transitive) preference functionals and to an arbitrary number of agents (or systems), provided that the preferences of the n group members or systems are mapped to points in the hypercube $[0,1]^n$.

In the proofs of Propositions 4 – 9, the agent is assumed to have DPU preferences in (5).

Proposition 4: Present bias holds if and only if $\theta \in (0,1)$.

Proof: (Sufficiency) We need to show that (7) implies (8):

(7) $V(y,p,0) = (1-\theta)w(p)u(y) + \theta py = V(c,p,\Delta) = (1-\theta)\delta^{\Delta}w(p)u(c) + \theta pc$

(8) $(1-\theta)\delta^{t}w(p)u(y) + \theta py < (1-\theta)\delta^{t+\Delta}w(p)u(c) + \theta pc$

Note that since $c > y$, equation (8) implies that $w(p)u(y) > \delta^{\Delta}w(p)u(c)$.

Also note that (7) can be rewritten as: $(1-\theta)\left(w(p)u(y) - \delta^{\Delta}w(p)u(c)\right) = \theta p(c-y)$

In addition, (8) can be rewritten as: $(1-\theta)\delta^{t}\left(w(p)u(y) - \delta^{\Delta}w(p)u(c)\right) < \theta p(c-y)$

Thus, $(1-\theta)\delta^{t}\left(w(p)u(y) - \delta^{\Delta}w(p)u(c)\right) < (1-\theta)\left(w(p)u(y) - \delta^{\Delta}w(p)u(c)\right)$.

The above inequality holds since $w(p)u(y) > \delta^{\Delta}w(p)u(c)$.

(Necessity) Under DPU, the agent has a constant discount factor if $\theta = 0$ or $\theta = 1$. ■

Proposition 5: For a concave power function u, the magnitude effect holds if and only if $\theta \in (0,1)$.

Proof: (Sufficiency) We need to show that (9) implies (10):

(9) $V(y,p,t) = (1-\theta)\delta^{t}w(p)u(y) + \theta py = V(c,p,s) = (1-\theta)\delta^{s}w(p)u(c) + \theta pc$

(10) $(1-\theta)\delta^{t}w(p)u(ry) + \theta pry < (1-\theta)\delta^{s}w(p)u(rc) + \theta prc$

Note that since $c > y$, equation (9) implies that $\delta^{t}w(p)u(y) > \delta^{s}w(p)u(c)$.

Also note that (9) can be rewritten as

(11) $(1-\theta)w(p)\big(\delta^{t}u(y) - \delta^{s}u(c)\big) = \theta p(c-y)$

Inequality (10) can be rewritten as: $(1-\theta)w(p)\big(\delta^{t}u(ry) - \delta^{s}u(rc)\big) < \theta pr(c-y)$

For concave power utility, (i.e., $u(z) = z^{\alpha}$, with $z > 0, \alpha < 1$), this inequality becomes

(12) $(1-\theta)r^{\alpha}(\delta^{t}w(p)y^{\alpha} - \delta^{s}w(p)c^{\alpha}) < \theta pr(c-y)$.

Note that by (11), we have $(1-\theta)(\delta^{t}w(p)y^{\alpha} - \delta^{s}w(p)c^{\alpha})/\theta p(c-y) = 1$.

Thus, (12) reduces to $r > r^{\alpha}$, which is satisfied since $r > 1$ and $\alpha < 1$.

(Necessity) If $\theta = 0$ or $\theta = 1$, the scaling constant factors out. ■

Proposition 6: Let $\mathbb{E}[(y,p,t)] > \mathbb{E}[(c,q,t)]$. Then for any concave power function u, the peanuts effect holds under DPU if and only if $\theta \in (0,1)$.

Proof: (Sufficiency) We need to show that (13) implies (14):

(13) $V(y,p,t) = (1-\theta)\delta^t w(p)u(y) + \theta py = V(c,q,t) = (1-\theta)\delta^t w(q)u(c) + \theta qc$

(14) $$(1-\theta)\delta^t w(p)u(ry) + \theta pry > (1-\theta)\delta^t w(q)u(rc) + \theta qrc$$

For $\mathbb{E}[(y,p,t)] > \mathbb{E}[(c,q,t)]$, equation (13) implies that $\delta^t w(q)u(c) > \delta^t w(p)u(y)$.

