11-2-2016

Economic Analysis with Systematically Biased Agents

Mark Schneider
Chapman University

Economic Analysis with Systematically Biased Agents

Comments
Working Paper 16-28

Economic Analysis with Security Biased Agents

BY MARK SCHNEIDER*

This Version: January 25, 2018

ABSTRACT

A tenet of behavioral economics is that biases are systematic and should have visible effects in economic applications. Expected utility maximization has been widely applied in economics, but progress has been slower incorporating 'systematically biased' agents into applications involving risk. This contrasts with the widespread application of present-biased preferences in intertemporal settings. To address this gap, we advocate a model of quasi-rank dependent probability weighting as a natural risky choice analog to quasi-hyperbolic time discounting. The model provides a formulation of 'security bias' – a disproportionate preference for lotteries with larger minimum payoffs, which unifies the certainty effect and loss aversion. The model satisfies stochastic dominance and transitivity and transforms individual rather than cumulative probabilities. We illustrate the model's tractability, demonstrating that it predicts the optimal purchase of insurance at actuarially unfair prices, the existence of gaps between buying and selling prices, and potential market failure due to security bias. The model also demonstrates that markets can generate unbiased prices even if all traders are systematically biased. A generalization of the model produces an asset pricing formula in which an asset's price depends on its fundamental value, a risk premium, a positive skewness premium, and a premium for robustness to model uncertainty. We also extend the model to time preferences, resulting in a 'three-factor' model of behavioral biases with factors reflecting (positive) skewness preference, present bias, and security bias. We demonstrate that this three-factor model predicts eight prominent behavioral anomalies in the literature for decisions under risk and over time.

Keywords: Risk Preference; Time Preference; Present Bias; Bias toward Concentration; Skewness Preference

JEL Codes: D03; D81

* Economic Science Institute, Chapman University. One University Drive, Orange, CA 92866. (e-mail: MarkSchneider@aya.yale.edu). Acknowledgments: I thank Manel Baucells, Enrico Diecidue and Charlie Sprenger for comments and encouragement regarding this research. I am solely responsible for any errors.

I. INTRODUCTION

The expected utility (EU) model has been the workhorse of economic analysis since it was axiomatized by John von Neumann and Oscar Morgenstern nearly seventy years ago[1]. It was swiftly applied to decisions involving insurance (Friedman and Savage, 1948), the existence of mixed strategy equilibria in non-cooperative games (Nash 1950a, 1951), cooperative bargaining theory (Nash, 1950b) and optimal portfolio selection (Tobin 1958; Markowitz, 1959). With the rise of information economics, EU became the micro foundation for analyzing games and markets under asymmetric information. It has penetrated the barriers into neighboring disciplines, finding application in political science, evolutionary biology, and sociology. Even one of its greatest critics has referred to EU as "the most important theory in the social sciences" (Kahneman 2011, p. 270).

Almost since its adoption, however, EU has been subject to persistent challenges, most notably due to Allais (1953). Allais introduced two paradoxes – the common consequence effect and the common ratio effect in which many people frequently reveal a bias toward certainty that violates the independence axiom - one of the central assumptions of EU. Motivated by the premise that more accurate assumptions would lead to more accurate predictions, or perhaps that EU imposed too stringent a definition of rationality, research in the late 1970's began to experiment with weakening the axioms of EU, or modifying its basic functional form to better accommodate empirical evidence such as the Allais paradoxes.

One of the earliest approaches to generalizing EU involved transforming individual probabilities in accordance with some probability weighting function to give the model greater flexibility (Edwards 1954; Karmarkar 1978; Kahneman and Tversky, 1979). This approach was abandoned however, when it was recognized that transforming individual probabilities seems to lead inevitably to violations of first-order stochastic dominance, a fundamental principle of rational choice which is rarely violated in experiments (Diecidue and Wakker, 2001).

Quiggin (1982) provided an elegant solution to the problem of transforming probabilities while preserving stochastic dominance. His rank dependent utility (RDU) model made the basic observation that to satisfy stochastic dominance, a probability weighting model should account for the rank of payoffs in a lottery and transform *cumulative* rather than individual probabilities. The probability transformations in RDU were subsequently incorporated into cumulative prospect

[1] See Bleichrodt et al. (2016) for the conclusion that John Nash and Jacob Marschak were the first to provide a complete axiomatization of EU in their 1950 papers in the same issue of Econometrica (Marschak 1950; Nash 1950b).

theory (CPT) by Tversky and Kahneman (1992) which is widely viewed today as the leading descriptive model of choices under risk (Fox et al., 2015).

Without delving into a debate regarding the strengths or limitations of CPT in describing human behavior, we make a different point – that CPT has generally lagged far behind EU in terms of economic applications. Reflecting on thirty years of prospect theory, Barberis (2013, p.173) comments, “It is curious, then, that so many years after the publication of the 1979 paper, there are relatively few well-known and broadly accepted applications of prospect theory in economics.”

This observation is in contrast to the domain of choices over time where the model of quasi-hyperbolic discounting has already been ‘plugged in’ to a wide variety of applications ranging from consumption and saving decisions (e.g., Angeletos et al., 2001; Diamond and Koszegi, 2003), problems of self-control (Laibson et al., 1998), addiction (Khwaja et al., 2007), credit card usage (e.g., Laibson, 1997; Meier and Sprenger, 2010) and bilateral bargaining (Kodritsch, 2014), even though it was popularized in a paper published nearly 20 years after prospect theory (Laibson, 1997). With regard to the model of quasi-hyperbolic discounting, Prelec (2006, p. 334) comments: “It is arguably the single most productive theoretical innovation associated with behavioral economics, overshadowing prospect theory in the breadth of applications.”

One might be led to wonder why prospect theory has seen relatively limited application, despite being described in survey articles as “the most successful general-purpose model currently available for predicting, describing, and interpreting decisions under risk” (Fox et al., 2015).

We suggest that a potential answer to this question lies in the observation made by Prelec. Both CPT and the quasi-hyperbolic model of time preference formalize systematic deviations from the standard models of rational choice. CPT captures a variety of different biases (e.g., the Allais paradoxes, loss aversion, risk aversion for gains and risk-seeking for losses) as arising from different components of the model (probability weighting, reference dependence, diminishing sensitivity). As the value function is defined over gains and losses rather than final wealth, and as multiple components of the model generate different effects, CPT does not always lend itself to produce clean comparisons with EU. In contrast, the quasi-hyperbolic model focuses on one important deviation from the standard discounted utility (DU) theory (present bias), and captures this bias through a single parameter which facilitates very clean comparisons with DU. Rather than taking a comprehensive approach, the quasi-hyperbolic model has its own descriptive limitations- its strength lies in the simplicity in which it captures an important behavior. The generality and comprehensiveness of CPT is impressive, but it is difficult to plug in CPT to game theoretic settings, or markets, or to represent agents in an economy and derive closed solutions the way that this can be done in the simple quasi-hyperbolic model. Indeed, Machina (1989) writes that the theoretical goal of descriptive decision models "is to show that non-expected utility models of individual decision making can be used to conduct analyses of standard economic decisions under uncertainty, such as insurance, gambling, investment, or search, in a manner that at least approximates the elegance and power of expected utility analysis" (p. 1623). For this purpose, it may be desirable to have a model of risky choice that is analogous to the quasi-hyperbolic model.

While the quasi-hyperbolic model is a standard generalization of DU which incorporates systematically biased agents, there is no standard recipe to incorporate systematically biased agents who deviate from EU into economic applications involving risk. Given the preceding discussion, one might raise the question of whether there is a natural analog to quasi-hyperbolic discounting for decisions under risk which generalizes EU to capture systematic biases, but which is also portable to a variety of economic applications. We argue that the answer to this question appears to be 'yes' and the resulting model appears to be one of 'quasi-rank dependent' utility (QRD). The QRD model formalizes 'security bias' – a disproportionate preference for lotteries with higher minimum payoffs. The QRD model satisfies basic axioms of rational choice such as transitivity and stochastic dominance, and transforms individual rather than cumulative probabilities. This observation may be surprising since it is widely believed that models which transform individual probabilities necessarily violate stochastic dominance or transitivity (Diecidue et al., 2004; Dhami, 2016). Yet QRD retains each of these properties. In contrast, models that overweight outcomes which occur with certainty do violate transitivity or stochastic dominance (Neilson, 1992; Schmidt 1998; Bleichrodt and Schmidt 2002; Diecidue et al., 2004; Andreoni and Sprenger, 2010). The QRD model is the special case of the non-extreme outcome additive (NEO-additive) model (Schmidt 2000; Chateauneuf et al., 2007; Webb and Zank 2011) which preserves risk aversion.

After outlining the quasi-rank dependent (QRD) model in Section II, we consider behavioral implications (Section III) and apply QRD to optimal insurance decisions (Section IV), prediction markets (Section V), and auctions (Section VI). Section VII extends QRD to account for optimism and pessimism. Section VIII applies the model to finance. Section IX extends QRD to both risk and time preferences. Section X discusses the related literature. Section XI concludes. Proofs for Sections III and IX are provided in the Appendix. Other proofs are in the main text.

II. QUASI-RANK DEPENDENT UTILITY THEORY

Let $X \subset \mathbb{R}$ denote a finite set of possible outcomes. A *lottery*, $f,$ is a probability distribution on X. Denote the set of lotteries by $\Delta(X)$. Consider model (1) where $\underline{x}(f)$ is the minimum outcome in the support of f, $U(f) = \sum_{x \in X} f(x) \cdot u(x)$ and $\theta \in [0,1]$:

$$V(f) = \theta[U(f)] + (1 - \theta)[u(\underline{x}(f))], \tag{1}$$

Model (1) takes the convex combination of the expected utility and the minimum utility of the lottery and can be interpreted as a disproportionate preference for lotteries with larger minimum payoffs. We refer to this behavior as 'security bias' since lotteries with higher minimum payoffs offer greater 'security' to the decision maker by limiting the worst-case scenario. Security bias can be quantified by $1 - \theta$. We show that security bias offers a unified explanation for the certainty effect (the Allais paradox and common ratio effect) and loss aversion (aversion to symmetric small mixed gambles) for choices under risk (in Section III), and that security bias generates the sign effect in intertemporal choice (in Section IX). If probabilities are subjective, security bias can also reflect a preference for robustness to mis-specified beliefs, and indeed, it reduces to Wald's (1950) maximin rule when $\theta = 0$, which is widely used in robust decision making under uncertainty.

Model (1) first appeared explicitly in Schmidt (2000), although a variant of (1) was introduced in Gilboa (1988). An analogous model to (1) for decisions under ambiguity is characterized in Kopylov (2009). Model (1) also appears as a special case of the model in Webb and Zank (2011) which allows for both optimistic and pessimistic behavior. However, despite the simplicity of (1) and its convenient properties, it has received almost no attention in applications.

There are several related behavioral interpretations for (1). The minimum payoff of a lottery may be salient in the mind of the decision maker and $1 - \theta$ then reflects disproportionate attention allocated to that payoff, where attention is a limited resource. A decision maker might also overweight the worst outcome in a lottery's support because it carries valuable information that is not shared by the other outcomes – namely it informs the decision maker prior to his decision what is the most he can be guaranteed from that lottery with certainty. Alternatively, a pessimistic decision maker who feels 'unlucky' or who 'expects the worse' may over-weight the worst outcome in a lottery's support. An axiomatic foundation for (1) is given by Webb and Zank (2011). Instead, we will focus on behavioral implications, applications, and extensions of the model.

To begin, we note that (1) has an equivalent representation as a model of quasi-rank dependent utility. Let X_f denote a random variable associated with f. Let $\succsim$ denote a binary (preference) relation on $\Delta(X)$, with strict preference and indifference represented by $\succ$ and $\sim$, respectively.

Definition 1: (Quasi-Rank Dependent Utility): *Under Quasi-Rank Dependent (QRD) utility theory, there exists utility function, u, probability weighting function π, with $\pi(0)=0, \pi(1)=1$, and $\sum_{x\in X}\pi(f(x))=1$, and a unique parameter $\theta \in [0,1]$ such that for any $f, g \in \Delta(X)$, $f \succsim g$ if and only if $V(f) \geq V(g)$, where for all x in the support of f:*

$$(2) \qquad V(f)=\sum_{x\in X}\pi(f(x))\cdot u(x),\ \pi(f(x))=\begin{cases}1-\theta+\theta f(x), & P(X_f \geq x)=1\\ \theta f(x), & P(X_f \geq x)<1\end{cases}$$

Note that the weights sum to 1. Somewhat surprisingly, the simple probability weighting function in (2) has not yet appeared in the literature. The formula in Definition 1 is a tractable "quasi-rank-dependent" probability weighting model which satisfies both stochastic dominance and transitivity. In addition, it is straightforward to test expected utility theory in this setup by testing if $\theta=1$. The parameter $\theta \in [0,1]$ indexes a decision maker's degree of security bias – the extent to which the agent deviates from expected utility maximization in the direction consistent with the certainty effect (Allais, 1953) and loss aversion (Kahneman and Tversky, 1979) with greater bias for lower values of θ. A rational unbiased agent corresponds to $\theta=1$. This approach transforms *individual* probabilities (rather than cumulative probabilities). It is 'quasi-rank-dependent' since the weight is different for the lowest ranked outcome, but all other outcomes receive the same weight. For further intuition, consider an analogy to choice over time. Note that (2) is somewhat analogous to the model of quasi-hyperbolic discounting (3) which has discount function $d(t)$, and consumption stream $(x_0, x_1, \ldots, x_T)$ is evaluated as (3) where $\beta \in [0,1]$:

$$(3) \qquad W(x_0, x_1, \ldots, x_T)=\sum_t d(t)\cdot u(x_t)\ \ and\ (t)=\begin{cases}1, & t=0\\ \beta\delta^t, & t>0\end{cases}.$$

A. Certainty Preference or Security Bias?

Previous work has led to the impression that the analog to quasi-hyperbolic discounting under risk is a model of 'certainty preference' with weights assigned to probabilities depending on whether $f(x)=1$ or $f(x)<1$. However, such models necessarily violate either stochastic dominance or transitivity (e.g., Diecidue et al., 2004). Yet because such models capture the certainty effect, which is viewed as a risky choice analog to present bias (Prelec and Loewenstein, 1991; Halevy 2008), it seems unavoidable that such models are the appropriate analog to quasi-hyperbolic discounting. This conclusion is disappointing since it is generally agreed that any normative model should satisfy both stochastic dominance and transitivity, and those axioms are rarely violated in experiments (e.g., Blavatskyy (2010), Regenwetter et al. (2011) and Baillon et al. (2014)).

