Individual Preferences, Monetary Gambles and the Equity Premium

Nicholas Barberis, Ming Huang, Richard Thaler
University of Chicago, Stanford University and University of Chicago

October 2002

Abstract

Many different preference specifications have been proposed as a way of addressing the equity premium. How should we pick between them? We suggest one possible metric, namely these utility functions' ability to explain *other* evidence on attitudes to risk. We consider some simple observations about attitudes to monetary gambles with just two outcomes and show that the vast majority of utility functions used in asset pricing have difficulty explaining these observations. However, utility functions with two features – first-order risk aversion and narrow framing – can easily explain them. We argue that by this metric at least, such utility functions may be very attractive to financial economists: they can generate substantial equity premia and at the same time, make sensible predictions about attitudes to monetary gambles.

1 Introduction

In their efforts to understand the historical equity premium, economists have made use of a large range of preference specifications. Power utility functions, preferences with external or internal habit, recursive utility and loss aversion preferences are just a few examples. In principle, all of these functional forms can match the historical premium for *some* choice of parameter values. How, then, can we pick between them? Which one offers the most attractive way of addressing the equity premium, or of capturing people's attitudes to stock market risk?

In the past, economists have typically judged these different preference specifications on their ability to match other variables of interest: the risk-free rate, say, or the volatility of returns. For example, the fact that when calibrated to match the historical premium, power utility preferences often also predict a counterfactually high risk-free rate, has been used as an argument against such preferences.

In this paper we suggest an additional metric for assessing different preference specifications, namely their ability to explain attitudes not only to stock market risk, but also to *other* simple risks. To this end, we start with some evidence on how people react to the simplest imaginable form of risk – monetary gambles with just two possible outcomes, a gain and a loss, with stakes both large and small – and ask what kinds of utility functions are able to explain these observations. In particular, we are interested in the fact that people often turn down small-stakes gambles whose outcomes are independent of other risks, even if they are actuarially favorable: for example, most people turn down a 50:50 bet offering a \$110 gain against a \$100 loss.

Surprisingly, we find that a wide range of utility functions, including the vast majority of those used in asset pricing, have great difficulty capturing these attitudes to monetary gambles. The only utility functions that *can* address these observations with any degree of satisfaction are those exhibiting both *first-order risk aversion* and *narrow framing*. First-order risk aversion means that a utility function is locally risk averse, unlike many standard preferences that are smooth and therefore locally risk-neutral; a utility function with a kink at the agent's current wealth, for example, would exhibit this feature. Narrow framing means that to some extent, the agent evaluates new gambles she is offered in isolation from other risks that she faces. More formally, her utility function depends on the outcomes of specific gambles she faces over and above what those outcomes mean for her aggregate wealth risk.

To see why these two features are necessary, consider first a utility function without first-order risk aversion, in other words, one which is locally smooth. Since the agent is locally risk-neutral, she will normally be very happy to accept a small, actuarially attractive gamble like the coin flip bet to win \$110 or lose \$100. In order to explain the commonly observed rejection of such gambles, then, we need to push risk aversion up to very high levels. However, risk aversion will then be *so* high as to make the agent reject some apparently very favorable gambles with larger stakes. To avoid such counterfactual predictions, we need to choose utility functions that are locally risk averse, not locally risk-neutral; in other words, utility functions exhibiting first-order risk aversion.

This argument for first-order risk aversion has already appeared in various guises in the literature. Our more novel contribution is to show that to explain attitudes to monetary gambles and in particular, aversion to small gambles, we need not only first-order risk aversion but narrow framing as well. The intuition for why first-order risk aversion is not enough is straightforward. Suppose that an investor with first-order risk aversion is offered a small, independent gamble to be resolved at some point in the future, and that the gamble is actuarially attractive. Now also make the reasonable assumption that the investor faces some pre-existing risks: labor income risk perhaps, house price risk, or other financial market risk. The investor will decide whether to take on the new gamble by merging it with her pre-existing risk and checking to see if the combination is attractive. It turns out that the combination *is* almost always attractive: since the new gamble is independent of the agent's other risks, it brings her useful diversification benefits. Even though she is first-order risk averse, she happily accepts it. The only way to get the agent to reject the gamble is, once again, to set risk aversion to extraordinarily high levels. However, this again implies, counterfactually, that the agent would reject some attractive gambles with larger stakes.

In order to explain the commonly observed aversion to a small gamble, then, it must be that the investor does not fully merge it with pre-existing risks, but that to some extent, she evaluates it in isolation. Put differently, her decision utility must depend on the outcome of the gamble over and above what that outcome means for the risk exposure of his aggregate wealth; more simply, her utility function must exhibit narrow framing.

In summary, then, we show that the vast majority of utility functions used to address the equity premium have great difficulty explaining typical attitudes to the simplest imaginable form of risk, namely monetary gambles with two outcomes. In our view, this is an important metric by which to assess utility functions and one that points to preference specifications based on first-order risk aversion and narrow framing as being attractive ways of modelling the equity premium: such utility functions can capture aversion to stock market risk *and* make sensible predictions about attitudes to other simple risks.

Towards the end of the paper, we respond to a possible critique of our approach, namely that in discrediting a vast array of standard utility functions, we appear to rely heavily on the widely observed aversion to small gambles like the 50:50 bet over a \$110 gain or \$100 loss. To some people, such gambles might appear unimportant: it is enough, they might say, for utility functions to capture attitudes to large gambles. We disagree with this point of view, and argue that for financial economists, small gambles *are* important, because stocks are, in many ways, a small gamble: for the typical investor, stock market holdings constitute only a small fraction of total wealth. We also provide analytical evidence to support this view: we take two utility specifications, calibrate them so that they deliver similar attitudes to large gambles but very different attitudes to small gambles and show that the preferences exhibiting greater aversion to small gambles also deliver a much higher equity premium. In short, attitudes to stocks appear to be closely tied to attitudes to small gambles, making it important to find preferences that can explain the latter.

Our research builds on earlier work investigating how, in static settings, individuals with various preference specifications react to monetary gambles of different sizes. Kandel and Stambaugh (1990) point out that power utility functions have trouble simultaneously explaining attitudes to both large and small scale gambles, while Rabin (2000) shows that this problem extends to all expected utility functions. One contribution of our research is to check whether these static arguments apply to the intertemporal setting used by financial economists. Another is to show that even more general types of preferences, including those exhibiting first-order risk aversion, also have difficulty capturing attitudes to monetary gambles.