Also note that (13) can be rewritten as

(15) $$(1-\theta)\delta^t\big(w(q)u(c) - w(p)u(y)\big) = \theta(yp - cq)$$

In addition, the inequality in (14) can be rewritten as:

(16) $$(1-\theta)\delta^t\big(w(q)u(rc) - w(p)u(ry)\big) < \theta r(yp - cq)$$

For a concave power utility function over gains, (i.e., $u(z) = z^\alpha$, with $z > 0, \alpha < 1$):

(17) $$(1-\theta)\delta^t r^\alpha (w(q)c^\alpha - w(p)y^\alpha) < \theta r(yp - cq).$$

Note that by (15), we have $(1-\theta)\delta^t\big(w(q)u(c) - w(p)u(y)\big)/\theta(yp - cq) = 1$.

Thus, (17) reduces to $r > r^\alpha$, which is satisfied since $r > 1$ and $\alpha < 1$.

(Necessity) If $\theta = 0$ or $\theta = 1$, the scaling constant factors out. ■

Proposition 7: Let $\mathbb{E}[(c,\alpha p,t)] > \mathbb{E}[(y,p,t)]$. Then time interacts with risk preference if and only if $\theta \in (0,1)$.

Proof: (Sufficiency) We need to show that (18) implies (19):

(18) $$(1-\theta)\delta^t w(p)u(y) + \theta py = (1-\theta)\delta^t w(\alpha p)u(c) + \theta\alpha pc$$

(19) $$(1-\theta)\delta^s w(p)u(y) + \theta py < (1-\theta)\delta^s w(\alpha p)u(c) + \theta\alpha pc.$$

Note that since $\mathbb{E}[(c,\alpha p,t)] > \mathbb{E}[(y,p,t)]$, we have $\alpha cp > py$, in which case equation (18) implies that $\delta^t w(p)u(y) > \delta^t w(\alpha p)u(c)$. Also note that (18) can be rewritten as

$$(1-\theta)\delta^t\big(w(p)u(y) - w(\alpha p)u(c)\big) = \theta p(\alpha c - y)$$

In addition, note that the inequality in (19) can be rewritten as:

$$(1-\theta)\delta^s\big(w(p)u(y) - w(\alpha p)u(c)\big) < \theta p(\alpha c - y)$$

Thus, $(1-\theta)\delta^s\big(w(p)u(y) - w(\alpha p)u(c)\big) < (1-\theta)\delta^t\big(w(p)u(y) - w(\alpha p)u(c)\big)$.

The above inequality holds since $w(p)u(y) > w(\alpha p)u(c)$.

(Necessity) If either $\theta = 0$ or $\theta = 1$, the discount factors in (18) and (19) each cancel. ■

Proposition 8: For any convex weighting function w, risk interacts with time preference if and only if $\theta \in (0,1)$.

Proof: (Sufficiency) We need to show that (20) implies (21):

(20) $(1-\theta)\delta^t w(p)u(y) + \theta py = (1-\theta)\delta^{t+\Delta} w(p)u(c) + \theta pc$

(21) $(1-\theta)\delta^t w(q)u(y) + \theta qy < (1-\theta)\delta^{t+\Delta} w(q)u(c) + \theta qc$

Note that since $c > y$, equation (20) implies $\delta^t w(p)u(y) > \delta^{t+\Delta} w(p)u(c)$, and therefore $u(y) > \delta^\Delta u(c)$. Also note that (20) can be rewritten as:

$$(1-\theta)\delta^t w(p)\left(u(y) - \delta^\Delta u(c)\right) = \theta p(c-y)$$

Note that (21) can be rewritten as: $(1-\theta)\delta^t w(q)\left(u(y) - \delta^\Delta u(c)\right) < \theta q(c-y)$

Then by (20), $\left((1-\theta)\delta^t\left(u(y) - \delta^\Delta u(c)\right)\right)/\theta(c-y) = p/w(p)$.

By (21), $\left((1-\theta)\delta^t\left(u(y) - \delta^\Delta u(c)\right)\right)/\theta(c-y) < q/w(q)$. Thus, if $w(q)/w(p) < q/p$ then (20) implies (21). Since $q \in (0,p)$, we can write $q = kp$, for $k \in (0,1)$. For any convex w with $w(0) = 0$, we have $w(kp + (1-k)0) < kw(p) + (1-k)w(0)$, which implies $w(q)/w(p) < q/p$.

(Necessity) If either $\theta = 0$ or $\theta = 1$, the probability weights cancel in (20) and (21). ■

Proposition 9: For any concave utility function u, with $u(0) = 0$, subendurance holds if and only if $\theta \in (0,1)$.