Upon closer inspection, however, it seems that such models do not capture certainty preference in an intuitive way: They allow for situations, for example, where a person chooses a guaranteed \$10 over a 50-50 chance of gaining \$11 or gaining \$12, merely because the former is 'certain'. It seems plausible, at least in such cases, that certainty is better viewed as the 'minimum guarantee' of a lottery. Under this view, certainty preference would predict choosing the non-degenerate lottery since the former guarantees \$10 whereas the latter *guarantees at least* \$11. This observation suggests that a more appropriate analog assigns different weights to the probability of outcome x depending on if $P(X_f \geq x) = 1$ or $P(X_f \geq x) < 1$.

It also appears that models which assign different weights to probabilities depending on whether $f(x) = 1$ or $f(x) < 1$ do not capture the behavior of most subjects who exhibit the Allais paradox. For instance, Incekara-Hafalir and Stecher (2016) conducted a novel test of Allais-style violations of EU in which they replaced the common consequence across a series of six decisions. Using a Savage matrix representation of lotteries, their subjects chose between safe lottery $(\$c, 0.89; \$8, 0.10; \$8, 0.01)$ and risky lottery $(\$c, 0.89; \$10, 0.10; \$0, 0.01)$ for $c \in \{0,5,8,10,16,20\}$. Let R denote a 'risky' choice and S denote a 'safe' choice. Ordering the six choices from those with the lowest value of c to the highest value of c, only the preference patterns RRRRRR and SSSSSS are consistent with EU. The classic Allais paradox corresponds to the case where the first letter in the sequence is R (where $c = 0$) and the third letter is S (where $c = 8$). Models which assign different weights to certain and uncertain outcomes predict a 'certainty effect' pattern of RRSRRR but not the 'zero effect' pattern RSSSSS. In contrast, QRD is consistent with a 'zero effect' pattern RSSSSS, but not with a certainty effect pattern. The standard RDU model permits the zero effect pattern, the certainty effect pattern, and the reverse of each pattern, and so is not very helpful in predicting which effect will dominate. In fact, Incekara-Hafalir and Stecher (2016) observed strong support for the zero effect pattern while none of the subjects in their experiment exhibited the certainty effect pattern[2]. Both of these findings are consistent with security bias as formalized by QRD. It thus appears that (2) may actually be the more appropriate

[2] Incekara-Hafalir and Stecher also find that the effect of security level is strongest when the minimum payoff is zero.

analog to (3). Such a conclusion is encouraging because (2) satisfies both stochastic dominance and transitivity. Furthermore, as we will illustrate in the following sections, (2) can be plugged into economic models involving risk as simply as quasi-hyperbolic discounting can be plugged into economic models involving time.

III. BEHAVIORAL IMPLICATIONS

We next explore four behavioral implications of QRD: continuity violations, loss aversion, the Allais paradox and the common ratio effect.

A. Violations of Continuity

The continuity axiom of EU implies that for any lotteries p, p', p'' such that $p \gtrsim p' \gtrsim p''$, there exists probability $\alpha \in [0,1]$ such that $\alpha p + (1-\alpha)p'' \sim p'$. Although continuity is a standard assumption, there are intuitive cases where it can be violated. Adapting an example from Levin (2006), suppose $p \coloneqq (w + \$10,\ 1)$, (i.e., p yields your current wealth w plus $10 with certainty), $p' \coloneqq (w, 1)$ (i.e., the status quo), and $p'' \coloneqq (0,1)$ (you lose all your wealth with certainty). Would most people really be willing to risk losing all their wealth for a chance of gaining $10? Does rationality really require such behavior? We show that QRD can explain violations of continuity in such cases. In this example, under QRD, continuity requires there to be some $\alpha \in [0,1]$ so that

$$V(\alpha p + (1-\alpha)p'') = \theta\alpha\big(u(w+10) - u(0)\big) + u(0) = u(w).$$

Clearly this equation need not hold when $\theta < 1$. For instance, we can normalize utilities so that $u(w+10) = 1$ and $u(0) = 0$. If $u(w) > 0.9$, then any $\theta \leq 0.9$, implies $p' \succ \alpha p + (1-\alpha)p''$ for all $\alpha \in [0,1]$ and the agent will not risk all his wealth for any $\alpha \in [0,1]$. Under QRD, continuity is satisfied for lotteries with the same worst-case scenario which may be a more natural condition.

B. Loss Aversion

A fundamental property of observed choices under risk is loss aversion which has been defined behaviorally by Kahneman and Tversky (1979) and Schmidt and Zank (2005) as aversion to 50-50 gain-loss bets. More formally, given a choice between lotteries f and g where $f \coloneqq (y, 0.5; -y, 0.5)$ and $g \coloneqq (x, 0.5; -x, 0.5)$, for any $x > y \geq 0$, *loss aversion holds* if $f \succ g$.

Proposition 1: *Let $u(-x) = -u(x)$ for all x. Then loss aversion holds if and only if $\theta \in [0,1)$.*

C. Rabin's Paradox

Rabin (2000) introduced a powerful critique of EU. Rabin proved a calibration theorem which implies, for instance, that an EU maximizer who rejects a 50-50 bet to lose $100 or win $125 at all wealth levels will also reject a 50-50 bet to lose \$600 or win \$1 million. Rabin's theorem is often taken as strong evidence against the assumption of a utility function over final wealth levels

Under the QRD model, the 50-50 lose \$100, gain \$125 bet is rejected at all wealth levels, $w >$ \$100, if $u(w) > (1-\theta/2)u(w-100) + (\theta/2)u(w+125)$ for all $w > \$100$. Under the assumption that u is strictly concave, it follows that:

$$u((1-\theta/2)(w-100) + (\theta/2)(w+125)) > (1-\theta/2)u(w-100) + (\theta/2)u(w+125)$$

for all $\theta \in (0,1)$. Also, note that for any $\theta \in (0, 8/9)$, we have:

$$u(w) > u((1-\theta/2)(w-100) + (\theta/2)(w+125)).$$

Thus, for any strictly concave utility function, a QRD agent with $\theta \in (0, 8/9)$ will always reject Rabin's small stakes gamble. Next consider the 50-50 gamble to lose \$600 or gain \$$z$. The gamble is accepted at current wealth level, w, if $(\theta/2)u(w+z) + (1-\theta/2)u(w-600) > u(w)$.

In order for the gamble to be accepted, it must also be the case that:

$$u(w + z(\theta/2) - 600(1-\theta/2)) > u(w),$$

which implies $\theta > 1200/(600+z)$. Under the preceding restrictions on θ, a QRD agent may reject the small stakes gamble at all wealth levels and choose the gamble with a 50-50 chance of losing \$600 or gaining \$$z$ for sufficiently large z. As concrete examples, fixing $\theta = 0.7$, if $u(x) = x$, the first gamble is always rejected and the second is accepted for all $z \geq \$1115$; if $u(x) = \ln(x)$, the first gamble is always rejected and the second is accepted. for all $z \geq \$1218$.

D. The Common Ratio Effect

In the following analysis, we provide general conditions on θ which explain two systematic deviations from EU – the common ratio effect and the Allais paradox. In both cases, we assume that the decision maker has QRD preferences.

A robust violation of EU, the common ratio effect (at certainty), is defined next.

Definition 2: (Common Ratio Effect): Let $f \coloneqq (y,1), f' \coloneqq (y,q;0,1-q)$, $g \coloneqq (x,p;0,1-p)$, $g' \coloneqq (x,qp;0,1-qp)$, for any $\in (0,x)$ *and* $p,q \in (0,1)$. The *common ratio effect holds* if $f \sim g$ implies $f' \prec g'$.

Proposition 2: *The common ratio effect holds if and only if* $\theta \in (0,1)$

In the classic version of the common ratio effect due to Kahneman and Tversky (1979), $(x,y,p,q) = (\$4000, \$3000, 0.8, 0.25)$. Definition 2 implies that an agent who is indifferent between lotteries f and g will strictly prefer g' over f'.

E. The Allais Paradox

The Allais paradox (also known as the common consequence effect) is defined as:

Definition 3: (Allais Paradox): Let $f \coloneqq (y, 1), f' \coloneqq (y, q; 0, 1-q)$, $g \coloneqq (x, p; y, 1-q; 0, q-p), g' \coloneqq (x, p, 0; 1-p)$, for any $y \in (0, x)$ and $p, q \in (0,1)$. The *Allais paradox holds* if $f \sim g$ implies $f' \prec g'$.

Proposition 3: *The Allais paradox holds if and only if* $\theta \in (0,1)$.

In the classic Allais paradox $(x, y, p, q) = (\$5 \text{ Million}, \$1 \text{ Million}, 0.10, 0.11)$. Definition 3 implies that an agent who is indifferent between lotteries f and g will strictly prefer g' over f'.

F. Stake Dependence of the Allais Paradox

The QRD model also makes predictions regarding the Allais paradox which distinguish it from the most widely used specification of CPT. While the Allais paradox is observed at the large stakes of Allais (1953) where $(x, y, p, q) = (\$5 \text{ Million}, \$1 \text{ Million}, 0.10, 0.11)$, and at the stakes used by Kahneman and Tversky (1979), where $(x, y, p, q) = (\$2500, \$2400, 0.33, 0.34)$, the paradox is greatly diminished at small stakes. In particular, when payoffs are scaled down to $(x, y, p, q) = (\$100, \$20, 0.10, 0.11)$, as done by Fan (2002), or to $(x, y, p, q) = (\$25, \$5, 0.10, 0.11)$, as done by Huck and Muller (2012), experimental subjects typically choose the riskier lottery in both choices. There is a strong intuitive basis for not observing the paradox at these small stakes: People are naturally willing to accept the 1% chance of receiving \$0 in exchange for a 10% chance of receiving \$100. The appeal of the sure-thing is not as strong at small stakes. It is only when the large payoff is not much larger than the sure-thing (as in Kahneman and Tversky's example), or when the sure-thing is a very large amount (as in the Allais example), that the paradox is likely to be observed. Thus, a more complete explanation of the Allais paradox should predict the paradox to occur at the large stakes observed by Allais, and Kahneman and Tversky, but should predict behavior consistent with EU at the smaller stakes used by Fan (2002) and Huck and Muller (2012). The standard version of CPT with a power value function defined over gains and losses cannot account for this aggregate pattern even given any rank-dependent probability weighting function. However, QRD naturally accommodates all four cases. For example, if $u(w + x) = \ln(w + x)$, for an agent with current wealth w, and $\theta = 0.9$, then for any w such that $\$5{,}000 \leq w \leq \$500{,}000$, QRD predicts the Allais paradox to be observed for the examples by Allais (1953) and Kahneman and Tversky (1979), and predicts behavior consistent with EU (the choice of the two riskier options) for the examples from Fan (2002) and Huck and Muller (2012).

IV. APPLICATION TO RISK REDUCTION AND INSURANCE PURCHASE

For decisions under risk, an important implication of quasi-rank dependent probability weighting is that a decision maker will pay a disproportionately higher premium for risk elimination than for an equivalent degree of risk reduction that does not eliminate the risk. This prediction also has empirical support. For instance, Botzen et al. (2013) find that households place a substantial premium on policies to eliminate flood risk relative to other opportunities which merely reduce the risk. A similar conclusion was reached by Viscusi et al. (2014) who found there to be a greater premium for policies that reduce cancer risks to zero relative to policies which reduce but do not eliminate the risk. We illustrate this preference for risk elimination in the context of insurance.

A. The Decision to Purchase Regular Insurance

Consider the following situation described by Blavatskyy (2014): A decision maker has a risk of losing D dollars with probability $q \in (0,1)$. The decision maker has the option of purchasing $x \in [0, D]$ units of insurance, where one unit of insurance costs c dollars, with $c \in (0,1)$, and pays the decision maker one dollar if the loss occurs. We consider the optimality of purchasing regular (full) insurance (i.e., the case where $x = D$) under the quasi-rank dependent model.

As noted by Blavatskyy, under regular insurance, the decision maker loses exactly cD dollars regardless of whether the loss occurs. If $x < D$, the decision maker loses cx dollars with probability $1 - q$ and loses $D + x(c - 1)$ dollars with probability q. Under the quasi-rank dependent model, the decision maker will purchase regular insurance if and only if the following inequality holds for any $x \in [0, D)$.

$$u(-cD) > \theta(1-q)u(-cx) + (1-\theta+\theta q)u\big(-D - x(c-1)\big).$$

The preceding inequality can be arranged as follows:

$$\frac{u(-cD)-u\big(-D-x(c-1)\big)}{(D-x)(1-c)}(1-c) > \frac{u(-cx)-u\big(-D-x(c-1)\big)}{D-x}(1-q)\theta.$$

The fraction on the left-hand side of the inequality is the slope of the utility function between points $-cD$ and $-D - x(c-1)$, and the fraction on the right-hand side is the slope of the utility function between points $-cx$ and $-D - x(c-1)$. For any strictly concave utility function u, the slope on the left hand side is always greater than the slope on the right hand of the inequality. Therefore, regular insurance will be optimal to purchase when $(1-c)/(1-q) \geq \theta$.

Under EU, we have $\theta = 1$, and regular insurance is optimal when $q \geq c$. This observation is the well-known result that a risk-averse expected utility maximizer will never purchase regular insurance at an actuarially unfair price. This implication cannot be reconciled under EU with the large premiums many people are willing to pay to eliminate risk. Rearranging the above expression, we have the following which holds for any concave utility function:

Proposition 4: *A risk-averse consumer with QRD preferences will purchase regular insurance at an actuarially unfair price $c > q$ if* (4) *holds:*

$$1 - (1 - q)\theta \geq c. \tag{4}$$

Note that there is always some $\theta^* < 1$ such that an agent will purchase regular insurance, even at an actuarially unfair price for any $\theta \leq \theta^*$.

The QRD model can also explain the finding from Sydnor (2010) that many people 'over-insure' for modest risks. Sydnor found that expected utility theory could not explain his observations from real insurance purchases that customers with a 4 percent probability of a loss were willing to pay \$95 to lower the deductible from \$1,000 to \$500. Let p denote the price of insurance with a \$1000 deductible, and let w denote the consumer's initial wealth. For simplicity and to isolate the role of security bias, let $u(x) = x$. Then a QRD consumer prefers to pay \$95 to lower the deductible from \$1000 to \$500 if the following inequality holds:

$$(1 - 0.96\theta)(w - p - 595) + 0.96(w - p - 95)\theta > (1 - 0.96\theta)(w - p - 1000) + 0.96(w - p)\theta.$$

This inequality holds under QRD for all $\theta < 0.84375$.