Epstein and Zin (1990) and Epstein (1992) argue that stocks are effectively a small risk, making it important that the preferences we use to model attitudes to stocks should do a good job capturing attitudes to small gambles. They use this reasoning as a way of motivating an investigation of first-order risk averse preferences. We agree with this line of thinking, but show that to capture attitudes to small gambles successfully, both first-order risk aversion *and* narrow framing are required, not first-order risk aversion alone.

In Section 2, we discuss common attitudes to simple monetary gambles and introduce various classes of utility functions whose ability to match those attitudes we are interested in. In Section 3, we show that without first-order risk aversion, it is hard to match these attitudes. In Section 4, we show that even first-order risk aversion is not enough and that narrow framing is required as well. In Section 5 we discuss our apparent over-reliance on the rejection of small gambles as well as possible reasons for why people engage in narrow framing in the first place. Section 6 concludes.

2 Attitudes to Monetary Gambles

Consider the small gamble

$$\underline{G}_S = (110, \frac{1}{2}; -100, \frac{1}{2}),$$

which we read as "get \$110 with probability $\frac{1}{2}$ and $-\$100$ with probability $\frac{1}{2}$". It is the premise of this paper that most people find this gamble and others like it unattractive. We base this premise on hundreds of experimental studies, some with real money and some based on hypothetical questions. Since economists are often skeptical of answers to hypothetical questions, and since experiments are rarely done with real money, we conducted a small experiment to add further support to our premise.

The subjects in our experiment are a group of part-time MBA students. They were asked to fill out a short survey that included one real money question, namely whether they would play the \$110/100 gamble above. They were told that if they wished to accept this gamble, they should indicate so on the experimental form, and then come to class the following week with the \$100 they would need to pay in case they lost the gamble. They were informed that if they won they would be paid immediately in cash. Of the 41 students that participated in the experiment, only 4 were willing to accept the gamble (9.75%).

Aversion to small losses is not confined to the laboratory, but occurs in numerous real-world settings as well. One example is automobile collision insurance, where people seem to fear small losses so much that they request remarkably low deductibles. Grgeta and Thaler (2002) find that in the 1994-96 period, more than half the purchasers of such insurance elected a deductible of \$250 or less. For the typical consumer, increasing the deductible from \$250 to \$500 would save about \$80 a year. To justify the lower deductible people would have to file claims one year in three, but in fact the probability of a claim is less than one in ten.[1]

In Sections 3 and 4, we investigate what kinds of utility functions can capture the commonly observed aversion to gambles like $\underline{G}_S$. Of course, many utility functions can explain this evidence simply by assuming sufficiently high risk aversion. To provide a reasonable upper bound on individual risk aversion, we introduce two new gambles; the first is again a small stakes bet, the second involves larger stakes:[2]

$$\begin{aligned} \overline{G}_S &= (400, \frac{1}{2}; -100, \frac{1}{2}) \\ G_L &= (20,000,000, \frac{1}{2}; -10,000, \frac{1}{2}). \end{aligned}$$

It is the premise of this paper that these bets are typically *accepted*. To provide some evidence for this, we presented these bets to the same subjects that participated in our real money experiment, though budget constraints prevented us from offering to play these gambles out. Instead, we simply asked the subjects to think hard about how they would actually choose. We found that these gambles were indeed much more attractive: 29 of the 41 subjects were willing to accept $\overline{G}_S$, and 30 were willing to take G_L.[3]

In summary, then, we are interested in what kinds of preference specifications can explain the three observations,

---

[1]Even though such real-world examples are very suggestive, it is hard to know whether they are driven by aspects of individual preferences, or by irrational beliefs: in our insurance example, it may simply be that people are mis-estimating the probability of a claim. We therefore focus primarily on trying to make sense of the laboratory evidence, where probabilities are known.

[2]The S subscript on $\underline{G}_S$ and $\overline{G}_S$ stands for "small," and the L subscript on G_L for "large." The underline in $\underline{G}_S$ denotes an unfavorable small gamble, and the overline in $\overline{G}_S$, a favorable small gamble.

[3]Typically, to get a majority of subjects to accept a 50-50 bet at moderate stakes requires that the gain be more than twice the loss.

I. $\underline{G}_S$ is rejected

II. $\overline{G}_S$ is accepted

III. G_L is accepted.

We do not insist that a utility function be able to explain these observations at *all* wealth levels. Rather, we make the weaker demand that they explain them over a range of reasonable wealth levels – neither too high nor too low. To be precise, we check observation I for wealth levels below $\$100,000$, observation II for wealth levels above $\$1,000$ and observation III for wealth levels above $\$50,000$.

When we check utility functions' ability to explain these observations, it can make a difference, for certain utility specifications, whether the three gambles are "immediate" or "delayed." A gamble is immediate if its uncertainty is resolved at once, before any further consumption decisions are made. A delayed gamble, on the other hand, might be played out as follows: in the case of $\underline{G}_S$, the subject is told that at some point in the next few months, she will be contacted and informed either that she has just won \$110 or that she has lost \$100, the two outcomes being equally probable and independent of other risks.

Although certain utility functions make a distinction between immediate and delayed gambles, we think that in reality, people do not treat the two kinds of bets very differently. To test this intuition we asked our sample of students one additional hypothetical question, which was whether they would accept $\underline{G}_S$, the \$110/100 gamble, if it were played out on a day picked at random during 2003. (The survey was conducted in October 2002.) The subjects largely shared our intuition: only 9 of the 41 subjects were willing to accept the delayed gamble.[4]

In view of this evidence, we insist that the preference specifications we consider be able to capture observations I-III in either case, immediate or delayed. For the first part of our analysis, we will only need to work with the computationally simpler immediate gambles: it turns out that many classes of utility functions have trouble explaining attitudes to immediate gambles alone. In cases where utility functions *are* able to capture attitudes to immediate gambles, we challenge them with delayed gambles as well.

2.1 Utility Functions

We now introduce the different classes of preferences whose ability to capture observations I-III we are interested in. We list them in increasing order of sophistication, along with the abbreviations we use to refer to them.