Proof: (Sufficiency) We need to show that (22) implies (23):

(22) $(1-\theta)\delta^{t+\Delta} w(p)u(c) + \theta pc = (1-\theta)\delta^t w(\lambda p)u(c) + \theta\lambda pc$

(23) $(1-\theta)\delta^{t+\Delta} w(p)u(y) + \theta py < (1-\theta)\delta^t w(\lambda p)u(y) + \theta\lambda py$

Since $pc > \lambda pc$, equation (22) implies $\delta^t w(\lambda p)u(c) > \delta^{t+\Delta} w(p)u(c)$. Also note that (22) can be rewritten as (24) and (23) can be rewritten as (25):

(24) $(1-\theta)\left(\delta^t w(\lambda p)u(c) - \delta^{t+\Delta} w(p)u(c)\right) = \theta pc(1-\lambda)$

(25) $\theta py(1-\lambda) < (1-\theta)\left(\delta^t w(\lambda p)u(y) - \delta^{t+\Delta} w(p)u(y)\right)$. From (24) and (25):

$$(1-\theta)\left(\delta^t w(\lambda p)u(c) - \delta^{t+\Delta} w(p)u(c)\right)y < (1-\theta)\left(\delta^t w(\lambda p)u(y) - \delta^{t+\Delta} w(p)u(y)\right)c.$$

For all $\theta \in (0,1)$, the above inequality reduces to, $u(c)/c < u(y)/y$. Since $y \in (0,c)$, we can write $y = kc$, for $k \in (0,1)$. For any concave u with $u(0) = 0$, we have $ku(c) + (1-k)u(0) < u(kc + (1-k)0)$ which implies $u(c)/c < u(y)/y$.

(Necessity) If $\theta = 0$ or $\theta = 1$, the utilities cancel in (22) and (23).

In the proofs of Propositions 11 – 14, we assume DPU preferences as in (6).

Proposition 11: *Let* $\mathbb{E}[(x,t,p)] > \mathbb{E}[(y,t,1)]$. *Then allocation interacts with risk preferences for all* $\theta \in [0,1)$.

Proof: We need to show that (26) implies (27):

(26) $(1-\theta)\delta^t u(y) + \theta y = (1-\theta)w(p)\delta^t u(x) + \theta px$

(27) $(1-\theta)(\delta^t u(y-\alpha(x-y)) + \theta y < (1-\theta)w(p)\delta^t u(x) + \theta px.$

Note that since $px > y$, equation (26) implies $u(y) > w(p)u(x)$. Given (26), it is clear that (27) holds if $(1-\theta)(\delta^t u(y-\alpha(x-y)) + \theta y < (1-\theta)\delta^t u(y) + \theta y$, which holds given System 1 is inequity-averse ($\alpha > 0$), and the result follows. ■

Proposition 12: *Let* $w(0.5) < 0.5$. *Then risk interacts with social preferences if and only if* $\theta \in (0,1)$.

Proof: To prove sufficiency, we need to show that (28) implies (29):

(28) $(1-\theta)(\delta^t u(x-\alpha|x|)) + \theta x = (1-\theta)(\delta^t u(x-y-\alpha|x-2y|)) + \theta(x-y)$

(29) $(1-\theta)w(0.5)\left(\delta^t u(x-\alpha|x|)\right) + 0.5\theta x > (1-\theta)w(0.5)(\delta^t u(x-y-\alpha|x-2y|) + 0.5\theta(x-y))$

Note that (28) implies $\delta^t u(x-y-\alpha|x-2y|) > \delta^t u(x(1-\alpha))$.

Note (28) can be written as $\theta y = (1-\theta)\delta^t[u(x-y-\alpha|x-2y|) - u(x(1-\alpha))]$.

Also (29) can be written: $0.5\theta y > (1-\theta)w(0.5)\delta^t[u(x-y-\alpha|x-2y|) - u(x(1-\alpha))]$.

Since $0.5\delta^t\left[u(x-y-\alpha|x-2y|) - u\left(x(1-\alpha)\right)\right] > w(0.5)\delta^t[u(x-y-\alpha|x-2y|) - u(x(1-\alpha))]$, it follows that (29) holds and $\theta \in (0,1)$ is sufficient to generate the effect. Necessity that $\theta \in (0,1)$ follows since the probability weights cancel when evaluating the two alternatives in the special cases of $\theta = 0$ and $\theta = 1$ and risk does not affect social preferences in those cases. ■

Proposition 13: *Allocation interacts with time preference for all* $\theta \in [0,1)$.

Proof: We need to show that (30) implies (31):

(30) $(1-\theta)w(p)u(y) + \theta py = (1-\theta)w(p)\delta^t u(x) + \theta px$

(31) $(1-\theta)w(p)u(y-\alpha(x-y)) + \theta py < (1-\theta)w(p)\delta^t u(x) + \theta px$

Since $x > y$, (30) implies $u(y) > \delta^t u(x)$. Given (30), it is clear that (31) holds if:

(32) $(1-\theta)w(p)u(y-\alpha(x-y)) + \theta py < (1-\theta)w(p)u(y) + \theta py$, which holds given System 1 is inequity-averse, and the result follows. ■

Proposition 14: *Time interacts with social preferences if and only if* $\theta \in (0,1)$.