B. Aversion to Probabilistic Insurance

Expected utility theory also predicts that a risk-averse agent will prefer 'probabilistic insurance' to regular insurance. Under probabilistic insurance, the decision maker pays a fraction of the insurance premium up front. If the loss occurs, there is a probability that the decision maker pays the remainder of the premium to receive full coverage from the insurance firm, but otherwise the insurance firm refunds the partial premium already paid by the decision maker and does not cover the loss. Kahneman and Tversky (1979) demonstrate that many people will not purchase probabilistic insurance, even though expected utility theory (with a concave utility function) implies that probabilistic insurance will be preferred to regular insurance.

As in the example from Kahneman and Tversky (1979), suppose that at wealth w, a person is indifferent between paying premium y to insure against a probability q of losing x, where $x > y > 0$. Then a risk-averse expected utility maximizer would be willing to pay a smaller premium ry to reduce the probability of losing x from q to $(1-r)q$ for $r \in (0,1)$. Kahneman and Tverksy note that in expected utility theory, "if one is indifferent between $(w-x, q; w, 1-q)$ and $(w-y)$, then one should prefer probabilistic insurance $(w-x, (1-r)q; w-y, rq; w-ry, 1-q)$ over regular insurance $(w-y)$."

Without loss of generality, set $u(w-x) = 0$ and $u(w) = 1$. Under the quasi-rank dependent model, the indifference condition implies

$$u(w-y) = \theta(1-q).$$

Probabilistic insurance is preferred to regular insurance if

$$u(w-y) < \theta rqu(w-y) + \theta(1-q)u(w-ry).$$

For any $\theta > 0$, substituting the indifference condition, we have

$$(1-q) < (1-q)rq\theta + (1-q)u(w-ry).$$

For $\theta = 1$, the above inequality holds if and only if u is concave, implying a preference for probabilistic insurance over regular insurance under expected utility theory. In contrast, we see that regular insurance is preferred to probabilistic insurance if $1 - rq\theta > u(w-ry)$, which can hold for $\theta < 1$. As the examples in this section illustrate, the quasi-rank dependent model can be very tractable in applications and provides a clean comparison to the expected utility predictions.

V. APPLICATION TO PREDICTION MARKETS

We consider a prediction market based on Wolfers and Zitzewitz (2006). Assets are Arrow-Debreu securities which pay \$1 if a target event occurs and \$0 otherwise. There are m traders, $j = 1, \ldots, m$, of whom $n < m$ are buyers and $m - n$ are sellers. There is heterogeneity in beliefs, where trader j believes the target event will occur with probability q_j. Wealth, w, is assumed to be independent of beliefs. Beliefs are drawn from a distribution $F(q)$. Traders are price-takers and pursue trading strategies which maximize their preferences. We assume that preferences are given by the quasi-rank dependent utility model. Traders are risk-averse with log utility and have the same degree of bias given by θ. Wealth is only affected by the outcome of the prediction market so there are no hedging motives for trading the security. Let p denote the price of the security and let x_j denote the quantity of the security purchased by trader j, for $j = 1, \dots, n$. Let y_j denote the quantity of the security sold by trader j, for $j = n+1, \dots m$. Buyers of the security solve (5):

$$(5) \qquad \max_x V_j = \theta q_j \ln(w + x_j(1-p)) + (1 - \theta q_j)\ln(w - x_j p)$$

for $j = 1, \dots, n$. Sellers of the security solve a similar maximization problem:

$$(6) \qquad \max_y V_j = (1 - \theta + \theta q_j)\ln(w - y_j(1-p)) + \theta(1 - q_j)\ln(w + y_j p)$$

for $j = n+1, \dots, m$. The optimal quantity, x_j^*, demanded by buyer $j = 1, \dots, n$ and the optimal quantity, y_j^*, offered by seller $j = n+1, \dots, m$ are given by:

$$(7) \qquad x_j^* = w\frac{\theta q_j - p}{p(1-p)}, \qquad y_j^* = w\frac{\theta(1-q_j)-(1-p)}{p(1-p)}.$$

From the formulas for x_j^* and y_j^*, a buyer j's demand is positive if $\theta q_j - p > 0$ and seller j's supply is positive if $\theta(1-q_j) - (1-p) > 0$. In equilibrium, supply equals demand, which implies:

$$(8) \qquad \int_{-\infty}^{1+\frac{p}{\theta}-\frac{1}{\theta}} w\frac{\theta(1-q)-(1-p)}{p(1-p)} f(q)dq = \int_{\frac{p}{\theta}}^{\infty} w\frac{\theta q - p}{p(1-p)} f(q)dq$$

From (7), we see that security bias implies that buying-selling price gaps will exist in equilibrium. A similar finding was derived by Dow and Werlang (1992) in the context of portfolio choice[3]:

Proposition 5: *Buying-selling price gaps exist in equilibrium: For QRD preferences, a trader with belief q will buy if and only if $p < \theta q$ and will sell if and only if $p > \theta q + 1 - \theta$.*

Note that when $\theta = 1$ (i.e., under EU preferences), there are no buying-selling price gaps: A trader with belief q will buy the security when $p < q$ and will sell when $p > q$. Also note that the size of the buying-selling price gap is equal to the degree of security bias $(1-\theta)$. For $\theta < 1$, security bias results in an efficiency loss: A buyer with belief q, could earn a positive subjective expected payoff by trading for all $q \in (p, p/\theta)$ and a seller with belief q could earn a positive subjective expected payoff by trading for all $q \in \left(1 + \frac{p}{\theta} - \frac{1}{\theta}, p\right)$. Yet no trade takes place for agents with $q \in \left(1 + \frac{p}{\theta} - \frac{1}{\theta}, \frac{p}{\theta}\right)$. In an extreme case, if security bias is sufficiently strong, it can even lead to market failure. For instance, if $p = 0.5$ and $\theta = 0.5$, then no trade occurs for any beliefs $q \in (0,1)$, even though welfare enhancing trade could occur for all buyers with $q > p$ and all sellers with $q < p$. For a less extreme example, if $p = 0.5$ and $\theta = 0.9$, then no trade occurs for beliefs $q \in \left(\frac{4}{9}, \frac{5}{9}\right)$.

[3] Dow and Werlang (1992) demonstrated that buying-selling price gaps for assets in financial markets hold under the more general Choquet Expected Utility preferences. The model of Bordalo et al. (2012a) generates buying-selling price gaps in the context of salience-based consumer choice.

Next, we consider the impact of security bias on market prices. Let $\overline{q} = \int_{-\infty}^{\infty} qf(q)dq$ (i.e., the average belief across the entire population of traders).

Definition 4: *The equilibrium market price, p, is unbiased if $p = \overline{q}$.*

Gjerstad (2005) and Wolfers and Zitzewitz (2006) show that in the EU case (i.e., $\theta = 1$), market prices will be unbiased. A natural question is whether prices can be unbiased despite the systematic security bias of market participants (i.e., for $\theta < 1$). That is, do equilibrium prices necessarily depend on θ if traders have preferences that depend on θ? We offer a simple illustration.

Proposition 6: *Suppose buyers and sellers maximize* (5) *and* (6), *respectively, and that supply and demand are each positive. If beliefs are uniformly distributed on* [0,1], *then equilibrium market level prices will be unbiased* (*i.e., $p = \overline{q}$*) *for any systematic individual level bias $\theta \in (0.5,1]$.*

Proof: The restriction $\theta \in (0.5,1]$ is necessary for supply and demand to each be positive[4]. Given our assumption that wealth is independent of beliefs, if $f(q)$ is uniform on [0,1], and supply and demand are positive, then (8) becomes:

$$\frac{w}{p(1-p)}\int_0^{1+\frac{p}{\theta}-\frac{1}{\theta}}(\theta(1-q)-(1-p))dq = \frac{w}{p(1-p)}\int_{\frac{p}{\theta}}^{1}(\theta q - p)\,dq.$$

The equilibrium price is then the price p that solves $(p+\theta-1)^2 = (p-\theta)^2$, yielding:

$$p = \frac{1-2\theta}{2-4\theta} = \frac{1}{2} = \overline{q}. \blacksquare$$

Although simple, Proposition 6 is a surprising result – aggregating preferences of *systematically biased* agents produces *unbiased* market level prices!

A primary finding in behavioral economics is that biases are systematic. Thus, they will not cancel out as noise in ways that random errors might. In contrast, many economists argue that biases will be eliminated by the market. In his book, *Misbehaving: The Making of Behavioral Economics*, Thaler (2015) writes: "I call this argument the invisible handwave…The vague argument is that markets somehow discipline people who are misbehaving. Handwaving is a must because there is no logical way to arrive at a conclusion that markets transform people into rational agents." Interestingly, Proposition 6 considers the case where biases are systematic in the same direction (in a manner consistent with the Allais paradox, and loss aversion) and shows that even

[4] Given that beliefs are uniformly distributed over [0,1], the restriction that supply and demand are positive implies that $\left[1+\frac{p}{\theta}-\frac{1}{\theta},\frac{p}{\theta}\right] \subset [0,1]$ or, equivalently, that $1-\theta < p < \theta$.

if all agents in a market are systematically biased, equilibrium prices can accurately aggregate beliefs and produce the *same* prices *as-if* all agents maximized expected utility.

Proposition 6 does not mean that biases do not affect market prices in general. The result is restrictive, particularly given the uniform assumption for beliefs and the assumption that traders have the same risk aversion and bias (although one might view this as a representative agent).

A subtle point related to Proposition 6 is that traders may revise their prior beliefs given the information revealed by market prices to form posterior beliefs. We have implicitly assumed that traders have fixed beliefs (i.e., their prior and posterior probabilities are equal). We next provide a simple and plausible illustration that equilibrium prices can be the same for prior and posterior beliefs, even if traders update their prior beliefs based on the information they extract from observing market prices. To distinguish prior and posterior beliefs, we now denote the former by q and the latter by $q(p)$. Consider a simple and plausible belief-updating rule discussed in Manski (2006) in which a trader's posterior belief is determined by a weighted average of her prior belief and the observed market price. That is:

(9) $$q(p) \coloneqq \lambda q + (1-\lambda)p.$$

where $\lambda \in [0,1]$. Under this rule, prior beliefs are updated in the direction of the market price, with $(1-\lambda)$ determining the degree to which beliefs are revised, including as special cases no revision ($(1-\lambda)=0$) and full revision ($(1-\lambda)=1$). If prior beliefs are uniformly distributed over [0,1], then the distribution of posterior beliefs is uniform over the interval $[(1-\lambda)p,\ \lambda+(1-\lambda)p]$ which is a subset of [0,1]. This interval encompasses many possible supports for the uniform distribution of posterior beliefs including those not centered at 0.5. Performing the same analysis as in Proposition 6, assuming supply and demand are positive and setting supply equal to demand in equilibrium yields:

$$\int_{(1-\lambda)p}^{\frac{p(1-\theta(1-\lambda))-(1-\theta)}{\theta\lambda}} \frac{(\theta(1-\lambda q-(1-\lambda)p)-(1-p))}{\lambda}\,dq = \int_{\frac{p(1-\theta(1-\lambda))}{\theta\lambda}}^{\lambda+(1-\lambda)p} \frac{(\theta(\lambda q+(1-\lambda)p)-p)}{\lambda}\,dq.$$

The equilibrium price is then the price p that solves the following equation:

$$(p((\lambda^2-1)\theta+1)+(\theta-1))^2 = (p((\lambda^2-1)\theta+1)-\lambda^2\theta)^2.$$

Solving for p yields:

$$p = \frac{1-2\theta-\lambda^4\theta^2+\theta^2}{2(1-2\theta-\lambda^4\theta^2+\theta^2)} = \frac{1}{2} = \frac{2(1-\lambda)p+\lambda}{2} = \overline{q}. \blacksquare$$

Thus, even if agents use an updating rule that takes a weighted average of their prior and the market price to form their posterior beliefs, the equilibrium price is the same as in Proposition 6.

VI. APPLICATION TO EQUILIBRIUM BIDDING AND REVENUE EQUIVALENCE

The QRD model provides a simple tool for investigating whether and how economic conclusions depend on the expected utility assumption. To further illustrate, consider an application to auctions: There are n bidders in a first price sealed bid private value auction for a single deterministic object. Bidder valuations are independently drawn from a uniform distribution over [0,1]. Bidders are QRD agents with linear utility and bias parameter $\theta \in (0,1]$. Then there is a symmetric equilibrium in which a bidder with value v bids $b(v) = \left(\frac{n-1}{n}\right) v$. This bidding strategy is well known although it has only been proven for the case $\theta = 1$ (i.e., where all bidders satisfy EU). Here we observe that the result holds more generally for any $\theta \in (0,1]$. To see this, suppose bidder B_1 has value v_1 and bids b_1. Since B_1 always wins the auction if $b_1 = (n-1)/n$, we can restrict our analysis to bids over the interval $[0, (n-1)/n]$. We need to establish that the strategy, $b(v)$, above, is a best response for B_1 if it is adopted by all other bidders. Suppose another bidder B_2 with value v_2 bids $b_2(v_2) = [(n-1)/n]v_2 < b_1$. Since v_2 is uniformly distributed over [0,1], B_1 bids higher than B_2 with probability $nb_1/(n-1)$. To win, B_1 must outbid all other bidders. The probability B_1 wins is:

$$q(b_1) = b_1^{n-1} \left(\frac{n}{n-1}\right)^{n-1}.$$

A QRD agent values each bid according to:

$$V(b_1) = \theta q(b_1)(v_1 - b_1) + (1 - \theta q(b_1))(0).$$

When finding the bid that maximizes $V(b_1)$, θ cancels and we obtain the same bid function $b(v)$ noted above which is well known for the expected utility case. But now the result is more general, holding for any degree of bias $\theta \in (0,1]$. Also θ does not affect bidding behavior in a second price sealed bid private value auction for a deterministic object[5]. It is then immediate that revenue equivalence for the first and second price auctions, originally established for risk-neutral EU agents, continues to hold for QRD agents with linear utility. Note that this conclusion holds for any probability $q(b_1)$ and so is robust to assuming a uniform value distribution.

---

[5] Karni and Safra (1989) have demonstrated that when the object up for auction is a lottery, bidding one's value in a second price auction is only an equilibrium if the bidders are expected utility maximizers.