[4]We are not sure why people are less averse to the delayed version of $\underline{G}_S$. Potential explanations include the fact that this question was hypothetical, producing somewhat more courage, and the possibility that some of the students may be graduating and anticipating more income and fewer tuition payments.

Expected utility preferences [EU]
Non-expected utility preferences:
  Recursive utility with EU certainty equivalent [R-EU]
  Recursive utility with non-EU, second-order risk averse certainty equivalent [R-SORA]
  Recursive utility with non-EU, first-order risk averse certainty equivalent [R-FORA]

Expected utility preferences are familiar enough. What about non-expected utility specifications? In an intertemporal setting, non-expected utility is typically implemented via a recursive structure in which time t utility, V_t, is defined through

$$V_t = W(C_t, \mu(\widetilde{V}_{t+1}|I_t)). \tag{1}$$

Here $\mu(\widetilde{V}_{t+1}|I_t)$ is the certainty equivalent of the distribution of future utility $\widetilde{V}_{t+1}$ conditional on time t information, and W is an "aggregator function" which aggregates current consumption C_t with a summary of future utility to give current utility.

We consider three kinds of recursive utility. They differ in the properties they impose on μ. One property that plays an important role is the *order* of risk aversion built into μ, and in particular whether μ exhibits "second-order risk aversion" or "first-order risk aversion," terms originally coined by Segal and Spivak (1990). An agent's utility function exhibits second-order risk aversion if the premium the agent pays to avoid an actuarially fair gamble $k\widetilde{\varepsilon}$ is, as $k \to 0$, proportional to k^2. In simple terms, such utility functions are smooth and the investor is almost risk-neutral for small risks. First-order risk averse utility functions, on the other hand, are preferences where the premium paid to avoid an actuarially fair gamble $k\widetilde{\varepsilon}$ is, as $k \to 0$, proportional to k. In this case, the investor is risk averse even over infinitesimal bets. A simple example of a utility function with this property is one exhibiting loss aversion, or a kink at the agent's current wealth.

Utility functions in the expected utility class can only exhibit second-order risk aversion. This is because, as Epstein (1992) points out, an increasing, concave utility function can only have a kink at a countable number of points. Non-expected utility functions, on the other hand, can exhibit either second-order risk aversion or first-order risk aversion, and it is important to consider the two cases separately.

We now describe the three kinds of recursive utility in more detail. First, we look at recursive utility preferences in which the certainty equivalent function μ has the expected utility form,

$$\mu(\widetilde{X}) = h^{-1}Eh(\widetilde{X}).$$

As noted above, we denote preferences in this class as R-EU.[5] Almost all implementations of recursive utility that have appeared in the asset pricing literature, including those of Epstein

[5]Note that even though μ is in the expected utility class, intertemporal utility V_t is still non-expected utility.

and Zin (1991a), Campbell (1996) and Campbell and Viceira (1999), are of the R-EU form. Researchers have made use of such preferences primarily because they offer a simple way of separating risk aversion and intertemporal elasticity of substitution, something which cannot be done satisfactorily within the expected utility class (Epstein, 1992).

Next we consider recursive utility in which μ is in the non-expected utility class and exhibits second-order risk aversion (R-SORA). Such preferences appear more rarely in the asset pricing literature: since the main objective of using recursive utility – the separation of risk aversion and intertemporal substitution – can be accomplished with the simpler R-EU preferences, fewer studies adopt R-SORA preferences. One exception is Epstein and Zin (1991b).

Finally, we consider recursive utility in which μ is again non-expected utility, but now exhibits first-order risk aversion (R-FORA). Such preferences have again only rarely appeared in the asset pricing literature, although they have been studied by Epstein and Zin (1990) and Bekaert, Hodrick and Marshall (1997) among others.

In Section 3, we show that utility functions without first-order risk aversion – in other words, the EU, R-EU and R-SORA classes – have difficulty explaining the attitudes to monetary gambles listed in observations I-III. In Section 4, we show that even utility functions with first-order risk aversion, namely those in the R-FORA class, have a hard time explaining the observations, and that a second ingredient, narrow framing, is required.

3 The Importance of First-order Risk Aversion

3.1 Expected Utility

In the expected utility framework, preferences are generally defined over an intertemporal consumption stream,

$$E(U(\tilde{C}_0, \tilde{C}_1, \ldots, \tilde{C}_T)). \tag{2}$$

Under mild conditions, one can show that optimizing expected utility over consumption leads to an indirect value function over wealth,

$$J(W_t; I_t, C_{-t}) = \max\ E_t(U(C_0, \ldots, C_t, \tilde{C}_{t+1}, \ldots, \tilde{C}_T)), \tag{3}$$

where $C_{-t} \equiv \{C_0, C_2, \cdots, C_{t-1}\}$ denotes the individual's past consumption history and I_t denotes information available at time t about the state of the economy. We make the reasonable assumption that the outcomes of our monetary gambles do not affect I_t and are independent of all other economic uncertainty.

We now ask whether the expected utility preferences in (2) can explain the attitudes to large and small gambles listed in observations I-III. The following proposition establishes that *no* utility function in this class can do so.

Proposition 1.

(a) Consider an individual with an expected utility preference in which future utility does not depend on past consumption, so that her value function is $J(W_t; I_t)$*. Suppose that for given* I_t*, she rejects* $\underline{G}_S$ *at wealth levels below* \$100,000*. Then she rejects* G_L *at all wealth levels.*

(b) Consider an individual with an expected utility preference in which future utility does depend on past consumption, so that her value function is $J(W_t; I_t, C_{-t})$*. Suppose that for given* I_t and C_{-t}*, she rejects* $\underline{G}_S$ *at wealth levels below* \$100,000*. Then she rejects* G_L *at all wealth levels.*

Proof: See Appendix.

In words, the proposition says that any utility function able to explain observation I – the rejection of $\underline{G}_S$, the \$110/100 bet – will inevitably fail to explain observation III, namely the acceptance of G_L, the $\$20,000,000/10,000$ bet.

The proposition covers a wide range of utility specifications, including most of those used in asset pricing. Part (a) of the proposition includes time-separable/state-independent utility of the form

$$U(C_0, \ldots, C_T) = \sum_{t=0}^{\infty} u_t(C_t)$$

as well as external habit dependence (Abel 1990, Campbell and Cochrane 1999). Part (b) covers internal habit dependence (Constantinides 1990, Sundaresan 1989).