Proof: To prove sufficiency, we need to show that (33) implies (34):

(33) $(1-\theta)w(p)\left(\delta^t u(x-\alpha|x|)\right) + \theta py = (1-\theta)w(p)\left(\delta^t u(x-y-\alpha|(2\mathrm{y}-\mathrm{x})|)\right)$

(34) $(1-\theta)w(p)(\delta^r u(x-\alpha|x|)) + \theta py > (1-\theta)w(p)(\delta^r u(x-y-\alpha|(2\mathrm{y}-\mathrm{x})|))$

Note that since $x > \mathrm{x} - y$, equation (33) implies $u(x - y - \alpha|2\mathrm{y} - x|) > u(x(1 - \alpha))$. Note that (33) can be rewritten as $\theta p y = (1 - \theta)w(p)\delta^t[u(x - y - \alpha|(2\mathrm{y} - \mathrm{x})|) - u(x(1 - \alpha))]$. Similarly, (34) can be written: $\theta p y > (1 - \theta)w(p)\delta^r[u(x - y - \alpha|(2\mathrm{y} - x)|) - u(x(1 - \alpha))]$. Note that (34) holds since $\delta^r < \delta^t$ for all $\delta \in (0,1)$, and thus $\theta \in (0,1)$ is sufficient to generate the effect. Necessity that $\theta \in (0,1)$ follows since the discount factors of System 1 and System 2 cancel when evaluating the two alternatives in the special cases of $\theta = 0$ and $\theta = 1$. ■

REFERENCES

ABDELLAOUI, M. (2000): "Parameter-free elicitation of utility and probability weighting functions." *Management Science,* 46, 1497-1512.

ABDELLAOUI, M., E. DIECIDUE, AND A. ONCULER (2011): "Risk preferences at different Time periods: An experimental investigation," *Management Science*. 57, 975-987.

ALLAIS, M. (1953): "Le Comportement de l'Homme Rationnel devant le Risque: Critiques des Postulats et Axioms de l'Ecole Americaine." *Econometrica*, 21, 503-546.

ANDERHUB, V., W., GÜTH, U., GNEEZY AND D. SONSINO (2001): "On the interaction of risk and time preferences: An experimental study." *German Econ. Review*, 2, 239-253.

ANDERSEN, S., G. HARRISON, M. LAU AND E. RUTSTROM (2008): "Eliciting risk and time preferences." *Econometrica*, 76, 583-618.

ANDERSEN, S., G. HARRISON, M. LAU AND E. RUTSTROM (2011): "Intertemporal utility and correlation aversion." Manuscript.

ANDERSSON, O., J. HOLM, J.R. TYRAN AND E. WENGSTROM (2016): "Risk aversion relates to cognitive ability: Preference or Noise?" Forthcoming, *Journal of the European Economic Association.*

ANDREONI, J. AND C. SPRENGER (2012): "Risk preferences are not time preferences." *American Economic Review*, 102, 3357-3376.

BAILLON, A., H. BLEICHRODT AND A. CILLO (2014): "A tailor-made test of intransitive choice." *Operations Research*, 63, 198-211.

BALLINGER, T.P., E. HUDSON, L. KARKOVIATA AND N.T. WILCOX (2011): "Saving behavior and cognitive abilities." *Experimental Economics*, 14, 349-374.

BARBERIS, N., A. MUKHERJEE AND B. WANG (2013): "First Impressions: System 1 Thinking and the Cross-section of Stock Returns." *Yale University working paper*.

BAUCELLS, M., F. HEUKAMP AND A. VILLASIS (2009): "Trading-off probability and time: Experimental evidence." *IESE Business School*.

BAUCELLS, M. AND F.H. HEUKAMP (2010): "Common ratio using delay." *Theory and Decision*, 68, 149-158.

BAUCELLS, M. AND F.H. HEUKAMP (2012): "Probability and time trade-off." *Management Science*, 58, 831-842.

BAUCELLS, M., N. OSADCHIY, AND A. OVCHINNIKOV (2016): "Behavioral anomalies in consumer wait-or-buy decisions and their implications for markdown management." Manuscript.

BENHABIB, J. AND A. BISIN (2005): "Modeling internal commitment mechanisms and self-control: A neuroeconomics approach to consumption-saving decisions." *Games and Economic Behavior*, 52, 460-492.