VII. EXTENSION TO OPTIMISM AND PESSIMISM

In the preceding analysis, we have considered the implications of a QRD agent in several different economic environments. One might also be interested in the behavior of an agent who exhibits both optimism and pessimism. The QRD model can be extended to the full NEO-EU model of Chateauneuf et al. (2007) by allowing for the possibility that an agent is sensitive to both the best and worst outcome of a lottery. For a lottery f, the NEO-EU model can be written as:

(10) $$V(f) = \theta U(f) + (1-\theta)[\alpha u(\overline{x}) + (1-\alpha)u(\underline{x})]$$

where $U(f)$ is the expected utility of f, $\alpha \in [0,1]$ represents the agent's degree of optimism, and $\overline{x}$ and $\underline{x}$ are, respectively, the most and least preferred outcomes in the support of f. Let $a \coloneqq (1-\theta)\alpha$ and let $b \coloneqq (1-\theta)(1-\alpha)$. The equivalent QRD representation to (10) is (11):

(11) $$V(f) = \sum_{x\in X} \pi\big(f(x)\big) \cdot u(x), \text{ where } \pi\big(f(x)\big) = \begin{cases} a + \theta f(x), & x = \overline{x} \\ \theta f(x), & x \in (\underline{x}, \overline{x}) \\ b + \theta f(x), & x = \underline{x} \end{cases}$$

Note that (11) is 'quasi-rank dependent' in that it transforms individual rather than cumulative probabilities and rank matters only for the best and worst outcomes. The model in (11) also satisfies transitivity and stochastic dominance.

The results from Section V on prediction markets extend to the general QRD preferences in (10). Buyers of the security now solve the following problem:

(12) $$\max_x V_j = ((1-\theta)\alpha + \theta q_j)\ln(w + x_j(1-p)) + ((1-\theta)(1-\alpha) + \theta(1-q_j))\ln(w - x_j p)$$

for $j = 1, \dots, n$. Sellers of the security solve a similar maximization problem:

(13) $$\max_y V_j = ((1-\theta)(1-\alpha) + \theta q_j)\ln(w - y_j(1-p)) + ((1-\theta)\alpha + \theta(1-q_j))\ln(w + y_j p)$$

for $j = n+1, \dots m$. The optimal quantity, x_j^*, demanded by buyer $j = 1, \dots n$ and the optimal quantity, y_j^*, offered by seller $j = n+1, \dots . m$ are now given by:

$$x_j^* = w\frac{\theta q_j - p + \alpha(1-\theta)}{p(1-p)}, \qquad y_j^* = w\frac{\theta(1-q_j) - (1-p) + \alpha(1-\theta)}{p(1-p)}.$$

In equilibrium, supply equals demand, which implies

(14) $$\int_{-\infty}^{1+\frac{p}{\theta}-\frac{1}{\theta}+\frac{\alpha}{\theta}-\alpha} w\frac{\theta(1-q)-(1-p)+\alpha(1-\theta)}{p(1-p)} f(q)dq = \int_{\frac{p-(\alpha(1-\theta))}{\theta}}^{\infty} w\frac{\theta q - p + \alpha(1-\theta)}{p(1-p)} f(q)dq$$

Next, we have the following generalization of Proposition 6:

Proposition 7: *Suppose buyers and sellers maximize* (12) *and* (13), *respectively, and that supply and demand are each positive. If beliefs are uniform on* [0,1], *then equilibrium market prices will be unbiased* (*i.e.*, $p = \overline{q}$) *for any* $\theta \in (0.5,1]$, *and any degree of optimism or pessimism,* $\alpha \in [0,1]$.

Proof: Proceeding as in Section V, computing the optimal solutions to (12) and (13), and setting demand equal to supply in equilibrium yields (14) for equilibrium prices. If beliefs are uniform on [0,1], then solving (14) for p yields:

$$p = \frac{1 - 2\alpha - 2\theta + 2\alpha\theta}{2(1 - 2\alpha - 2\theta + 2\alpha\theta)} = 0.5 = \overline{q}. \blacksquare$$

Thus, even if we allow for *any* $\alpha \in [0,1]$, distorting probabilities in the direction of optimism or pessimism does not affect market prices when the distribution of beliefs is uniform.

VIII. APPLICATION TO FINANCE

We next apply the model to the domain of finance. Chateauneuf et al. (2007) considered a stylized setting with one risky asset and one risk-free asset and demonstrated that model (10) could help explain the equity premium puzzle. Here we generalize their approach to an arbitrary number of risky assets and consider the implications of the model for some classical asset pricing anomalies.

We consider a representative agent economy with a representative investor who has initial wealth w_0 and can invest in J > 1 risky assets and a riskless asset (bond). Asset j trades at price p_j and pays dividend x_{js} if state s occurs. There are S possible states of nature. The bond has a normalized price of 1 and pays return r in every state. The investor has subjective prior π over states in which state $s \in \{1, \dots, S\}$ occurs with probability $\pi_s > 0$. The investor has preferences from (10, 11) with temporal discount factor δ and chooses a portfolio (holdings of assets) to maximize (15):

$$(15) \qquad \max_{\{\lambda_j\}} u(c_0) + \delta\left(\mathbb{E}_\pi[u(c_s)]\theta + \left[\alpha \max_{s\in S} u(c_s) + (1-\alpha) \min_{s\in S} u(c_s)\right](1-\theta)\right)$$

subject to the constraints: $c_0 = w_0 - \sum_j \lambda_j p_j$, $c_s = w_1 + \sum_j \lambda_j x_{js}$, $\lambda_j \geq 0$, where $\lambda_j, j = 1, \dots, J$ is the investor's holding of asset j (short sales are not permitted), and w_1 is a deterministic component of the investor's wealth in Period 1. Substituting the constraints into the objective function, (15) becomes (16) where the μ_j values are Karush-Kuhn-Tucker (KKT) multipliers:

$$(16) \qquad \max_{\{\lambda_j\}} u\left(w_0 - \textstyle\sum_j \lambda_j p_j\right) + \delta \mathbb{E}_\pi\left[u\left(w_1 + \textstyle\sum_j \lambda_j x_{js}\right)\right]\theta$$
$$+ \delta\left[\alpha \max_{s\in S} u\left(w_1 + \textstyle\sum_j \lambda_j x_{js}\right) + (1-\alpha) \min_{s\in S} u\left(w_1 + \textstyle\sum_j \lambda_j x_{js}\right)\right](1-\theta)\} - \textstyle\sum_j \lambda_j \mu_j.$$

A difficulty in extending the model beyond the case of one risky asset is that when maximizing (16), the partial derivatives of the max and min functions will not generally exist everywhere. To address this issue and for illustration purposes, we make the following simplifying assumption: We let all assets have the same (unique) worst state and the same (unique) best state. This amounts to an assumption that asset returns are highly correlated in the best and worst market conditions, consistent with the notion that economies transition through good times and bad times that affect the entire market. Note that this assumption places no restriction on the relationship between asset returns in intermediate states, nor does it restrict the relationship between returns of different assets in the best and worst states.

Denote the state with the best and worst returns by $\overline{s}$ and $\underline{s}$, respectively. The assumption that all assets have their best returns in the same state and their worst returns in the same state is a strong assumption. There is, however, some surprising empirical support for this assumption, at least as an approximation. For instance, Kim et al. (2015) present the data in Figure 1, which depicts the average correlation across the daily returns of the Thomson Reuters Datastream Global Equity Indices (Industry Classification Benchmark, Level 2) which spans 10 different industries, conditioned on the market returns from 1973 to 2010. Kim et al. note, "The most distinctive feature is that the average correlations during the extreme downside and the extreme upside are considerably higher than the correlations under the normal market states. The principal implication is that while industry index prices move relatively independently during normal market states, they go down (up) simultaneously during extremely bad (good) times. (p.21)".

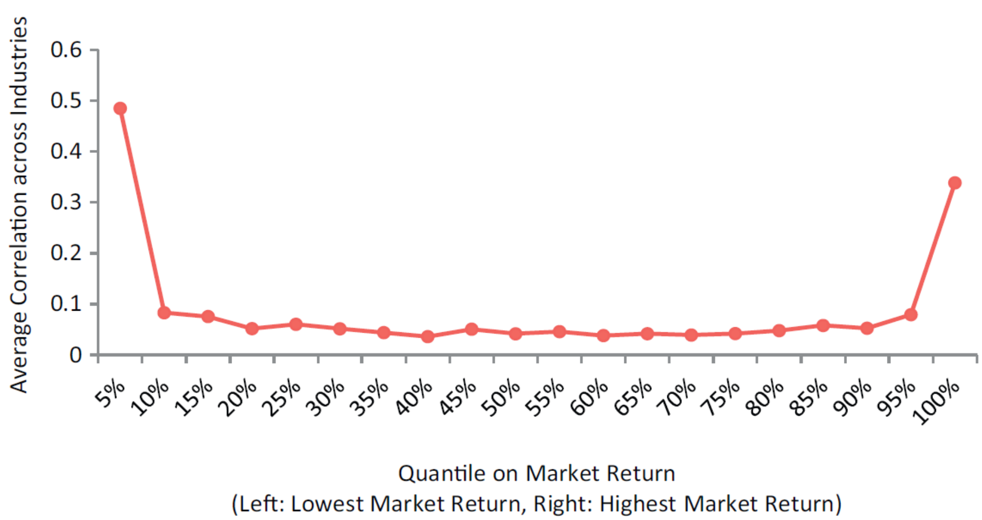

Figure 1. Average correlation across industries from 1973 to 2010 conditioned on market returns. (Reproduced from Kim et al., 2015)

Let $u(\overline{c})$ and $u(\underline{c})$ denote the maximum utility and minimum utility across all states in Period 1, and let $x_{j\overline{s}}$ and $x_{j\underline{s}}$ likewise denote the maximum and minimum randomized payoffs across states.

Under our correlation assumption, and given $\lambda_j \geq 0$ for all j with at least one strict inequality, we have $\sum_j \lambda_j x_{j\overline{s}} > \sum_j \lambda_j x_{js} > \sum_j \lambda_j x_{j\underline{s}}$ for all $s \neq \overline{s}, \underline{s}$. Setting the derivative with respect to each λ_j equal to zero yields the first order conditions:

$$-p_j u'(c_0) + \delta\left(\mathbb{E}_\pi\left[u'(c_s)x_{js}\right]\theta + \left[\alpha u'(\overline{c})x_{j\overline{s}} + (1-\alpha)u'\left(\underline{c}\right)x_{j\underline{s}}\right](1-\theta)\right) - \mu_j = 0 \text{ for all j.}$$

The KKT conditions require $\mu_j \lambda_j = 0$ for all j. Consider the case where the representative investor holds the market portfolio (consisting of all securities in the market). Then $\lambda_j > 0$ for all j, and thus $\mu_j = 0$ for all securities. Thus, the market portfolio is a local optimum. However, we are only assured that the solution is a global maximum when the objective function is concave (i.e., when $\alpha = 0$). Nevertheless, as the KKT conditions are necessary for a global maximum, the equilibrium price of any asset j can be solved for directly in the general case, yielding (17):

$$\text{(17)} \qquad p_j = \delta\left(\mathbb{E}_\pi[u'(c_s)x_s]\theta + \left[\alpha u'(\overline{c})x_{j\overline{s}} + (1-\alpha)u'(\underline{c})x_{j\underline{s}}\right](1-\theta)\right)/u'(c_0).$$

Next, define $\delta_s \coloneqq \delta u'(c_s)/u'(c_0)$. The quantity δ_s is the familiar 'stochastic discount factor' widely studied in asset pricing models. Also, define $\delta_{\overline{s}} \coloneqq \delta u'(\overline{c})/u'(c_0)$ and $\delta_{\underline{s}} \coloneqq \delta u'(\underline{c})/u'(c_0)$. The pricing formula for asset j can then be written:

$$\text{(18)} \qquad p_j = \mathbb{E}_\pi\left[\delta_s x_{js}\right]\theta + \left[\alpha(\delta_{\overline{s}})(x_{j\overline{s}}) + (1-\alpha)(\delta_{\underline{s}})(x_{j\underline{s}})\right](1-\theta).$$

Using a covariance decomposition, we can write:

$$\text{(19)} \qquad p_j = \mathbb{E}_\pi[\delta_s]\mathbb{E}_\pi[x_{js}]\theta + \operatorname{cov}\left(\delta_s, x_{js}\right)\theta + \left[\alpha(\delta_{\overline{s}})(x_{j\overline{s}}) + (1-\alpha)(\delta_{\underline{s}})(x_{j\underline{s}})\right](1-\theta).$$

In formula (19), the price of an asset depends on (i) the fundamental value of the asset (discounted expected dividend); (ii) a risk premium; (iii) a positive skewness premium; (iv) an ambiguity robustness premium; and (v) the investor's degree of optimism, α. To see this, note that we can rewrite (19) as (20):

$$\text{(20)} \qquad p_j = \mathbb{E}_\pi[\delta_s]\mathbb{E}_\pi[x_{js}] + \mathcal{R}_j + \alpha\mathcal{S}_j + (1-\alpha)\mathcal{M}_j$$

$$\textit{where} \quad \mathcal{R}_j = [\mathbb{E}_\pi\left[(\delta_s)(x_{js})\right] - \mathbb{E}_\pi[\delta_s]\mathbb{E}_\pi[x_{js}]]\theta = \operatorname{cov}\left(\delta_s, x_{js}\right)\theta$$

$$\mathcal{S}_j = \left[[(\delta_{\overline{s}})(x_{j\overline{s}})] - \mathbb{E}_\pi[\delta_s]\mathbb{E}_\pi[x_{js}]\right](1-\theta)$$

$$\mathcal{M}_j = \left[[(\delta_{\underline{s}})(x_{j\underline{s}})] - \mathbb{E}_\pi[\delta_s]\mathbb{E}_\pi[x_{js}]\right](1-\theta)$$

In (20), $\mathcal{R}_j$ is the risk premium, as in standard consumption-based asset pricing models which is determined by the covariance of the returns on asset j with consumption. The quantity $\mathcal{S}_j$ can be interpreted as a positive skewness premium: For two assets with the same values of $\mathbb{E}_\pi[\delta_s], \mathbb{E}_\pi[x_{js}]$, and $\delta_{\bar{s}}$, the asset with more positive skewness (in the sense of having a higher maximum payoff) will be priced higher. The quantity $\mathcal{M}_j$ can be interpreted as a premium for assets that are more robust to model uncertainty: For two assets with the same values of $\mathbb{E}_\pi[\delta_s], \mathbb{E}_\pi[x_{js}]$, and $\delta_{\underline{s}}$, the asset which is more robust to ambiguity (in the sense of having a higher minimum payoff) will be priced higher. Note that in (20), prices depend on both market fundamentals (expected dividends) and 'animal spirits' (the market degree of optimism, α).