Proposition 1 can be thought of as an intertemporal generalization of a recent result of Rabin (2000), who shows that in a static one-period setting, no EU specification with an increasing concave utility function can explain both observations I and III. The intuition for Rabin's finding, and hence also for Proposition 1, is straightforward. An individual with the preferences in Proposition 1 is locally risk-neutral; since gamble $\underline{G}_S$ involves small stakes, she would normally take it without hesitating. To get her to reject it, in accordance with observation I, we need to make her locally risk averse. Moreover, since she must reject $\underline{G}_S$ over a wide range of wealth levels, she must also be locally risk averse over a wide range of wealths. Proposition 1 simply states that this immediately implies a level of global risk aversion so high that she even rejects the apparently favorable large gamble, G_L.

The proof of Proposition 1 depends crucially on a property of the expected utility preferences in (2) and (3), namely that for fixed I_t and C_{-t}, the utility difference between two wealth levels does not depend on current wealth: the increase in utility from having $\$21,000$ rather than $\$20,000$ is the same, whether current wealth is $\$10,000$ or $\$20,000$. Therefore, knowing that someone will turn down a small gamble like $\underline{G}_S$ at a wealth level of $\$20,000$

provides valuable information about how, at a wealth level of $\$10,000$, she would react to a large risk like G_L that might bring her into the neighbourhood of $\$20,000$.

At first sight, it might seem from Proposition 1 that Rabin's (2000) argument transfers easily to the intertemporal setting. However, this is not completely true. The argument works much better for certain types of utility functions than for others. As is reasonable in a one-period context, Rabin (2000) considers utility functions that are defined over wealth alone. In an intertemporal setting, value functions often depend not only on wealth but, as shown in (3), on state variables I_t and past consumption C_{-t} as well. In order to apply Rabin's argument, then, we need the assumption given in each part of the proposition, namely that *keeping these other variables fixed*, $\underline{G}_S$ is rejected at a range of wealth levels. The problem is that this assumption may sometimes be difficult to verify.

Consider an individual with internal habit preferences, covered in part (b) of the proposition. There, we assume that for fixed C_{-t}, the investor rejects $\underline{G}_S$ at a range of wealth levels. To provide evidence that this assumption actually holds, we would want to ask people with different wealth, but the *same* past consumption, how they feel about $\underline{G}_S$. The problem is now clear: it is very hard to find a group of subjects to do this experiment on, because people with different wealth will also tend to have different past consumption. Since it is difficult to show that the premise of the proposition is true for internal habit preferences, using the proposition to dismiss such preferences may be too harsh. This caveat does not let habit-based preferences off the hook though, because they are still subject to more general criticisms that we make later of all utility functions displaying second-order risk aversion, whether expected utility or non-expected utility.[6]

Initial indications of the problem with EU preferences appear in Kandel and Stambaugh (1990), who show that in a one-period setting, power utility preferences

$$U(W) = \frac{W^{1-\gamma}}{1-\gamma}$$

have trouble simultaneously capturing attitudes to both small- and large-scale risks. Whatever value of γ is chosen, Kandel and Stambaugh (1990) show that the resulting preferences make counterintuitive predictions either about large-scale or about small-scale wealth gambles. Rabin (2000) and Proposition 1 above show that this problem arises not only for power utility functions but for all expected utility specifications: if they are calibrated to fit attitudes to small-scale gambles, they will be unable to fit attitudes to large-scale gambles.

[6]Rubinstein (2001) points out that Rabin's (2000) argument applies only when utility is defined over wealth, not when it is defined over wealth changes, say. In general, this critique is not relevant to our analysis. Financial economists define utility over consumption streams and as discussed in the main text, such utility functions lead quite generally to value functions defined over wealth, not changes in wealth. However, there is a sense in which the difficulty we raise in the case of internal habit preferences is very similar to the difficulty raised by Rubinstein (2001). In the case of internal habit, the value function $J(W_t; I_t, C_{-t})$ comes close to being a function of wealth changes, since past consumption C_{-t} is likely to be closely related to past wealth.

Example

We illustrate the proposition with a simple example. Consider an investor with power utility preferences

$$U(C_0, \ldots, C_T) = \sum_{t=0}^{\infty} \rho^t \frac{C_t^{1-\gamma}}{1-\gamma} \tag{4}$$

and with i.i.d. investment opportunities. The investor's value function is then given by

$$J(W_t) = \Gamma \frac{W_t^{1-\gamma}}{1-\gamma} \tag{5}$$

for some constant Γ (Ingersoll, 1982).

We now check that any γ able to explain observation I will be unable to explain observation III. To see this, note that the investor rejects an immediate gamble $\tilde{v}$ iff

$$E(J(W_t + \tilde{v})) < J(W_t).$$

For $\gamma > 1$, and $\tilde{v} = (x, \frac{1}{2}; -y, \frac{1}{2})$, this reduces to

$$(W_t + x)^{1-\gamma} + (W_t - y)^{1-\gamma} > 2W_t^{1-\gamma}. \tag{6}$$

Suppose that the investor's preferences fit observation I, so that she rejects $\underline{G}_S$, the \$110/100 gamble, at any wealth level below \$100,000. Then, in particular, she rejects $\underline{G}_S$ at a wealth level *equal* to \$100,000. A simple computation shows that the lowest integer value of risk aversion γ that satisfies (6) for

$$W_t = 100,000,\ x = 110,\ y = 100$$

is $\gamma = 90$. But at this level of risk aversion, the investor rejects G_L, the \$20,000,000/10,000 bet: for $\gamma \geq 90$, inequality (6) is violated when

$$x = 20,000,000,\ y = 10,000,$$

whatever the investor's initial wealth. In other words, if the investor's risk aversion γ is high enough to explain the rejection of $\underline{G}_S$ in observation I, it will be so high as to make it impossible to explain observation III.

3.2 Non-expected Utility

Having shown in Proposition 1 that EU preferences are unable to explain observations I-III, we turn to non-expected utility specifications.

Recursive utility with expected utility certainty equivalent [R-EU]

We begin with the following proposition, which shows that the first type of recursive utility preference – R-EU – *cannot* explain observations I and III.