BENJAMIN, D.J., S.A. BROWN AND J.M. SHAPIRO (2013): "Who is 'behavioral'? Cognitive ability and anomalous preferences." *Journal of the European Economic Association*, 11,1231-1255.

BERNHEIM, B.D. AND A. RANGEL (2005): "Behavioral Public Economics: Welfare and Policy Analysis with Fallible Decision Makers." Economic Institutions and Behavioral Economics. P. Diamond and H. Vartiainen. Princeton University Press.

BHATIA, S. (2016). "Decision making in environments with non-independent dimensions." Forthcoming. *Journal of Behavioral Decision Making*.

BLEICHRODT, H. AND J.L. PINTO (2000): "A parameter-free elicitation of the probability weighting function in medical decision analysis." *Management Science*, *46*, 1485-1496.

BOLTON, G.E., AND A. OCKENFELS (2010). "Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment," *American Economic Review*, 100, 628-633.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2013): "Salience and consumer choice." *Journal of Political Economy*, 121, 803-843.

BRACHA, A. AND D.J. BROWN (2012): "Affective decision making: A theory of optimism bias." *Games and Economic Behavior, 75*, 67-80.

BROCK, M.J., A. LANGE, AND E.Y. OZBAY (2013). "Dictating the risk: Experimental evidence on giving in risky environments." *American Economic Review*, 103-415-437.

BRUHIN, A., H. FEHR-DUDA, AND T. EPPER (2010): "Risk and rationality: Uncovering heterogeneity in probability distortion." *Econometrica*, *78*, 1375-1412.

BURKS, S.V., J.P. CARPENTER, L. GOETTE, AND A. RUSTICHINI (2009): "Cognitive skills affect economic preferences, strategic behavior, and job attachment." *Proceedings of the National Academy of Sciences*, 106, 7745-7750.

CALVILLO, D.P. AND J.N. BURGENO (2015). "Cognitive reflection predicts the acceptance of unfair ultimatum game offers." *Judgment and Decision Making*, 10, 332-341.

CAMERER, C., HO., T. AND CHONG, J.K. (2004): "A cognitive hierarchy model of games". *The Quarterly Journal of Economics. 119*, 816-898.

CAPRARO, V., B. CORGNET, A.M. ESPIN, AND R. HERNAN-GONZALEZ (2017): "Deliberation favours social efficiency by making people disregard their relative shares: evidence from USA and India." *Royal Society Open Science*, 4: 160605, 1-12.

CARPENTER, J., M. GRAHAM, AND J. WOLF (2013): "Cognitive ability and strategic sophistication." *Games and Economic Behavior*, 80, 115-130.

CHARNESS, G., AND M. RABIN (2002). "Understanding social preferences with simple tests." *Quarterly Journal of Economics*, 117, 817-869.

CHEUNG, S.L. (2015): "Comment on "Risk Preferences Are Not Time Preferences": On the Elicitation of Time Preference under Conditions of Risk." *The American Economic Review*, 105, 2242-2260.

CHEW, S.H. AND EPSTEIN, L.G. (1990). "Nonexpected utility preferences in a temporal framework with an application to consumption-savings behaviour." *Journal of Economic Theory*, 50, 54-81.

COKELY, E.T., C.M. KELLEY (2009): "Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation." *Judgment and Decision Making*, 4, 20-33.

CONSTANTINIDES, G.M. (1990): "Habit Formation: A resolution of the equity premium puzzle." *Journal of Political Economy*, 98, 519-543.

CORGNET, B., M. DESANTIS, AND D. PORTER (2015): "Revisiting Information Aggregation in Asset Markets: Reflective Learning & Market Efficiency." *ESI Working Paper* 15-15.

CUEVA, D., I. ITURBE-ORMAETXE, E. MATA-PEREZ, G. PONTI, M. SARTARELLI, H. YU, AND V. ZHUKOVA (2016). “Cognitive (ir)reflection: New experimental evidence.” *Journal of Behavioral and Experimental Economics,* 64-81-93.

DECK, C. AND S. JAHEDI (2015): “The effect of cognitive load on economic decision making: A survey and new experiments.” Manuscript.

DEJARNETTE, P., D. DILLENBERGER, D. GOTTLIEB, AND P. ORTOLEVA (2015): “Time Lotteries.” Manuscript.

DENES-RAJ, V. AND S. EPSTEIN (1994): “Conflict between intuitive and rational processing: when people behave against their better judgment.” *Journal of personality and social psychology*, 66, 819 – 829.