The equilibrium pricing formula for p_j in (20) includes four important special cases: If the investor has subjective expected utility preferences ($\theta = 1$), then $r = 1/\mathbb{E}_\pi[\delta_s]$ and prices are:

(21) $$\tilde{p}_j^{SEU} = \frac{\mathbb{E}_\pi[x_{js}]}{r} + \text{cov}(\delta_s, x_{js}).$$

If the investor has Hurwicz (1951) preferences ($\theta = 0$), based on the Hurwicz criterion for robust decision making, then equilibrium prices do not depend on the investor's beliefs and are:

(22) $$\tilde{p}_j^{Hurwicz} = \alpha(\delta_{\bar{s}})(x_{j\bar{s}}) + (1-\alpha)(\delta_{\underline{s}})(x_{j\underline{s}}).$$

If the investor follows Wald's (1950) maximin criterion for robust decisions under uncertainty ($\theta = 0, \alpha = 0$), then equilibrium prices, (23), do not depend on the utility function or beliefs:

(23) $$\tilde{p}_j^{Wald} = \frac{x_{j\underline{s}}}{r}.$$

If the investor has linear utility, equilibrium prices from (20) reduce to (24):

(24) $$\tilde{p}_j = (1/r)(\mathbb{E}_\pi[x_{js}] + (1-\theta)[\alpha(x_{j\bar{s}} - \mathbb{E}_\pi[x_{js}]) + (1-\alpha)(x_{j\underline{s}} - \mathbb{E}_\pi[x_{js}])]).$$

Formula (24) was derived in a setting with a single asset by Chateauneuf et al. (2007). We have derived the same formula as a special case of (20) for an arbitrary number of assets, and have highlighted the correlation assumption that makes this derivation possible. In (24), the positive skewness premium is $(x_{j\bar{s}} - \mathbb{E}_\pi[x_{js}])$, and the ambiguity robustness premium is $(x_{j\underline{s}} - \mathbb{E}_\pi[x_{js}])$.

Chateauneuf et al. (2007) noted that the equity premium under (24) when $\alpha = 0$ exceeds the equity premium under SEU. They defined the equity premium as the ratio of the expected return from the single risky asset in their economy to the price for that asset multiplied by the risk-free rate. However, they did not provide a calibration indicating whether the magnitude of the equity premium is sufficient to account for the equity premium puzzle of Mehra and Prescott (1985).

Recently, Doeswijk et al. (2017) estimated the historical returns of what they termed the global market portfolio (GMP) for over half a century (from 1960 through 2015). They found (Doeswijk et al., Table 2) that the GMP had an average annual real return[6] of 5.04%, a maximum annual real return of 33.64%, and a minimum annual real return of -25.25%. During their observation period, the real return on U.S. Treasury Bills was 1.13%. As a simple illustration, we apply the definition of the equity premium from Chateauneuf et al. to the data on the GMP from Doeswijk et al. (2017), using the support of real returns of the GMP from Doeswijk et al. as the support of real returns for the market portfolio held by the representative investor. We can calculate the equity premium as $\mathbb{E}_\pi[r_{ms}]/(\tilde{p}_m r)$ where $\mathbb{E}_\pi[r_{ms}]$ is the expected return on the GMP, $\tilde{p}_m$ is the price per share of the GMP (from (24)) and r is the risk-free rate. To illustrate, let $\delta = 1$. Then if $\theta = 1$, the equity premium is 0% (since all agents are risk-neutral and maximize expected return). However, for $\theta = 0.8$ and $\alpha = 0$, (QRD model (2)), the equity premium is 6.12%. This is close to the 6.18% equity premium estimated by Mehra and Prescott (1985) and it obtains *even when utility is linear*. In contrast, Mehra and Prescott find that for EU with a power utility function, even a risk aversion coefficient of 10, far above typical empirical estimates, cannot produce an equity premium greater than 0.35%, far below the 6.18% premium they observed. For $\theta = 0.75$ and $\alpha = 0$, the equity premium is 7.769%. If α is small but positive (e.g., $\alpha = 0.10$), then for $\theta = 0.75$, the equity premium is 6.17%. Thus, security bias can rationalize the equity premium puzzle.

Returning to the general formula in (20), we now seek to derive a formula for the expected return on asset j. Since returns satisfy $r_{js} = x_{js}/p_j$, they satisfy the relationship:

$$(25) \quad 1 = \mathbb{E}_\pi[\delta_s]\mathbb{E}_\pi[r_{js}]\theta + \text{cov}(\delta_s, r_{js})\theta + \left[\alpha(\delta_{\bar{s}})(r_{j\bar{s}}) + (1-\alpha)(\delta_{\underline{s}})(r_{j\underline{s}})\right](1-\theta),$$

which is a special case of (19). For a risk-free asset (that pays the same return, r, in every state):

$$(26) \quad 1 = r\left(\mathbb{E}_\pi[\delta_s]\theta + \left[\alpha(\delta_{\bar{s}}) + (1-\alpha)(\delta_{\underline{s}})\right](1-\theta)\right).$$

For linear utility and $\delta = 1$, the equilibrium risk-free rate from (26) is given by $r = 1$. Thus, QRD can generate both a high equity premium and a low-risk free rate in equilibrium, a combination of predictions that has proved to be challenging in models of consumption-based asset pricing.

---

[6] We used the arithmetic average real return from Doeswijk et al. of 5.04%. Similar results (approximately 6% equity premium) obtain with the same parameter values if one uses their reported compounded average real return of 4.38%.

To derive a formula for expected returns, equating (25) and (26) and rearranging terms yields:

$$(27) \quad \mathbb{E}_{\pi}[r_{js}] - r = \frac{-\mathrm{cov}(\delta_s, r_{js})\theta + \left[\alpha(\delta_{\bar{s}})(r - r_{j\bar{s}}) + (1-\alpha)(\delta_{\underline{s}})(r - r_{j\underline{s}})\right](1-\theta)}{\mathbb{E}_{\pi}[\delta_s]\theta}.$$

Next, define the following notation:

$$(28) \quad \beta_{1,j} := \left(\frac{\mathbb{E}_{\pi}\left[\delta_s\left(r_{js} - \mathbb{E}_{\pi}[r_{js}]\right)\right]}{\mathrm{var}(\delta_s)}\right); \beta_{2,j} := \left(\frac{\delta_{\bar{s}}(r_{j\bar{s}} - r)}{\mathrm{var}(\delta_s)}\right); \ \beta_{3,j} := \left(\frac{\delta_{\underline{s}}(r_{j\underline{s}} - r)}{\mathrm{var}(\delta_s)}\right).$$

Also, let $D(\delta_s) := \mathrm{var}(\delta_s)/\mathbb{E}_{\pi}[\delta_s]$ denote the index of dispersion associated with the stochastic discount factor. Then the relationship in (27) can be expressed as a beta asset pricing model:

Proposition 8: *In a representative agent economy with equilibrium price* (20), *the equilibrium expected return on any asset j is given by* (29) *with $\beta_{1,j}$, $\beta_{2,j}$, and $\beta_{3,j}$ from* (28):

$$(29) \quad \mathbb{E}_{\pi}[r_{js}] = r - \beta_{1,j}[D(\delta_s)] - \beta_{2,j}[\alpha(1-\theta)D(\delta_s)/\theta] - \beta_{3,j}[(1-\alpha)(1-\theta)D(\delta_s)/\theta].$$

In (28), the *risk factor* $\beta_{1,j}$ can be written equivalently as $\beta_{1,j} = \left(\frac{\mathrm{cov}(\delta_s, r_{js})}{\mathrm{var}(\delta_s)}\right)$ which is the same as in the standard consumption-based asset pricing models. The interesting content of (29) is the emergence of two new asset pricing factors that come out of our analysis: $\beta_{2,j}$ can be viewed as a *positive skewness factor* that reduces the expected returns in equilibrium for assets with high potential (high maximum payoff); $\beta_{3,j}$ can be viewed as an *ambiguity robustness factor* that reduces the expected returns in equilibrium for assets that are more robust to model uncertainty (in the sense of having a higher minimum payoff). As a result, value stocks and small-cap stocks will tend to be underpriced in equilibrium since they are often thought to be vulnerable to bankruptcy (low minimum payoff). As a consequence, they will earn higher expected returns in equilibrium, consistent with the findings of Fama and French (1992). In contrast, growth stocks are viewed to have great potential (high maximum payoff) and so, will be priced higher by the market and will earn lower returns in equilibrium under (29). Thus, from (28, 29) in equilibrium, assets with greater risk (positive covariance with consumption implying negative covariance with the stochastic discount factor) earn a higher expected return, assets with greater positive skewness (higher maximum return $r_{j\bar{s}} > r$ and thus higher $\beta_{2,j}$) earn lower expected returns, and assets with greater robustness to ambiguity (higher minimum return $r_{j\underline{s}} < r$, and thus higher $\beta_{3,j}$) earn lower expected returns. In this respect, positive skewness factor, $\beta_{2,j}$, and ambiguity robustness factor, $\beta_{3,j}$, that emerge from our analysis may capture some of the information embedded in the Fama-French (1992) factors.

IX. A THREE-FACTOR MODEL OF BEHAVIORAL BIASES

We have considered an intuitive analogy between quasi-rank dependent probability weighting and quasi-hyperbolic time discounting. We now show that there is a plausible formal relation between these models by extending QRD to incorporate both risk and time preferences. To do so, we assume three payoffs are particularly salient when evaluating an uncertain consumption stream: the worst possible outcome, the best possible outcome, and the 'certain-immediate' outcome – the most the agent can be guaranteed in the present period. This is consistent with the intuition that a person may focus on the worst-case scenario, the best-case scenario, and the 'offer on the table'.

A. *Extension to Choice over Time*

Formally, let $X \subset \mathbb{R}$ denote a finite set of possible outcomes, and let $T = \{0,1,2,\dots,\mathrm{T}\}$ denote a finite set of discrete time periods. A *consumption sequence* $x_j \coloneqq \{x_{j1},\dots,x_{j\mathrm{T}}\}$ is a sequence of dated outcomes. An outcome in X received in period t from a consumption sequence x_j is denoted x_{jt}. We index consumption sequences by $j \in \{1,2,\dots,n\}$. Denote the set of consumption sequences by C. A *stochastic consumption plan*, $f\colon C \rightarrow [0,1]$, is a probability distribution on C, where $f(x_j)$ is the probability that the decision maker receives consumption sequence j if he chooses f. Denote the set of stochastic consumption plans by Ω. The minimum outcome available from a stochastic consumption plan f in period 0 is denoted $\underline{x}_0$ and the maximum and minimum outcomes that can be obtained from f across all periods $t \in \{0,1,2,\dots,\mathrm{T}\}$ are denoted by $\overline{x}$ and $\underline{x}$, respectively, where we omit the dependence on f for notational convenience. Let $\succsim$ denote a weak preference relation on Ω with $\succ$ and $\sim$ denoting strict preference and indifference, respectively.

We propose that three payoffs are particularly salient to a decision maker when evaluating a stochastic consumption plan, f: the best outcome, $\overline{x}$, the worst outcome, $\underline{x}$, and the 'certain-immediate' outcome, $\underline{x}_0$, which combines two properties - certainty, and immediacy, and represents the most the agent can be guaranteed in the present period from f. We account for these payoffs by generalizing (2) and (10, 11) to encompass both risk and time preferences. In particular, we consider a decision maker who evaluates stochastic consumption plans such that for any $f, g \in \Omega$, $f \succsim g$ if and only if $V(f) \geq V(g)$, where:

$$V(f) = \theta \sum_t \delta^t \,[U(f)] + (1-\theta)[\alpha u(\overline{x}) + \beta u(\underline{x}_0) + \gamma u(\underline{x})] \qquad (30)$$

where $U(f) = \sum_j f(x_{jt}) \cdot u(x_{jt})$, and $\alpha, \beta, \gamma, \theta, \delta \in [0,1]$.

In (30), a decision maker exhibits a disproportionate preference for alternatives with higher minimum payoffs, higher maximum payoffs, and higher immediate payoffs. In particular, the decision maker maximizes a weighted average of discounted expected utility and 'salient utility' by overweighting the best, worst, and immediate outcomes.

Formula (30) provides a 'three-factor model' of behavioral biases, perhaps analogous in spirit to the five-factor model of personality traits in psychology. The three factors, $\alpha, \beta,$ and γ, have intuitive interpretations as indexing the decision maker's degree of (positive) *skewness preference*, *present bias*, and *security bias*, respectively. Security bias unifies a bias toward certainty in the Allais paradoxes, loss aversion for mixed gambles, the sign effect for choice over time, and a preference for robustness to ambiguity when probabilities are subjective. Skewness preference produces systematic deviations from risk aversion for choices under risk and produces systematic deviations from consumption smoothing for choices over time. When $(\alpha, \beta, \gamma) = (0,1,0)$, model (30) reduces to quasi-hyperbolic discounting for choices over time. When $(\alpha, \beta, \gamma) = (0,0,1)$, (30) reduces to QRD from (2) for choices under risk. By estimating $\alpha, \beta,$ and γ, one can determine the relative strength of skewness preference, present bias, and security bias for a decision maker.

When $\theta = 0.5$, the three-factor model in (30) can be equivalently written as:

(31) $$V(f) = \sum_t \delta^t \, [U(f)] + \alpha u(\overline{x}) + \beta u(\underline{x}_0) + \gamma u(\underline{x}).$$

The three-factor model has the appealing feature that it is linear in behavioral biases and so should be convenient when investigating the effects of particular biases on economic behavior.

If the consumer's true preferences satisfy the axioms of rational choice (i.e., are consistent with discounted expected utility theory), the parameter θ represents the degree to which the consumer deviates from the optimal dynamic programming solution to his preference maximization problem by systematically overweighting the best-case outcome, the worst-case outcome, and the certain-immediate outcome. Under this interpretation, the three-factor model provides a coherent and general alternative to the neo-classical view of perfect optimizing behavior that applies to both risk and time preferences and can also be applied to a setting with purely subjective probabilities. When $\theta = 1$, the consumer perfectly optimizes. For lower values of θ, the consumer increasingly focuses on the salient payoffs and places less weight on maximizing discounted expected utility. In the other extreme where $\theta = 0$, the decision maker more closely resembles a 'greedy algorithm' that seeks to obtain the most that can be guaranteed immediately and neglects all future consequences and all uncertain outcomes except those yielding the best or worst payoff in the stochastic consumption plan. This interpretation suggests the testable prediction that when two alternatives have the same best, worst, and immediate outcomes, behavior will conform more closely to discounted expected utility theory.

To illustrate the three-factor model, we will adopt a simple specification of (30) where $\alpha = \beta = \gamma = 1$. Note that this specification still gives greater weight to security bias than to skewness preference since $\underline{x}_0 = \underline{x}$ for decisions involving only risk. Fixing $\alpha, \beta,$ and γ to be equal, the only parameters in (30) are the agent's discount factor, δ, utility function, u, and bias parameter, θ.