Proposition 2. Suppose that an individual has the recursive utility preferences

$$V_t = W(C_t, \mu(\widetilde{V}_{t+1}|I_t))$$

with $\mu(\widetilde{V}_{t+1}|I_t)$ the certainty equivalent of the distribution of future utility $\widetilde{V}_{t+1}$ conditional upon information at time t, and where μ has the expected utility form

$$\mu(\widetilde{X}) = h^{-1}Eh(\widetilde{X})$$

for some increasing, concave h, so that the value function is

$$J(W_t; I_t) = \max\, W(C_t, \mu(\widetilde{V}_{t+1}|I_t)).$$

Suppose that for given I_t, the individual rejects $\underline{G}_S$ at wealth levels below \$100,000. Then she rejects G_L at all wealth levels.

Proof: See Appendix.

In words, the proposition says that if an R-EU preference specification is calibrated to match observation I – the rejection of $\underline{G}_S$, the \$110/100 bet – it fails to match observation III, in that it predicts the rejection of G_L, the $\$20,000,000/10,000$ bet. The idea behind the proof is straightforward. Even though the preferences in Proposition 2 are non-expected utility, attitudes to risk are governed by the certainty equivalent functional μ, which *is* in the expected utility class. Therefore just as expected utility functions cannot explain observations I and III – our result in Proposition 1 – so recursive utility with an expected utility functional μ cannot explain them either.

Example

To illustrate Proposition 2, consider an investor with the following preferences, which belong to the class studied in the proposition,

$$\begin{aligned} W(C,\mu) &= ((1-\beta)C^{\rho} + \beta\mu^{\rho})^{\frac{1}{\rho}},\ \rho < 1,\ \rho \neq 0 \qquad (7) \\ \mu(\widetilde{V}) &= (E(\widetilde{V}^{1-\gamma}))^{\frac{1}{1-\gamma}}, \qquad (8) \end{aligned}$$

and with i.i.d. investment opportunities. Epstein and Zin (1989) show that in this case, the investor's value function takes the form

$$J(W_t) = \Gamma W_t \qquad (9)$$

for some constant Γ.

We now check that any calibration of (7) able to explain observation I is unable to explain observation III. An investor with recursive utility evaluates an immediate gamble $\widetilde{v}$ in the following way. He inserts an infinitesimal time step Δt at time t and checks whether the utility from taking the gamble,

$$W(0,\mu(J(W_{t+\Delta t}))) = W(0,\mu(J(W_t+\widetilde{v}))) = W(0,\mu(\Gamma(W_t+\widetilde{v}))) = W(0,\Gamma\mu(W_t+\widetilde{v})), \quad (10)$$

is greater than the utility from *not* taking the gamble

$$W(0,\mu(J(W_{t+\Delta t}))) = W(0,\mu(J(W_t))) = W(0,\mu(\Gamma W_t)) = W(0,\Gamma\mu(W_t)). \quad (11)$$

The decision therefore comes down to comparing $\mu(W_t+\widetilde{v})$ and $\mu(W_t)$. Given the form of μ in (8), an investor with $\gamma > 1$ will reject an immediate gamble $(x, \frac{1}{2}; -y, \frac{1}{2})$ at wealth level W_t iff

$$(W_t + x)^{1-\gamma} + (W_t - y)^{1-\gamma} > 2W_t^{1-\gamma},$$

exactly the condition that emerged in the example following Proposition 1. Precisely the same reasoning shows that the preferences in (7)-(8) cannot simultaneously explain observations I and III.

Recursive Utility with second-order risk averse certainty equivalent [R-SORA]

We now turn to the second kind of recursive utility preference, R-SORA. In this case, it is impossible to prove that such preferences can *never* explain observations I-III. In particular, the Rabin (2000) argument can no longer be applied to the same extent as before. As mentioned earlier, the expected utility preferences in (2) and (3) have a very useful property, namely that the utility difference between two wealth levels does not depend on current wealth. As a result, attitudes to small risks at one wealth level provide very valuable information about attitudes to larger risks at other wealth levels. Without the EU assumption, however, this logic fails: the knowledge that someone rejects $\underline{G}_S$ at a wealth of \$20,000 provides little information about their attitudes to G_L at \$10,000 wealth.

While R-SORA preferences can, in principle, explain observations I-III, we argue that they can only do so for extreme parameterizations. One reason for this is already well-known in the literature. R-SORA preferences are locally smooth, which means that the investor is locally risk-neutral: she will accept an infinitesimally small, actuarially fair gamble. While $\underline{G}_S$ is not literally infinitesimal, it is virtually so; R-SORA preferences therefore need to adopt extreme parameters to explain its rejection.

This argument is a standard one. But we can go further. While it is certainly a challenge for R-SORA preferences to explain observation I, it is an even greater challenge for them to explain *both* observations I and II. The reasoning is essentially a "local" version of Rabin's argument. Suppose that R-SORA preferences are calibrated to explain the rejection of $\underline{G}_S$, the \$110/100 bet. They must therefore exhibit a very high level of risk aversion, making it very likely that the individual will reject, at the same wealth level, a more attractive gamble with larger stakes, such as

$$(2000, \frac{1}{2}; -500, \frac{1}{2}). \qquad (12)$$

For the standard case where the certainty equivalent functional μ is homogeneous, rejecting the gamble in (12) at wealth level W is equivalent to rejecting $(400, \frac{1}{2}; -100, \frac{1}{2})$ at wealth level $W/5$. To summarize, an R-SORA specification which explains the rejection of $\underline{G}_S$ at one wealth level will typically predict the rejection of $\overline{G}_S$, the \$400/100 bet, at a lower wealth level, making it especially hard for such preferences to explain observations I and II simultaneously.[7]

Example

To see the difficulties faced by R-SORA preferences, consider a simple example of a utility function in this class,

$$W(C, \mu) = ((1-\beta)C^{\rho} + \beta\mu^{\rho})^{\frac{1}{\rho}},\ \rho < 1,\ \rho \neq 0 \qquad (13)$$

where μ takes a form suggested by Chew and MacCrimmon (1979) and Chew (1983), namely "weighted utility". Given a gamble

$$\widetilde{V} = (x_1, p_1; x_2, p_2; \ldots; x_n, p_n),$$

μ is defined as

$$\mu(\widetilde{V}) = \left(\frac{p_1 x_1^{1-\gamma+\delta} + \ldots + p_n x_n^{1-\gamma+\delta}}{p_1 x_1^{\delta} + \ldots + p_n x_n^{\delta}} \right)^{1/(1-\gamma)},\ \gamma \neq 1. \qquad (14)$$

Risk aversion increases as γ increases or as δ falls. When $\delta = 0$, these preferences reduce to the standard power utility specification.