DOHMEN, T., A. FALK, D. HUFFMAN, AND U. SUNDE (2010): “Are risk aversion and impatience related to cognitive ability?” *American Economic Review*. 100, 1238-1260.

DREBER, A., D. FUDENBERG, D.K. LEVINE, AND D.G. RAND (2014): “Altruism and Self-control.” Manuscript.

EPPER, T. AND H. FEHR-DUDA (2015): “The Missing Link: Unifying Risk Taking and Time Discounting.” Manuscript.

EPPER, T. AND H. FEHR-DUDA (2016): “A Tale of Two Tails: On the Coexistence of Overweighting and Underweighting of Rare Extreme Events.” Manuscript.

EPSTEIN, L.G. AND S.E. ZIN (1989): “Substitution, risk aversion and the temporal Behavior of consumption and asset returns: A theoretical framework.” *Econometrica,* 57, 937-969.

ERICSON, K.M. AND J. NOOR, (2015): “Delay Functions as the Foundation of Time Preference: Testing for Separable Discounted Utility.” National Bureau of Economic Research, *NBER working paper* No. 21095.

EXLEY, C.L. (2016): “Excusing selfishness in charitable giving: The role of risk,” *The Review of Economic Studies*, 83, 587-628.

FEHR, E. AND K.M. SCHMIDT (1999): “A theory of fairness, competition, and cooperation.” *Quarterly Journal of Economics,* 114, 817-868.

FEHR-DUDA, H., A. BRUHIN, T. EPPER, R. SCHUBERT (2010): “Rationality on the rise: Why relative risk aversion increases with stake size.” *Journal of Risk and Uncertainty*, 40, 147-180.

FISHBURN, P. (1965): “Independence in utility theory with whole product sets.” *Operations Research*, 13, 28-45.

FREDERICK, S. (2005): “Cognitive Reflection and Decision Making.” *Journal of Economic Perspectives*. 19, 25-42.

FREDERICK, S., G. LOEWENSTEIN AND T. O’DONOGHUE (2002): “Time Discounting and Time Preference: A Critical Review.” *Journal of Economic Literature,* 40, 351-401.

FUDENBERG, D., AND D.K. LEVINE (2006): “A dual-self model of impulse control.” *American Economic Review*, 96, 1449-1476.

FUDENBERG, D., AND D.K. LEVINE (2011): “Risk, delay, and convex self-control costs.” *American Economic Journal: Microeconomics*, 3, 34-68.

FUDENBERG, D. AND D.K. LEVINE (2012): “Timing and Self-Control.” *Econometrica*. 80, 1-42.

FUDENBERG, D., LEVINE, D.K. AND Z. MANIADIS (2014): “An approximate dual-self model and paradoxes of choice under risk.” *Journal of Economic Psychology* 41, 55-67.

GONZALEZ, R., AND G. WU (1999): “On the Shape of the Probability Weighting Function.” *Cognitive Psychology* 38, 129-166.

GUL, F. AND W. PESENDFORFER (2001): “Temptation and Self-Control.” *Econometrica.* 69, 1403-1435.

GUL, F. AND W. PESENDFORFER (2004): “Self-control and the theory of consumption.” *Econometrica*. 72, 119-158.

HALEVY, Y. (2008): “Strotz Meets Allais: Diminishing Impatience and the Certainty Effect.” *American Economic Review*. 98, 1145-1162.

HARRISON, G.W. AND E.E. RUTSTRÖM (2009): “Expected utility theory and prospect theory: One wedding and a decent funeral.” *Experimental Economics,* 12, 133-158.

HARROD, R.F. (1948): *Towards a dynamic economics, some recent developments of economic theory and their application to policy* (No. HB171 H28).

HARSANYI, J.C. (1955): “Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility.” *The Journal of Political Economy,* 63, 309-321.

HEY, J.D. AND C. ORME (1994): “Investigating generalizations of expected utility theory using experimental data.” *Econometrica,* 1291-1326.

HOLGER, G., BIELE, G.P., HEEKEREN, H.R. AND H. UHLIG (2016): “Cognitive Load increases Risk Aversion.” SFB Discussion Paper No. SFB649DP2016-011.

HSEE, C.K., AND Y. ROTTENSTREICH (2004): "Music, Pandas, and Muggers: On the Affective Psychology of Value." *Journal of Experimental Psychology. General*. 133, 23-30.

JACKSON, M.O. AND L. YARIV (2015): "Collective Dynamic Choice: The Necessity of Time Inconsistency." *American Economic Journal: Microeconomics,* 7, 150-178.

JULLIEN, D. (2016): "Under uncertainty, over time and regarding other people: Rationality in 3D." GREDEG working paper no. 2016-20.