We will show that the three-factor model explains four parallels between risk and time preferences: (i) Present bias (Laibson, 1997) and the certainty effect (Kahneman and Tversky, 1979); (ii) the finding that risk interacts with time preferences (Keren and Roelofsma, 1995) and that time interacts with risk preferences (Baucells and Heukamp, 2010; Abdellaoui et al., 2011); (iii) the fourfold pattern of risk attitudes (Tversky and Kahneman, 1992) and a bias toward concentration (Koszegi and Szeidl, 2013), and (iv) loss aversion (Kahneman and Tversky, 1979) and the sign effect (Prelec and Loewenstein, 1991). We will illustrate the three-factor model using the same simple specification. In addition to setting $\alpha = \beta = \gamma = 1$, we let $u(x) = x$, and set the agent's annual discount factor at $\delta = 0.95$. In our parametric specification, the classical examples of behaviors (i), (ii), and (iii) each hold for all $\theta \in [0.00, 0.90]$. In addition, given any $\theta \in (0,1)$, a sufficient condition for both the sign effect and loss aversion is $\alpha < \gamma$ (skewness preference is weaker than security bias). We assume in all propositions to follow that the decision maker has preferences given by the three-factor model in (30).

B. Present Bias and the Certainty Effect

Since Prelec and Lowenstein (1991), there has been an intuitive analogy between the certainty effect and present bias. In Section III, we saw that QRD provides a simple account of the Allais paradox and common ratio effect with a certain payoff. Next, we consider present bias. For all the following analyses, we let preferences be given by the three-factor model (30). Let (x, p, t) denote a stochastic consumption plan that pays $x > 0$ with probability p at time t and pays 0 otherwise.

Definition 5: (Present Bias): Let $f \coloneqq (y, 1,0), f' \coloneqq (y, 1, s), g \coloneqq (x, 1, t),\ g' \coloneqq (x, 1, t + s)$, for any $y \in (0, x), and\ t, s \in (0, \infty)$. *Present bias holds* if $f \sim g$ implies $f' \prec g'$.

Let $DU(f)$ denote the discounted utility of f, and set the time horizon for the decision maker to span periods in the interval $[0, t + s]$. For all periods not specified by each consumption plan, the plan yields a payoff of 0.

Proposition 9: *Let $DU(f) = DU(g)$. Then present bias holds for all $\alpha, \beta, \gamma, \theta \in (0,1)$.*

For the present bias example in Table 1, we have $(x, y, t, s) = (110, 100, 4\ weeks, 26\ weeks)$. Definition 5 implies an agent who is indifferent between A and B in Choice Set 5, will strictly prefer B from Choice Set 6. Under our parametric specification ($\alpha = \beta = \gamma = 1, u(x) = x$, and an annual discount factor of $\delta = 0.95$), the preference for A in Choice Set 5 and for B in Choice Set 6 holds for all $\theta \in [0.00, 0.90]$.

C. Interaction Effects between Risk and Time Preferences

Systematic interaction effects have also been documented for risk and time preferences. Table 1 illustrates seven empirical phenomena of interest. We have already discussed the Allais paradox, common ratio effect and present bias. The fourth phenomenon in Table 1 captures an interaction between risk and time in which a delay is added to decisions under risk (Baucells and Heukamp 2010). Since adding the three-month delay shifts preference toward the riskier lottery, we say 'time affects choice under risk.' The fifth phenomenon captures a different interaction effect in which uncertainty is introduced into decisions over time (Keren and Roelofsma 1995). Since introducing uncertainty shifts preference toward the delayed lottery, we say 'risk affects choice over time.' The sixth observation is that delaying all outcomes by the same amount reduces the common ratio effect. The seventh finding is that making all outcomes uncertain reduces present bias.

Table 1. Choices between Options A and B involving Risk and Time

Observation	CS	Option A	vs.	Option B
Allais paradox (Allais, 1953)	1 2	**(1 M, for sure, now)** (1 M, 11%, now)		(1 M, 89%; 5 M, 10%, now) **(5 M, 10%, now)**
Common ratio effect (Baucells & Heukamp, 2010)	3 4	**(9, for sure, now)** (9, 10%, now)		(12, 80%, now) **(12, 8%, now)**
Present bias (Keren & Roelofsma, 1995)	5 6	**(100, for sure, now)** (100, for sure, 26 weeks)		(110, for sure, 4 weeks) **(110, for sure, 30 weeks)**
Time affects choice under risk (Baucells & Heukamp, 2010)	7 8	**(9, for sure, now)** (9, for sure, 3 months)		(12, 80%, now) **(12, 80%, 3 months)**
Risk affects choice over time (Keren & Roelofsma, 1995)	9 10	**(100, for sure, now)** (100, 50%, now)		(110, for sure, 4 weeks) **(110, 50%, 4 weeks)**
Common ratio depends on delay (Baucells & Heukamp, 2010)	11 12	(9, for sure, 3 months) (9, 10%, 3 months)		**(12, 80%, 3 months)** **(12, 8%, 3 months)**
Present bias depends on risk (Keren & Roelofsma, 1995)	13 14	(100, 50%, now) (100, 50%, 26 weeks)		**(110, 50%, 4 weeks)** **(110, 50%, 30 weeks)**

Complementary probabilities correspond to payoffs of 0. Majority responses of experimental subjects are highlighted in bold. "CS" denotes "Choice Set".

As shown in Table 1, systematic interaction effects involving risk and time have been observed in which incorporating risk into intertemporal choices produces more patient behavior (Keren and Roelofsma, 1995) and incorporating delays into risky choices reduces risk aversion (Baucells and Heukamp, 2010; Abdellaoui et al., 2011). Such behaviors cannot be explained by any models in which probability weighting functions and time discount functions are multiplicatively separable. This results in a descriptive limitation of the quasi-hyperbolic model. As Levine (2012, p. 95) writes: "The only problem with the model is that it predicts that present bias should not depend on whether or not the reward is uncertain. Unfortunately this is not the case."

The problem is naturally resolved in the more general setting presented here. In addition to predicting that present bias depends on whether the reward stream is uncertain, the three-factor model also predicts that the common ratio effect depends on whether or not the lotteries are delayed. Both of these predictions are supported experimentally. The first prediction concerns the effect of delaying outcomes in decisions under risk, which systematically reduces risk aversion (Baucells and Heukamp, 2010; Abdellaoui et al., 2011). We can formalize this effect as follows:

Definition 6: (Time affects Choice under Risk): Let $f \coloneqq (y,1,0), f' \coloneqq (y,1,t), g \coloneqq (x,p,0)$, $g' \coloneqq (x,p,t)$, for any $y \in (0,x),\ p \in (0,1), and\ t \in (0,\infty)$. *Time reduces risk aversion* if $f \sim g$ implies $f' \prec g'$.

Let $DEU(f)$ denote the discounted expected utility of f, and set the time horizon for the decision maker to span periods in the interval $[0,t]$. For all periods not specified by each stochastic consumption plan, the plan yields a payoff of 0.

Proposition 10: *Let $DEU(f) = DEU(g)$. Then time reduces risk aversion for all $\alpha, \beta, \gamma, \theta \in (0,1)$*

In the experimental example in Table 1, $(x,y,p,t) = (12, 9, 0.8, 3\, months)$. Definition 6 implies that an agent who is indifferent between A and B in Choice Set 7, will strictly prefer B from Choice Set 8. Under our running parametric specification ($u(x) = x$, and an annual discount factor of $\delta = 0.95$), the preference for A in Choice Set 7 and for B in Choice Set 8 holds for all $\theta \in [0.00, 0.90]$.

A second prediction of the three-factor model concerns how uncertainty affects choice over time. In particular, we consider the finding by Keren and Roelofsma that introducing risk into intertemporal decisions reduces impatience. This behavior can be defined as follows:

Definition 7: (Risk affects Choice over Time): Let $f \coloneqq (y, 1, 0), f' \coloneqq (y, p, 0), g \coloneqq (x, 1, t)$, $g' \coloneqq (x, p, t)$, for any $y \in (0, x),\ p \in (0,1), and\ t \in (0, \infty)$. *Risk reduces impatience* if $f \sim g$ implies $f' \prec g'$.

Proposition 11: *Let $DEU(f) = DEU(g)$. Then risk reduces impatience for all $\alpha, \beta, \gamma, \theta \in (0,1)$.*

In the experimental example in Table 1 $(x, y, p, t) = (110, 100, 0.5, 4\, weeks)$. Definition 7 implies that an agent who is indifferent between A and B in Choice Set 9, will strictly prefer B from Choice Set 10. Under our running parametric specification ($u(x) = x$, and an annual discount factor of $\delta = 0.95$), the preference for A in Choice Set 9 and for B in Choice Set 10 holds for all $\theta \in [0.00, 0.90]$.

The predictions that time affects choice under risk and that risk affects choice over time are less obvious than more familiar behaviors (certainty effect, present bias), but these interaction effects between risk and time are supported by the empirical data, as illustrated in Table 1. Moreover, the opposite preference patterns are not predicted, revealing these behaviors to be strong implications of the three-factor model.

D. The Fourfold Risk Pattern and a Bias Toward Concentration

A robust property of observed risk preferences is the fourfold pattern observed by Tversky and Kahneman (1992) and illustrated in Table 2. Under the fourfold pattern of risk preferences, a decision maker is risk-averse for gains of high probability and losses of low probability, but is risk-seeking for gains of low probability and losses of high probability. An example of the fourfold pattern from Tversky and Kahneman (1992) is provided in Table 2.

Table 2. Choices illustrating the Fourfold Risk Pattern and a Bias Toward Concentration

Observation	CS	Option A	vs.	Option B
	1	**(95, for sure)**		(100, 95%; 0, 5%)
Fourfold Risk Pattern	**2**	**(-5, for sure)**		(-100, 5%; 0, 95%)
(Tversky and Kahneman, 1992)	**3**	(-95, for sure)		**(-100, 95%; 0, 5%)**
	4	(5, for sure)		**(100, 5%; 0, 95%)**
	5	**(100, 0, 0, …, 0, 0, 0)**		(0, 1, 1 ,… 1, 1, 1)
Bias Toward Concentration	**6**	**(0, 0, 0, …, 0, 0, 100)**		(1, 1, 1, … 1, 1, 0)
(Koszegi and Szeidl, 2013)	**7**	(-100, 0, 0,…, 0, 0, 0)		**(0, -1, -1,…, -1, -1, -1)**
	8	(0, 0, 0,…, 0, 0, -100)		**(-1,-1, -1,…, -1, -1, 0)**

The prototypical response pattern is highlighted in bold. “CS” denotes “Choice Set”.

In the example, a person chooses \$95 with certainty over a 95% chance of \$100 and a 5% chance of \$0 (risk aversion for high probability gains), and chooses a guaranteed \$5 loss over a 5% chance of losing \$100 (risk aversion for low probability losses). However, the decision maker also prefers a 95% chance of losing \$100 over a guaranteed loss of \$95 (risk seeking for loses of high probability) and prefers a 5% chance of winning \$100 over a \$5 with certainty (risk seeking for gains of low probability).

Under (30) with $u(x) = x$, the fourfold pattern illustrated in Table 2 holds for all $\alpha, \theta \in (0,1)$. Formally, for decisions involving only risk (and no time) we have the following definition:

Definition 8: When comparing lotteries $f \coloneqq (x, p; 0, 1-p)$ and $E(f) \coloneqq (xp, 1)$, the *fourfold pattern of risk attitudes* holds if $E(f) \succ f$ for sufficiently high p and $f \succ E(f)$ for sufficiently low p when $x > 0$, but $f \succ E(f)$ for sufficiently high p and $E(f) \succ f$ for sufficiently low p when $x < 0$.

Proposition 12: *Let $u(x) = x$. Then the fourfold pattern of risk attitudes holds for all $\alpha, \gamma, \theta \in (0,1)$.*

A related pattern of behavior in intertemporal choice referred to as a ‘bias toward concentration’ was identified by Koszegi and Szeidl (2013). They recognized that people often prefer consumption sequences with a few large advantages relative to consumption sequences with many small advantages and prefer consumption sequences with many small disadvantages relative to those with a few large disadvantages.

A ‘fourfold’ pattern of time preference due to a bias toward concentration is illustrated in the bottom panel of Table 2. In each of choices 5 – 8, a consumer chooses between gains or losses over a horizon of 100 periods. In Choice Set 5, a consumer prefers \$100 in period 0 over a steady stream of \$1 payoffs in periods 1 through 100. This behavior is reminiscent of a preference for a large lump-sum payment from winning a lottery instead of receiving an annuity. Indeed, people often prefer lump-sum gains instead of annuities (e.g., Brown, Casey, and Mitchell, 2007). In Choice Set 6, the consumer prefers \$100 in period 100 over a stream of \$1 payoffs in periods 0 through 99, thereby exhibiting future bias for delayed concentrated gains. This behavior is consistent with goal-seeking behavior in which small short-term gains from leisure activities are foregone when working to achieve a personal or career goal. In Choice Set 7, the consumer prefers to make \$1 payments in each of periods 1 through 100 than to pay \$100 outright in period 0. This behavior reflects a preference for financing and making regular installment payments, such as on a car or a mortgage or a consumer durable, rather than paying the full cost up front. In Choice Set 8, the consumer prefers to pay \$1 in each of periods 0 through 99 than to pay \$100 in period 100, exhibiting future bias for delayed concentrated losses. This behavior is reminiscent of the decision to invest a small cost to exercise regularly to avoid large health costs in the future. Whether a consumer chooses to exercise may depend on whether the consumer adopts this 'many period' frame, or views each decision in isolation as a binary choice to exercise now or later.

The fourfold pattern of risk preferences and the bias toward concentration can each be described more succinctly as a twofold pattern: For risk, the fourfold pattern is equivalent to a *preference for positively skewed probability distributions* (producing a preference for state lottery tickets) and an *aversion to negatively skewed distributions* (producing a preference to purchase insurance). For time, the bias toward concentration is equivalent to a *preference for positively skewed consumption sequences* (producing a preference for large lump-sum gains over dispersed small gains) and an *aversion to negatively skewed consumption sequences* (producing a preference for small monthly payments over a single large payment). More formally, we have the following:

Definition 9: Consider a choice between consumption sequences $A \coloneqq (nx, 0, \ldots, 0)$ that yields payoff nx in period 0 and pays 0 in periods 1 through n, and $B \coloneqq (0, x, \ldots, x)$ that pays 0 in period 0 and yields payoff x in each of periods 1 through n. Next, consider a choice between consumption sequences $A' \coloneqq (0, \ldots, 0, nx)$ that yields payoff nx in period n and pays 0 in periods 0 through $n-1$, and $B' \coloneqq (x, \ldots, x, 0)$ that pays 0 in period n and yields x in each of periods 0 through $n-1$. A *bias toward concentration* holds if, for sufficiently large n, $A \succ B$ and $A' \succ B'$ for $x > 0$, and $B \succ A$ and $B' \succ A'$ for $x < 0$.