When investment opportunities are i.i.d., Epstein and Zin (1989) show that the individual's value function is given by

$$J(W_t) = \Gamma W_t.$$

An investor with such preferences therefore decides whether or not to take on an immediate gamble in the way laid out in equations (10) and (11). In particular, she accepts a gamble $\tilde{v} = (x, \frac{1}{2}; -y, \frac{1}{2})$ iff

$$\mu(W_t + \tilde{v}) > \mu(W_t),$$

or if

$$\left(\frac{(W_t + x)^{1-\gamma+\delta} + (W_t - y)^{1-\gamma+\delta}}{(W_t + x)^{\delta} + (W_t - y)^{\delta}} \right)^{1/(1-\gamma)} > W_t. \qquad (15)$$

[7] In the more general case where the certainty equivalent functional μ is not homogeneous, the same kind of argument can still be applied: massive aversion to local risks is likely to imply implausible aversion to larger bets. The nice feature of homogeneity is that we can scale these larger bets down into equivalent smaller bets, which are easier to play for real money in the laboratory.

The area shaded with "+" signs in the top panel of Figure 1 shows the range of values of γ and δ consistent with observation I, in other words, with the investor rejecting $\underline{G}_S$, the \$110/100 bet, at all wealth levels below \$100,000. Mathematically, these are the values of γ and δ for which inequality (15) fails for all $W_t < 100,000$ when

$$x = 110, \; y = 100.$$

The diagram shows, just as we predicted, that extreme values of γ and δ are needed just to explain observation I. For example, when $\delta = 0$, so that the preferences in (13)-(14) collapse to power utility preferences, we see, consistent with our earlier evidence, that $\gamma \geq 90$ is needed to explain the rejection of $\underline{G}_S$. The shaded area is concentrated in the bottom-right of the panel because risk aversion increases as we move towards the south-east, or as γ increases and δ falls.

This shaded area reappears, again marked with "+" signs, in the bottom panel of Figure 1. This panel also shows, now with "x" signs, the values of γ and δ consistent with observation II, i.e., for which the investor would accept $\overline{G}_S$, the \$400/100 bet, at all wealth levels above \$1,000. Mathematically, these are the values of γ and δ for which inequality (15) holds when $W_t > 1,000$ and

$$x = 400, \; y = 100.$$

This region is located in the top-left corner of the picture: once risk aversion climbs too high, the investor is no longer willing to accept $\overline{G}_S$.

The picture shows that while it is difficult to explain just observation I, explaining both observations I and II is even harder, with only a thin sliver of parameter values in the upper right-hand corner able to do the task. The intuition, once again, is that if the investor is so risk averse as to reject the \$110/100 bet at some wealth level, she will probably reject more attractive gambles with larger stakes, such as a \$2000/500 bet, at the same wealth level, and therefore also the \$400/100 bet at lower wealth levels.

Recursive Utility with first-order risk averse certainty equivalent [RU-FORA]

Finally, we turn to the last class of recursive utility preferences, R-FORA, in which the certainty equivalent μ is non-expected utility and exhibits first-order risk aversion. Such preferences certainly do a better job addressing observations I-III than the other specifications we have seen so far. In particular, they have no trouble explaining the attitudes in observations I-III, *so long as the gambles are played out immediately*, a critical caveat we return to shortly.

The intuition for why R-FORA preferences can explain atttitudes to immediate gambles is straightforward. The essence of the difficulty with EU, R-EU, and R-SORA preferences is that the investor is risk-neutral to small gambles, forcing us to push risk aversion up to dramatically high levels in order to explain the rejection of $\underline{G}_S$, the \$110/100 bet. An agent with R-FORA preferences, on the other hand, is by definition, locally risk averse. Risk aversion over large gambles does not, therefore, need to be increased very much to ensure that $\underline{G}_S$ is rejected.

Example

To illustrate, consider an investor with the following specific R-FORA preferences:

$$W(C,\mu) = ((1-\beta)C^\rho + \beta\mu^\rho)^{1/\rho}, \tag{16}$$

where μ takes a form developed by Gul (1991),

$$\mu(\widetilde{V}_{t+1})^{1-\gamma} = E(\widetilde{V}_{t+1}^{1-\gamma}) + (\lambda - 1)E((\widetilde{V}_{t+1}^{1-\gamma} - \mu(\widetilde{V}_{t+1})^{1-\gamma})1(\widetilde{V}_{t+1} < \mu(\widetilde{V}_{t+1})). \tag{17}$$

These preferences are often referred to as "disappointment aversion" preferences. The investor gets disutility if the outcome of the gamble $\widetilde{V}$ falls below its certainty equivalent μ, where the degree of disutility is governed by λ. This parameter effectively controls how sensitive the agent is to losses as opposed to gains. Any $\lambda > 1$ implies first-order risk aversion.

For i.i.d. investment opportunities, Epstein and Zin (1990) show that the investor's value function is given by

$$J(W_t) = \Gamma W_t. \tag{18}$$

We now check that (16)-(17) can easily be parameterized to explain all of observations I-III for the case of *immediate* gambles. As in our earlier examples of recursive utility, the investor evaluates an immediate gamble $\widetilde{v}$ by checking whether the utility from taking the gamble, given in (10), is greater than the utility from not taking the gamble, given in (11), which again reduces to comparing $\mu(W_t + \widetilde{v})$ and $\mu(W_t)$. Given the functional form of μ, a little algebra shows that observations I-III can be simultaneously explained if there exist γ and λ such that

$$(W_t + 110)^{1-\gamma} + \lambda(W_t - 100)^{1-\gamma} > (1+\lambda)W_t^{1-\gamma} \tag{19}$$

holds for all wealth levels below $\$100,000$,

$$(W_t + 400)^{1-\gamma} + \lambda(W_t - 100)^{1-\gamma} < (1+\lambda)W_t^{1-\gamma} \tag{20}$$

hold for all wealth levels above $\$1,000$ and

$$(W_t + 20,000,000)^{1-\gamma} + \lambda(W_t - 10000)^{1-\gamma} < (1+\lambda)W_t^{1-\gamma} \tag{21}$$

holds for all wealth levels above $\$50,000$. Condition (19) ensures that $\underline{G}_S$ is rejected, condition (20) that $\overline{G}_S$ is accepted and condition (21) that G_L is also accepted, all at the appropriate wealth levels. A quick computation confirms that all three of (19), (20) and (21) can be satisfied with $\gamma = 2$ and $\lambda = 2$. The intuition is that since λ controls sensitivity to losses as opposed to gains, we need λ to be somewhere between 1.1 and 4 so that the $\$110/100$ bet is rejected and the $\$400/100$ bet accepted.