KAHNEMAN, D. (2003): "Maps of bounded rationality: Psychology for behavioral economics." *American Economic Review*, *93*(5), 1449-1475.

KAHNEMAN, D. (2011). *Thinking, Fast and Slow*. New York. Farrar, Straus, and Giroux.

KAHNEMAN, D. AND S. FREDERICK (2002): "Representativeness revisited: Attribute substitution in intuitive judgment." In T. Gilovich, D. Griffin, D. Kahneman, Heuristics and biases: The psychology of intuitive judgment, 49-81. Cambridge.

KAHNEMAN, D. AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47, 263-292.

KEENEY, R.L. AND R. NAU (2011): "A theorem for Bayesian group decisions." *Journal of risk and Uncertainty,* 43, 1-17.

KEENEY, R. L. AND H. RAIFFA (1993): *Decisions with multiple objectives: Preferences and value trade-offs*. New York. Cambridge.

KEREN, G. AND P. ROELOFSMA (1995): "Immediacy and Certainty in Intertemporal Choice." *Organizational Behavior and Human Decision Processes*, 63, 287-297.

KOOPMANS, T.C. (1960): "Stationary ordinal utility and impatience." *Econometrica*, 28, 287-309.

KOVARIK, J. (2009). "Giving it now or later: Altruism and discounting." *Economics Letters*, 102, 152-154.

KRAWCZYK, M. AND LE LEC, F (2010): "'Give me a chance!' An experiment in social decision under risk. *Experimental Economics*, 13, 500-511.

KREPS, D.M. AND PORTEUS, E.L. (1978): "Temporal resolution of uncertainty and dynamic choice theory." *Econometrica,* 46, 185-200.

LAIBSON, D. (1997): "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112, 443-477.

LEDER, J. AND T. BETSCH (2016): "Risky choice in interpersonal context: Do people dare because they care?" *Journal of Economic Psychology*, 52, 1-23.

LEIGH, J. (1986): "Accounting for tastes: Correlates of risk and time preferences." *Journal of Post Keynesian Economics*, 9, 17-31.

LELAND, J. (1994): "Generalized similarity judgments: An alternative explanation for choice anomalies." *Journal of Risk and Uncertainty*. 9, 151-172.

LELAND, J. (2002): "Similarity Judgments and Anomalies in Intertemporal Choice." *Economic Inquiry*, 40, 574-581.

LOEWENSTEIN, G. AND D. PRELEC (1992): "Anomalies in Intertemporal Choice: Evidence and an Interpretation." *The Quarterly Journal of Economics*, 107, 573-597.

LOEWENSTEIN, G., S. RICK AND J.D. COHEN (2008): "Neuroeconomics," *Annual Review of Psychology*, 59, 647-672.

LOEWENSTEIN, G., O'DONOGHUE, T. AND S. BHATIA (2015): Modeling the interplay between affect and deliberation. *Decision*, *2*, 55-81.

LUCKMAN, A., C. DONKIN, AND B.R. NEWELL (2017): "People wait longer when the alternative is risky. The relation between preferences in risky and inter-temporal choice." Forthcoming, *Journal of Behavioral Decision Making*,

MARKOWITZ (1952): "The Utility of Wealth," *Journal of Political Economy*, 60, 151-158

MAYER, N.D., AND Z.L. TORMALA (2010): ""Think" versus "feel" framing effects in persuasion." *Personality and Social Psychology Bulletin*, 36, 443-454.

MCCLURE, S.M., LAIBSON, D.I., LOEWENSTEIN, G., AND J.D. COHEN (2007): "Time discounting for primary rewards," *The Journal of Neuroscience*, 27, 5796-5804.

MIAO, B. AND S. ZHONG (2015): "Comment on "Risk Preferences Are Not Time Preferences": Separating Risk and Time Preference." *The American Economic Review*, 105, 2272-2286.

MONGIN, P. AND M. PIVATO (2015). "Ranking multidimensional alternatives and uncertain prospects." *Journal of Economic Theory*, 157, 146-171.

MUKHERJEE, K. (2010): "A Dual System Model of Preferences under Risk," *Psychological Review*. 117, 243-255.

NEYS W.D., N. NOVITSKIY, L. GEERAERTS, J. RAMAUTAR, AND J. WAGEMANS (2011): "Cognitive control and individual differences in economic ultimatum decision-making."*PLoS*

ONE, 6, 1-9.

O'DONOGHUE, T. AND M. RABIN (1999): "Doing it Now or Later." *The American Economic Review*, 89, 103-124.

OECHSSLER, J., ROIDER, A., AND P.W. SCHMITZ (2009): "Cognitive abilities and behavioral biases." *Journal of Economic Behavior and Organization,* 72, 147-152.