We analyze the undiscounted case which may be viewed as an approximation.

Proposition 13: *Let $u(x) = x$, and $\delta = 1$. Then a bias toward concentration holds for all $n > 1 + \max(\beta/\alpha, \beta/\gamma)$, for any $\alpha, \beta, \gamma, \theta \in (0,1)$.*

The quantity β/α in Proposition 13 reflects a tradeoff between the effect of present bias and a bias toward concentration. Higher values of β imply greater present bias while higher values of α imply greater concentration bias for gains. Similarly, higher values of γ imply greater concentration bias for losses. The three-factor model thus provides an approach to unify present bias and concentration bias.

E. Loss Aversion and the Sign Effect

We close this section with a fourth striking parallel between risky and intertemporal choice – the gain-loss asymmetry of loss aversion (Kahneman and Tversky, 1979) for choice under risk, and the gain-loss asymmetry referred to as the sign effect (Prelec and Loewenstein, 1991) for choices over time. For any $\theta \in (0,1)$, a sufficient condition for loss aversion (aversion to symmetric mixed gambles) under the three-factor model is $\alpha < \gamma$ (skewness preference is weaker than security bias). The sign effect for choice over time is the finding that people are more patient for losses than for gains. Prelec and Loewenstein (1991) define the sign effect as follows:

Definition 10: The *sign effect* holds if for all $y > x > 0, t > r \geq 0$, $(x,r) \sim (y,t)$ implies $(-x,r) \succ (-y,t)$.

Proposition 14: *Let $u(x) = x$. Then for any $\theta \in (0,1)$, the sign effect holds if and only if $\alpha < \gamma$.*

In contrast, the sign effect cannot be explained by a conventional behavioral approach that incorporates a loss aversion parameter into the value function of the decision maker. We demonstrate this in the appendix after the proof of Proposition 14.

In sum, the parametric restrictions $\gamma > \alpha$, and $\alpha, \beta, \gamma \in (0,1)$ generate the certainty effect and present bias, the fourfold pattern of risk preferences and a bias toward concentration, loss aversion and the sign effect, and observed interaction effects between risk and time. Table 3 summarizes the predictions of the three-factor model from (30) for these eight robust behaviors in the literature on risk and time preferences for any $\theta \in (0,1)$ and indicates the parameters driving each behavior.

Table 3. Predictions of the Three-Factor Model of Behavioral Biases

Risk	**Time**
Loss Aversion (γ)	Sign Effect (γ)
Time affects Risk Preference (β)	Risk affects Time Preference (β)
Certainty Effect (Allais, Common Ratio) (γ)	Present Bias, Time Inconsistency (β)
Skewness Preference (Fourfold Pattern) (α, γ)	Skewness Preference (Bias toward Concentration) (α, β, γ)

(α represents (positive) skewness preference; β represents present bias; γ represents security bias)

X. RELATED LITERATURE

The QRD model is a simple one-parameter generalization of expected utility theory. It preserves three basic properties of expected utility analysis: transitivity, first-order stochastic dominance, and risk aversion, while accounting for the Allais paradoxes, loss aversion, and intuitive violations of continuity. Although the QRD model is simpler than leading models such as RDU and CPT, it is quite similar to earlier models in the literature which assumed that decision makers care about a 'security level' (Lopes 1986; Gilboa 1988; Jaffray 1988; Cohen 1992). Schmidt (2000) appears to be the first to explicitly state the form in (1), although it can be viewed as a special case in Gilboa (1988). Kopylov (2009) axiomatizes an analogous model for ambiguity. Models which provide a special role for the best and worst outcomes are characterized by Chateauneuf et al. (2007) for ambiguity and by Webb and Zank (2011) for choice under risk. Despite its simplicity, the QRD model has received little application, although Chateauneuf et al. (2007) apply their model to simultaneous gambling and insurance purchase[7] and to the equity premium puzzle.

It is peculiar that there is a large literature on non-expected utility models, but a relatively small literature on importing such models into applications. A possible reason is that leading descriptive models such as CPT are much less tractable than EU, whereas more tractable models of the certainty effect (e.g., Schmidt 1998) violate transitivity or stochastic dominance. In this respect, QRD may offer a compromise between descriptive adequacy and tractability[8].

One alternative model of choice under risk that has gained popularity in applications is the Koszegi-Rabin model of reference-dependent preferences (Koszegi and Rabin, 2006; 2007). That model can also explain loss aversion and the Allais paradox. The loss-aversion parameter, λ, in the Koszegi-Rabin model is consistent with rank dependent utility theory for $\lambda < 2$, but can violate stochastic dominance if $\lambda > 2$. The model also makes the strong assumption that agents have rational expectations regarding the reference point. In contrast, QRD does not violate stochastic dominance and makes no assumptions regarding rational expectations. Under QRD there is a natural reference point that is directly observable and well-specified for each lottery – the

[7] Chateauneuf et al. consider the decision of *whether* to purchase insurance and lottery tickets, but do not consider the *optimality* of purchasing full insurance.

[8] Throughout we have considered our analysis apart from CPT. But note that QRD (both in (2) and in (11)) is in fact a special case of rank-dependent utility theory, which itself is a special case of CPT for the domain of gains. Our analysis however makes no direct use of cumulative probability weighting functions or reference-dependent value functions, both of which are central components of CPT. However, given that QRD is technically a special case of CPT, our analysis highlights a simple approach to making prospect theory tractable in economic applications.

minimum outcome in the lottery's support. In addition, QRD is naturally extended to a three-factor model of behavioral biases that provides a formal relationship between three important principles of behavior – present bias, security bias, and positive skewness preference. These three biases interact to predict eight parallels between choices under risk and over time that violate the standard discounted expected utility model (as summarized in Table 3).

The three factor model organizes a larger set of the leading anomalies than alternative models. It also provides an alternative way to formalize loss aversion. Of course the three-factor model does not explain all prominent behavioral anomalies. For instance, QRD does not generate the 'lives-saved'/'lives-lost' framing effect of Tversky and Kahneman (1981), whereas this can be explained by prospect theory and salience theory (Bordalo et al., 2012b). Under additional assumptions, salience theory also can explain preference reversals between choice and pricing tasks (Lichtenstein and Slovic, 1971).

In the domain of intertemporal choice, the hyperbolic discounting model of Loewenstein and Prelec (1992) can explain present bias and the sign effect. The probability-time tradeoff model (Baucells and Heukamp, 2012) can explain the risk-time interaction effects, the certainty effect and present bias but does not explain other anomalies such as the sign effect or the bias toward concentration. Interestingly, two models which have been imported to a variety of economic applications – quasi-hyperbolic discounting (Laibson, 1997) and the focusing model of Koszegi and Szeidl (2013) each explain only one of the anomalies in Table 3. This demonstrates that data-fitting is not the only important ingredient of a successful decision model. Mathematical tractability is also important. Variations of the salience model of Bordalo et al. (2012b) have also been widely applied, although that model has not been extended to choices over time. The three-factor model proposed here has the appealing property that it accounts for many of the leading anomalies while being portable to economic applications. While it is true that model (30) has a large number of parameters, simpler versions of the model are sufficient to generate the main effects. For instance, quasi-hyperbolic discounting is a one-parameter extension of DU and the quasi-rank dependent model of security bias is a one-parameter extension of EU, and both are special cases of the three-factor model. In addition, we demonstrated that setting $\alpha = \beta = \gamma = 1$ in (30) still generates seven of the behavioral anomalies in Table 3 even though this specification has the same number of parameters as quasi-hyperbolic discounting.

XI. DISCUSSION

In this paper, we have motivated quasi-rank dependent (QRD) utility theory as a modeling tool for studying security biased agents in economic applications. The QRD model satisfies transitivity, stochastic dominance, and risk aversion, and is simpler than leading probability weighting models since it transforms individual rather than cumulative probabilities. The QRD model has an intuitive interpretation (i) as a pessimistic decision criterion that generalizes EU and Wald's maximin rule, and (ii) as a natural risky choice analog to quasi-hyperbolic discounting. The model also provides a unified explanation of the Allais paradox and loss aversion (aversion to small mixed gambles) as well as expressing a preference for robustness to model uncertainty when probabilities are subjective.

The QRD model was generalized in Section IX to a three-factor model of behavioral biases (with factors of positive skewness preference, present bias, and security bias) that unifies quasi-hyperbolic time discounting and quasi-rank dependent probability weighting. The three-factor model highlights three ways in which human decision making systematically deviates from maximizing discounted expected utility: by exhibiting a disproportionate preference for alternatives with higher minimum outcomes, higher maximum outcomes and higher immediate outcomes. We demonstrated that the three-factor model explains eight prominent behavioral anomalies and predicts close parallels between behaviors for decisions under risk and over time.

After describing the QRD model and the interpretations of security bias, we illustrated QRD in several economic applications. In the domain of insurance, we showed that the QRD model can explain the purchase of full insurance at actuarially unfair prices, which is not explained by EU. For prediction and betting markets, we demonstrated, somewhat surprisingly, that it is possible for a market to accurately aggregate traders' beliefs and generate *unbiased* prices, even if all traders are *systematically* biased. We also demonstrated that QRD predicts gaps between buying prices and selling prices and that security bias can contribute to efficiency loss and market failure. With the extended model in (10, 11), we generalized the analysis of Chateauneuf et al. (2007) from one to many risky assets through a strong correlation assumption. We derived equilibrium asset prices and found them to depend on the asset's fundamental value, a risk premium, a positive skewness premium, an ambiguity robustness premium, and the market's level of optimism. We also provided a simple calibration of the model with empirical data on the global market portfolio and found that it can explain the magnitude of the observed equity premium while generating a low risk-free rate.

We believe there are many novel questions and avenues for research that warrant further study. One promising research agenda may include work that addresses the basic question of 'when' do biases matter. These include questions like: What types of games or institutions are robust to systematically biased agents? Under what conditions do markets fail when traders are systematically biased? Under what conditions can biases produce Pareto efficient outcomes? Another promising line of research might address the question of 'how' do biases matter. How does behavior under different institutions change if agents are systematically biased? How does systematic bias affect the welfare of biased and unbiased agents?

A third line of research could apply QRD as a 'robustness check' on standard economic models. For instance, how do biases affect economic conclusions regarding behavior in games or markets with asymmetric information? We suspect in many cases, biases will result in simple generalizations that keep qualitative economic predictions intact as was the case for auctions in Section VI and for the aggregation of investor biases into market prices in Section V. But there may also be cases where biased agents produce qualitatively different economic outcomes than what EU predicts such as the optimal purchase of insurance at actuarially unfair prices in Section IV, gaps between buying prices and selling prices in equilibrium in Section V, and a large equity premium and the pricing of an asset's skewness and robustness to ambiguity in Section VIII. As the preceding questions illustrate, there seems to be much room for applying QRD to perform economic analysis with security biased agents.

APPENDIX: PROOFS OF PROPOSITIONS

Proposition 1: *If $u(-x) = -u(x)$ for all x, loss aversion holds if and only if $\theta \in [0,1)$.*

Proof: Under QRD, loss aversion implies: $(1-\theta/2)(u(-y)-u(-x)) > (\theta/2)(u(x)-u(y))$. If $u(-x) = -u(x)$ for all x, the above inequality holds if and only if $\theta \in [0,1)$. ■

Proposition 2: *The common ratio effect holds if and only if $\theta \in (0,1)$*

Proof: Note that for Definition 3, the QRD model implies:

$f \sim g$ *iff* $u(y) = \theta\big(pu(x) + (1-p)u(0)\big)$.

$f' \prec g'$ *iff* $\theta\big(qu(y) + (1-q)u(0)\big) < \theta\big(qpu(x) + (1-qp)u(0)\big)$.

Note $f' \prec g'$ holds if and only if $\theta u(y) < \theta\big(pu(x) + (1-p)u(0)\big) = u(y)$.

This inequality holds under QRD if and only if $\theta \in (0,1)$. ■

Proposition 3: *The Allais paradox holds if and only if $\theta \in (0,1)$.*

Proof: Note that for Definition 2, the QRD model implies:

$f \sim g$ *iff* $u(y) = \theta\big(pu(x) + (1-q)u(y) + (q-p)u(0)\big)$.

$f' \prec g'$ *iff* $\theta\big(qu(y) + (1-q)u(0)\big) < \theta\big(pu(x) + (1-p)u(0)\big)$.

Note that $f' \prec g'$ holds if and only if $\theta qu(y) < \theta\big(pu(x) + (q-p)u(0)\big)$. Adding $(1-q)\theta u(y)$ to both sides of this inequality yields:

$$\theta u(y) < \theta\big(pu(x) + (1-q)u(y) + (q-p)u(0)\big) = u(y).$$

Since $u(y) = 0$ if $\theta = 0$, this inequality holds if and only if $\theta \in (0,1)$. ■

Proposition 9: *Let $DU(f) = DU(g)$. Then present bias holds for all $\alpha, \beta, \theta \in (0,1)$.*

Proof: Let $\alpha, \beta, \theta \in (0,1)$. The value of each consumption plan is given by:

$$V(f) = \theta\big(DU(f)\big) + (1-\theta)(\alpha+\beta)y \text{ and } V(g) = \theta\big(DU(g)\big) + (1-\theta)\alpha x.$$

Note that $f \sim g$ and $DU(f) = DU(g)$ implies $(1-\theta)(\alpha+\beta)y = (1-\theta)\alpha x$. Next, note that $DU(f) = DU(g)$ implies $DU(f') = DU(g')$, and so $f' \prec g'$ if and only if $(1-\theta)\alpha y < (1-\theta)\alpha x$ which holds for all $\alpha, \beta, \theta \in (0,1)$, given that $(1-\theta)(\alpha+\beta)y = (1-\theta)\alpha x$. ■

Proposition 19: *Let $DEU(f) = DEU(g)$. Then time reduces risk aversion for all $\alpha, \beta, \theta \in (0,1)$.*

The proof of Proposition 10 is analogous to the proof of Proposition 9 and so is omitted.

Proposition 11: *Let $DEU(f) = DEU(g)$. Then risk reduces impatience for all $\alpha, \beta, \theta \in (0,1)$.*

The proof of Proposition 11 is analogous to the proof of Proposition 9 and so is omitted.