4 The Importance of Narrow Framing

At first sight, then, it appears that preferences with first-order risk aversion can explain observations I-III. However, we now show that they can only do so in a very special case, namely when the monetary gambles are *immediate*. In the more realistic and general setting where the gambles are played out with some delay, however small or large, they have a much harder time explaining observations I-III. In other words, while they can easily explain aversion to small, immediate gambles, they have great difficulty – in a sense that we make precise below – capturing aversion to small, *delayed* gambles. This is a serious concern because as we saw in Section 2, people seem to be just as averse to the \$110/100 bet when it is played out quickly as when it is played out with delay. More generally, most real-world risks are delayed, making it important to get attitudes to such gambles right.

Before giving a precise statement of the difficulty with R-FORA preferences, we give an informal example to illustrate the idea. Consider a simple one-period utility function exhibiting first-order risk aversion,

$$w(x) = \begin{cases} x & \text{for} \quad x \geq 0 \\ 2x & \phantom{\text{for}} \quad x < 0 \end{cases}.$$

It is easy for such a utility function to explain why someone might reject the small, *immediate* gamble $(110, \frac{1}{2}; -100, \frac{1}{2})$: the individual would assign the gamble a value of $110(\frac{1}{2}) - 2(100)(\frac{1}{2}) = -45$, the negative outcome signalling that the gamble should be rejected. But how would this individual deal with the more realistic case of a small, *delayed* gamble?

In answering this, it is important to recall the essential difference between an immediate and a delayed gamble. The difference is that over the time interval that the uncertainty surrounding the delayed gamble is being played out, the individual is likely to be facing *other* sources of risk at the same time, such as labor income risk, house price risk, or risk from other financial investments. This is not true for the immediate gamble.

For the R-FORA preferences in (16)-(17), this distinction can have a big impact on whether a gamble is accepted. Suppose that the individual is facing the pre-existing risk $(30,000, \frac{1}{2}; -10,000, \frac{1}{2})$, to be resolved at the end of the period, and is wondering whether or not to take on an independent delayed gamble $(110, \frac{1}{2}; -100, \frac{1}{2})$, whose uncertainty is also to be resolved at the end of the period. The correct way for him to think about this is to merge the new gamble with the pre-existing gamble, and to check whether the combined gamble offers a high utility. Since the combined gamble is

$$(30,110, \frac{1}{4}; 29,900, \frac{1}{4}; -9,890, \frac{1}{4}; -10,100, \frac{1}{4}),$$

the comparison is between

$$30,000(\frac{1}{2}) - 2(10,000)(\frac{1}{2}) = 5000$$

and

$$30,110(\frac{1}{4}) + 29,900(\frac{1}{4}) - 2(9890)\frac{1}{4} - 2(10,100)\frac{1}{4} = 5007.5.$$

The important point here is that the combined gamble offers *higher* utility. In other words, the investor would want to *accept* the small delayed gamble, even if she would reject an immediate gamble with the same stakes. The intuition is that since the investor is already facing some pre-existing risks, adding a small, independent gamble represents a form of diversification, which she willingly takes on.

This simple example suggests that even if the certainty equivalent μ exhibits first-order risk aversion, it may be very difficult to explain the rejection of $\underline{G}_S$, the \$110/100 bet. In Proposition 3 below, we make the nature of this difficulty precise. In brief, while an individual with R-FORA utility acts in a first-order risk averse manner toward immediate gambles, she acts in a *second-order* risk averse manner towards independent, delayed gambles, so long as she is already facing other pre-existing risks.

This immediately reintroduces the same two difficulties that we saw in Section 3 when discussing preferences with second-order risk aversion (R-SORA). First and foremost, since the agent is second-order risk averse over delayed gambles, and since the delayed gamble $\underline{G}_S$ is virtually infinitesimal, she will be very keen to accept it. In order to explain why it is typically rejected, extreme parameterizations are required.

Second, if the agent rejects the delayed gamble $(110, \frac{1}{2}; -100, \frac{1}{2})$ at some wealth level, the local risk aversion will need to be so large that she will probably also reject, at the same wealth level, a more attractive gamble with larger stakes, such as $(2000, \frac{1}{2}; -500, \frac{1}{2})$. Together with the standard homogeneity assumption on the certainty equivalent μ, this implies that she will reject $(400, \frac{1}{2}; -100, \frac{1}{2})$ at lower wealth levels, a counterfactual prediction. We illustrate both of these difficulties in an example following the proposition.

While Proposition 3 is proven for just one implementation of first-order risk aversion, namely the Gul (1991) implementation through the recursive utility, the argument used in the proof is very general and applies readily to other formalizations.

Proposition 3: Suppose an individual has first-order risk aversion preferences as implemented through the recursive utility framework of Gul (1991) laid out in (16)-(17) above, where the function W is strictly increasing and differentiable with respect to both arguments, and where h is strictly increasing with a positive first derivative and negative second derivative.

Suppose that the individual is offered an actuarially favorable gamble $k\tilde{\varepsilon}$ to pay off between time t and $t+1$, and that the payoffs do not affect, and are independent of, I_t and the future economic uncertainties. Finally, suppose also that prior to taking the gamble, the distribution of the agent's $t+1$ utility value $\tilde{V}_{t+1}$ does not have finite mass at μ.

Then, the individual will be second-order risk averse over the new gamble; in other words, *for sufficiently small k, she will accept it.*

Proof: See Appendix.