ONAY, S. AND A. ONCULER (2007): "Intertemporal Choice under Timing Risk: An Experimental Approach," *Journal of Risk and Uncertainty*, 34, 99 – 121.

PONTI, G. AND I. RODRIGUEZ-LARA (2015): "Social preferences and cognitive reflection: evidence from a dictator game experiment." *Frontiers in Behavioral Neuroscience*, 9, 1-8.

PRELEC, D. (1998): "The probability weighting function." *Econometrica*, 66, 497-527.

PRELEC, D. AND G. LOEWENSTEIN (1991): "Decision Making Over Time and Under Uncertainty: A Common Approach." *Management Science*, 37, 770-786.

QUIGGIN, J. (1982): A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3, 323-343.

RAND, D.G., PEYSAKHOVICH, A., KRAFT-TODD, G.T., NEWMAN, G.E., WURZBACHER, O., NOWAK, M.A., AND GREENE, J.D. (2014): "Social heuristics shape intuitive cooperation." *Nature Communications*, 5: 4677, 1-12.

RAND, D. (2016): "Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation." Forthcoming, *Psychological Science.*

REGENWETTER, M., DANA, J. AND C.P. DAVIS-STOBER (2011): Transitivity of preferences. *Psychological Review*, 118, 42-56.

ROTTENSTREICH, Y. AND C.K. HSEE (2001): Money, kisses, and electric shocks: On the affective psychology of risk. *Psychological science*, *12*(3), 185-190.

RUBINSTEIN, A. (1988): Similarity and decision-making under risk (Is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory*, *46*, 145-153.

RUBINSTEIN, A. (2003): "Economics and Psychology"? The Case of Hyperbolic Discounting. *International Economic Review*, *44*(4), pp.1207-1216.

RUBINSTEIN, A. (2007): "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal*. 117, 1243-1259.

RUBINSTEIN, A. (2013): "Response time and decision making: An experimental study." *Judgment and Decision Making*, 8, 540-551.

SAITO, K. (2011): “Strotz meets Allais: Diminishing impatience and the certainty effect: Comment.” *American Economic Review*, 101(5), 2271-2275.

SAITO, K. (2013): “Social preferences under risk: Equality of opportunity versus equality of outcome.” *American Economic Review*, 103, 3084-3101.

SCHOLTEN, M. AND D. READ (2014): “Prospect theory and the “forgotten” fourfold pattern of risk preferences. *Journal of Risk and Uncertainty*, 48, 67-83.

SCHULZ, J.F., U. FISCHBACHER, C. THONI AND V. UTIKAL (2014): “Affect and fairness: Dictator games under cognitive load.” *Journal of Economic Psychology*, 1, 77-87.

SHIV, B. AND A. FEDORIKHIN (1999): “Heart and mind in conflict: The interplay of affect and cognition in consumer decision making.” *Journal of Consumer Research*, 26, 278-292.

STANOVICH, K.E., AND R.F. WEST (2000): “Individual differences in reasoning: Implications for the rationality debate?” *Behavioral and Brain Sciences*, 23, 645-665.

STARMER, C., 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of economic literature*, *38*, 332-382.

TRAEGER, C.P. (2013): “Once Upon a Time Preference – How Rationality and Risk Aversion Change the Rationale for Discounting.” CES Working Paper No. 3793.

TSUKAYAMA, E. AND A.L. DUCKWORTH, (2010): Domain-specific temporal discounting and temptation. *Judgment and Decision Making*, *5*, 72 - 82.

TVERSKY, A. AND D. KAHNEMAN (1992): “Advances in Prospect Theory: Cumulative representation of uncertainty.” *Journal of Risk and Uncertainty*, 5, 297-323.

VAN DEN BOS, W. AND S.M. MCCLURE (2013): “Towards a general model of temporal discounting.” *Journal of the Experimental Analysis of Behavior*, 99, 58-74.

WAKKER, P.P. (2010): “*Prospect Theory: For Risk and Ambiguity*.” Cambridge.

WALTHER, H. (2010):"Anomalies in intertemporal choice, time-dependent uncertainty and expected utility–A common approach.” *Journal of Economic Psychology*, 31, 114-130.

WU, G. AND R. GONZALEZ (1996): “Curvature of the probability weighting function.” *Management Science*, 42, 1676-1690.

ZUBER, S. (2016): “Harsanyi’s theorem without the sure-thing principle: On the consistent aggregation of monotonic Bernoullian and Archmiedean preferences.” *Journal of Mathematical Economics*, 63, 78-83.