Proposition 12: *Let $u(x) = x$. Then the fourfold risk pattern holds for all $\alpha, \gamma, \theta \in (0,1)$.*

Proof: Under the three-factor model, for $x > 0$, we have $E(f) \succ f$ if $xp > \theta xp + (1-\theta)(\alpha x)$. For $\alpha < 1$, there is sufficiently large p such that this inequality holds. Similarly, we have $f \succ E(f)$ if $\theta xp + (1-\theta)(\alpha x) > xp$. For $\alpha > 0$, there is sufficiently low p such that this inequality holds. For $x < 0$, we have $E(f) \succ f$ if $xp > \theta xp + (1-\theta)(\gamma x)$. For $\gamma > 0$, there is sufficiently low p such that this inequality holds. Similarly, we have $f \succ E(f)$ if $\theta xp + (1-\theta)(\gamma x) > xp$. For $\gamma < 1$, there is sufficiently high p such that this inequality holds. ■

Proposition 13: *Let $u(x) = x$, and $\delta = 1$. Then a bias toward concentration holds if $n > 1 + \max(\beta/\alpha, \beta/\gamma)$, for any $\alpha, \beta, \gamma, \theta \in (0,1)$.*

Proof: Given $u(x) = x, \delta = 1$, and $\alpha, \beta, \theta \in (0,1)$, For $x > 0, V(A) = \theta nx + (1-\theta)(\alpha+\beta)nx$ and $V(B) = \theta nx + (1-\theta)\alpha x$. Thus, A $\succ$ B for $x > 0$. For $x < 0$, $V(A) = \theta nx + (1-\theta)(\beta+\gamma)nx$ and $V(B) = \theta nx + (1-\theta)\gamma x$. Thus, B $\succ$ A for $x < 0$ for all $n \geq 1$. Next, $V(A') = \theta nx + (1-\theta)\alpha nx$ and $V(B') = \theta nx + (1-\theta)(\alpha+\beta)x$ for $x > 0$. For $x < 0$, $V(A') = \theta nx + (1-\theta)\gamma nx$ and $V(B') = \theta nx + (1-\theta)(\beta+\gamma)x$. Then for all $n > 1 + \beta/\alpha$ A′ $\succ$ B′ for $x > 0$ and for all $n > 1 + \beta/\gamma$, B′ $\succ$ A′ for $x < 0$. ■

Proposition 14: *Let $u(x) = x$. Then for any $\theta \in (0,1)$, the sign effect holds if and only if $\alpha < \gamma$.*

Proof: We consider the case where $r > 0$. The case where $r = 0$ follows analogously. Under the three-factor model, $(x,r) \sim (y,t)$ if and only if $x\theta\delta^r - y\theta\delta^t = (y-x)(1-\theta)\alpha$. In addition, $(-x,r) \succ (-y,t)$ if and only if $x\theta\delta^r - y\theta\delta^t < (y-x)(1-\theta)\gamma$. By substitution, $(x,r) \sim (y,t)$ implies $(-x,r) \succ (-y,t)$ if and only if $\alpha < \gamma$. ■

Remark: As indicated in Footnote 8, the sign effect cannot be explained by a conventional behavioral approach that employs a loss aversion parameter in the agent's value function. To illustrate, suppose instead of security bias, that the agent has loss-averse preferences given by:

$$u(x) = \begin{cases} x, & x \geq 0 \\ \lambda x, & x < 0 \end{cases}$$

for $\lambda > 1$, but otherwise maximizes discounted utility. Then we have $(x,r) \sim (y,t)$ if and only if $x\delta^r = y\delta^t$. In addition, we have $(-x,r) \succ (-y,t)$ if and only if $-x\lambda\delta^r > -y\lambda\delta^t$ if and only if $x\delta^r < y\delta^t$ which contradicts $x\delta^r = y\delta^t$.

REFERENCES

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.

Andreoni, J. and C. Sprenger. 2010. "Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena." Manuscript.

Angeletos, G.-M., D. Laibson, A. Repetto, J. Tobacman, and S. Weinberg. 2001. "The hyperbolic consumption model: Calibration, simulation, and empirical evaluation." *The Journal of Economic Perspectives*, 3, 47-68.

Baillon, A., H. Bleichrodt, A. Cillo. 2014. "A Tailor-made Test of Intransitive Choice." *Operations Research*, 63, 198-211.

Barberis, N.C. 2013. "Thirty years of prospect theory in economics: A review and assessment." *The Journal of Economic Perspectives*, 27, 173-195.

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.

Blavatskyy, P. 2010. "Modifying the mean-variance approach to avoid violations of stochastic dominance." *Management Science*, 56, 2050-2057.

Blavatskyy, P. 2014. "A Theory of Decision-Making under Risk as a Tradeoff between Expected Utility, Expected Utility Deviation, and Expected Utility Skewness." Manuscript.

Bleichrodt, H., and U. Schmidt. 2002. "A context-dependent model of the gambling effect." *Management Science*, 48, 802-812.

Bleichrodt, H., L. Chen, I. Moscati, and P.P. Wakker. 2016. "Nash was a first to axiomatize expected utility." *Theory and Decision*, forthcoming.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2012a. "Salience and experimental tests of the endowment effect." *American Economic Review*, 102, 47-52.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2012b. "Salience Theory of Choice under Risk." *Quarterly Journal of Economics*, 127, 1243-1285.

Botzen, W.J., J.C. Aerts, and J.C. van den Bergh. 2013. "Individual Preferences for Reducing

Flood Risk to Near Zero through Elevation." *Mitig. Adapt. Strateg. Glob. Change.* 18, 229-244.

Brown, J.R., M.D. Casey, and O.S. Mitchell. 2007. "Who values the social security annuity? New evidence on the annuity puzzle." NBER Retirement Research Center Paper NB 07-02.

Chateauneuf, A., J. Eichberger, and S. Grant. 2007. "Choice under uncertainty with the best and worst in mind: Neo-additive capacities." *Journal of Economic Theory*, 137, 538-567.

Cohen, M. 1992. "Security level, potential level, expected utility: a three-criteria decision model under risk." *Theory and Decision*, 33, 101-134.

Dhami, S. 2016. *Foundations of Behavioral Economic Analysis*. Oxford University Press

Diamond, P., and B. Koszegi. 2003. "Quasi-hyperbolic discounting and retirement." *Journal of Public Economics* 87, 1839-1872.

Diecidue, E., and P.P. Wakker. 2001. "On the intuition of rank-dependent utility." *Journal of Risk and Uncertainty*, 23, 281-298.

Diecidue, E., U. Schmidt, and P.P. Wakker. 2004. "The Utility of Gambling Reconsidered," *Journal of Risk and Uncertainty*, 29, 241-259.

Doeswijk, R., T. Lam, and L. Swinkels. 2017. "Historical returns of the market portfolio." Manuscript.

Dow, J. and S.R. Werlang. 1992. "Uncertainty aversion, risk aversion, and the optimal choice of portfolio." *Econometrica*, 60, 197-204.

Edwards, W. 1954. "The theory of decision making." *Psychological Bulletin*, 51, 380-417.

Fama, E.F., and K.R. French. 1992. "The Cross-Section of Expected Stock Returns". *Journal of Finance*, 47, 427-465.

Fan, C-P. 2002. "Allais paradox in the small." *Journal of Economic Behavior and Organization*, 49, 411-421.

Fox, C.R., C. Erner, and D.J. Walters. 2015. "Decision under Risk: From the Field to the Lab and Back." Chapter in G. Keren and G. Wu (Eds.), *Handbook of Judgment and Decision Making*.

Friedman, M. and L.J. Savage. 1948. "The utility analysis of choices involving risk." *The Journal of Political Economy*, 56, 279-304.

Gilboa, I. 1988. "A combination of expected utility and maxmin decision criteria." *Journal of Mathematical Psychology*, 32, 405-420.

Gjerstad, S. 2005. "Risk aversion, Beliefs, and Prediction Market Equilibrium." Manuscript.

Halevy, Y. 2008. "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect."

American Economic Review. 98, 1145-1162.

Huck, S. and W. Muller. 2012. “Allais for all: Revisiting the paradox in a large representative sample.” *Journal of Risk and Uncertainty*, 44, 261-293.

Hurwicz, L. 1951. “A Class of Criteria for Decision-Making under Ignorance,” *Cowles Commission Discussion Paper*, Statistics No. 356.

Incekara-Hafalir, E. and J.D. Stecher. 2016. “A Zero Effect explains deviations from Expected Utility in Allais-type tasks.” Manuscript, Carnegie Mellon University.

Jaffray, J.-Y. 1988. “Choice under risk and the security factor: an axiomatic model.” *Theory and Decision*, 24, 169-200.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York. Farrar, Straus, Giroux.

Kahneman, D. and A. Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica*, 47, 263-292.

Karmarkar, U.S. 1978. “Subjectively weighted utility: A descriptive extension of the expected utility model.” *Organizational Behavior and Human Performance*, 21, 61-72.

Karni, E. and Z. Safra. 1989. “Dynamic Consistency, Revelations in Auctions, and the Structure of Preferences.” *Review of Economic Studies*, 56, 421-434.

Keren, G. and P. Roelofsma. 1995. “Immediacy and Certainty in Intertemporal Choice.” *Organizational Behavior and Human Decision Processes,* 63, 287-297.

Khwaja, A., D. Silverman, and F. Sloan. 2007. “Time preference, time discounting and smoking decisions.” *Journal of Health Economics*, 26, 927-949.

Kim, W.C., J.H. Kim, J.M. Mulvey, and F.J. Fabozzi. 2015. “Focusing on the worst state for robust investing.” *International Review of Financial Analysis*, 39, 19-31.

Koszegi, B., M. Rabin. 2006. “A model of reference-dependent preferences.” *Quarterly Journal of Economics*, 121, 1133-1165.

Koszegi, B. M. Rabin. 2007. “Reference-dependent risk attitudes.” *American Economic Review*, 97, 1047-1073.

Koszegi, B., A. Szeidl. 2013. “A model of focusing in economic choice.” *Quarterly Journal of Economics*, 128, 53-104.

Kodritsch, S. 2014. “On time-inconsistency in bargaining.” Manuscript.

Kopylov, I. 2009. “Choice deferral and ambiguity aversion.” *Theoretical Economics*, 4, 199-225.

Laibson, D. 1997: “Golden Eggs and Hyperbolic Discounting,” *The Quarterly Journal of*

Economics, 112, 443-477.

Laibson, D., A. Repetto, J. Tobacman, R.E. Hall, W.G. Gale, and G.A. Akerlof. 1998. "Self-control and saving for retirement." *Brookings papers on economic activity*, 91-196.

Levin, J. 2006. *Choice under Uncertainty.* Manuscript.

Levine, D.K. 2012. *Is Behavioral Economics Doomed? The Ordinary versus the Extraordinary*. Open Book Publishers, United Kingdom.

Lichtenstein, S., and P. Slovic. 1971. "Reversals of preference between bids and choices in gambling decisions." *Journal of Experimental Psychology*, 89, 46-55.

Loewenstein, G., and D. Prelec. 1992. "Anomalies in Intertemporal Choice: Evidence and Interpretation." *Quarterly Journal of Economics*, 107, 573-597.

Lopes, L.L. 1986. "What naïve decision makers can tell us about risk." In: Daboni, Luciano, Montesano, Aldo, Lines, Marji (Eds.), *Recent Developments in the Foundations of Utility and Risk Theory.* Reidel, Netherlands, 311-326.

Machina, M.J. 1989. "Dynamic consistency and non-expected utility models of choice under uncertainty." *Journal of Economic Literature*, 27, 1622-1668.

Manski, C.F. 2006. "Interpreting the predictions of prediction markets." *Economics Letters*, 91, 425-429.

Markowitz, H.M. 1959. "*Portfolio selection: Efficient diversification of Investments."* Yale.

Marschak, J. 1950. "Rational behavior, uncertain prospects, and measurable utility." *Econometrica*, 18, 111-141.

Mehra, R., and E.C. Prescott. 1985. The equity premium: A puzzle. *Journal of Monetary Economics*, 15, 145-161.

Meier, S., and C. Sprenger. 2010. "Present-biased preferences and credit card borrowing." *American Economic Journal: Applied Economics*, 2, 193-210.

Nash, J.F. 1950a. "Equilibrium points in n-person games." *Proceedings of the National Academy of Sciences, USA*, 36, 48-49.

Nash, J.F. 1950b. "The bargaining problem." *Econometrica, 18,* 155-162.

Nash, J.F. 1951. "Non-cooperative games." *Annals of Mathematics*, 54, 286-295.

Neilson, W.S. 1992. "Some Mixed Results on Boundary Effects," *Economics Letters*, 39, 275-278

Prelec, D. and G. Loewenstein. 1991. "Decision Making Over Time and Under Uncertainty: A Common Approach." *Management Science*, 37, 770-786.

Prelec, D. 2006. "Rebuilding the Boat while Staying Afloat: The Modeling Challenge for Behavioral Economics." *Journal of Marketing Research*, 43, 332-336.

Quiggin, J. 1982. "A theory of anticipated utility." *Journal of Economic Behavior & Organization*, 3, 323-343.

Regenwetter, M., J. Dana, and C.P. Davis-Stober. 2011. "Transitivity of preferences." *Psychological Review*, 118, 42-56.

Schmidt, U. 1998. "A Measurement of the Certainty Effect," *Journal of Mathematical Psychology*, 42, 32-47.

Schmidt, U. 2000. "The certainty effect and boundary effects with transformed probabilities." *Economics Letters*, 29-33.

Schmidt, U. and H. Zank. 2005. "What is Loss Aversion?" *Journal of Risk and Uncertainty*, 30, 157-167.

Sydnor, J. 2010. (Over)insuring modest risks. *American Economic Journal: Applied Economics*, 2, 177-199.

Thaler, R. 2015. *Misbehaving: The Making of Behavioral Economics*. Norton.

Tobin, J. 1958. "Liquidity preference as behavior towards risk." *The Review of Economic Studies*, 25, 65-86.

Tversky, A. and D. Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science*, 211, 453-458.

Tversky, A. and D. Kahneman. 1992. "Advances in Prospect Theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, 5, 297-323.

Viscusi, W.K., J. Huber, and J. Bell. 2014. "Assessing whether there is a cancer premium for the value of a statistical life." *Health Economics*, 23, 384-396.

Wald, A. 1950. *Statistical decision functions.* Wiley, New York.

Webb, C.S., and H. Zank. 2011. "Accounting for optimism and pessimism in expected utility." *Journal of Mathematical Economics*, 47, 706-717.

Wolfers, J. and E. Zitzewitz. 2006. "Interpreting prediction market prices as probabilities." IZA Discussion Papers, No. 2092.