Example

We now illustrate the difficulties faced by R-FORA preferences with the help of a more formal example. The analysis closely mirrors the computations we did for the example of R-SORA preferences in Section 3. We first show that it is very difficult for the R-FORA preferences to explain the rejection of the delayed gamble $\underline{G}_S = (110, \frac{1}{2}; -100, \frac{1}{2})$; and that it is even more difficult to explain both the rejection of $\underline{G}_S$ *and* the acceptance of the delayed gamble $\overline{G}_S = (400, \frac{1}{2}; -100, \frac{1}{2})$ – in fact, we show that it is impossible to explain both facts.

To do the computations, we again consider an investor with the R-FORA preferences in (16)-(17). We assume, for simplicity, that the only investment opportunity available to the investor is a risky asset with gross return $\widetilde{R}$, which each period takes the value 1.2 or 0.98 with equal probability and which is i.i.d. over time. As before, the investor's value function takes the form

$$J(W_t) = \Gamma W_t.$$

In order to decide whether the agent takes on a delayed gamble, we need to compare the utility from not taking it,

$$W(C_t, \mu(J(\widetilde{W}_{t+1}))) = W(C_t, \mu(\Gamma \widetilde{W}_{t+1})) = W(C_t, \Gamma \mu(\widetilde{W}_{t+1})), \tag{22}$$

to the utility from taking it,

$$W(\widehat{C}_t, \mu(J(\widetilde{W}_{t+1} + \tilde{v}))) = W(\widehat{C}_t, \mu(\Gamma(\widetilde{W}_{t+1} + \tilde{v}))) = W(\widehat{C}_t, \Gamma \mu(\widetilde{W}_{t+1} + \tilde{v})). \tag{23}$$

The hat over $\widehat{C}_t$ is a reminder that if the investor takes on the gamble, her optimal consumption choice will be different from what it is when she does not take the gamble.

Figure 2, which has the same structure as Figure 1, presents the results. The shaded area in the top panel shows the range of values of γ and λ for which the agent rejects the delayed gamble $\underline{G}_S = (110, \frac{1}{2}; -100, \frac{1}{2})$. The figure confirms that extreme values are required to explain this rejection.

This shaded area reappears in the bottom panel, again marked with "+" signs. The panel also shows, marked with "x" signs, the range of values of γ and λ for which the agent accepts $\overline{G}_S = (400, \frac{1}{2}; -100, \frac{1}{2})$. As the figure shows, there is no overlap between the two shaded regions. In other words, there are *no* parameter values for which these R-FORA preferences can explain both observations I and II.

The intuition bears repeating: in the presence of pre-existing risk, the investor acts in a second-order risk averse manner towards small, delayed gambles. It therefore takes huge risk aversion to explain why a delayed gamble is rejected, so much, in fact, that she rejects more attractive gambles with larger stakes, such as $(2000, \frac{1}{2}; -500, \frac{1}{2})$, at that wealth level, and therefore also $(400, \frac{1}{2}; -100, \frac{1}{2})$ at lower wealth levels.

4.1 Incorporating Narrow Framing

So far, we have shown that three simple observations – I, II and III – pose considerable difficulties for almost every utility specification that has appeared in the asset pricing literature. What, then, *can* satisfy these observations? Clearly, first-order risk aversion is a necessary ingredient: we need it to explain why small gambles like $\underline{G}_S$, played out immediately, are rejected. However, the analysis earlier in this section shows that first-order risk aversion is not enough. Its weakness is that when an agent evaluates a small, delayed gamble, she merges it with pre-existing risks and since the resulting diversification is attractive, she is happy to accept it. To explain the *rejection* of such delayed gambles, then, it must be that the agent does *not* fully merge the gamble with pre-existing risks, but to some extent, evaluates it in isolation. More formally, her preferences must depend on the outcome of the gamble over and above what the outcome implies for aggregate wealth risk, a feature we call *narrow framing.*

We now check that preferences incorporating both first-order risk aversion and narrow framing can easily explain observations I-III, whether the gambles are played out immediately or with delay. Preferences of this type were originally proposed by Kahneman and Tversky (1979) and have been used in the context of asset pricing by Benartzi and Thaler (1995) and Barberis, Huang and Santos (2001). Here, we adopt a more tractable specification proposed by Barberis and Huang (2002), in which time t utility is given by

$$V_t = W\left[C_t, \mu(V_{t+1}) + b_0 E_t(\sum_i \overline{v}(G_{i,t+1}))\right] \tag{24}$$

where

$$\begin{aligned} W(c,y) &= ((1-\beta)C^{1-\gamma} + \beta y^{1-\gamma})^{1/(1-\gamma)} \\ \overline{v}(x) &= \begin{cases} x & \text{for} \quad x \geq 0 \\ \lambda x & \phantom{\text{for}} \quad x < 0 \end{cases} \\ \mu(V) &= (E(V^{1-\gamma}))^{1/(1-\gamma)}, \end{aligned} \tag{25}$$

and where $G_{i,t+1}$ are specific gambles faced by the investor whose uncertainty will be resolved between time t and $t+1$.[8]

[8] The specification in Barberis and Huang (2002) has numerous advantages over earlier formulations like that in Barberis, Huang and Santos (2001): it allows for a more tractable partial equilibrium analysis, offers a natural way of checking attitudes to monetary gambles and individual preferences no longer depend on an aggregate consumption scaling term.

The term prefixed by b_0 in (24) shows that relative to the usual recursive specification in (1), we now allow utility to depend on outcomes of gambles $G_{i,t+1}$ over and above what those outcomes mean for aggregate wealth risk. In other words, we allow for narrow framing, with the parameter b_0 controlling the degree of narrow framing.

Barberis and Huang (2002) show that just like an investor with the R-FORA preferences in (16)-(17), an investor with the preferences in (24)-(25) is first-order risk averse over immediate gambles. This time, however, the first-order risk aversion does not come from μ itself: note that the μ in (25) is of the expected utility class, and therefore second-order risk averse. This time, first-order risk aversion comes from the kink in $\overline{v}$ at the origin.

Barberis and Huang (2002) also show that an investor with the preferences in (24)-(25) is first-order risk averse not only to immediate gambles but to delayed gambles as well. The intuition is that when considering a delayed gamble, the investor does not fully merge it with pre-existing risks, but because of the $\overline{v}$ term in (24), evaluates it in isolation to some extent. The piece-wise linearity of $\overline{v}$ then induces first-order risk aversion over the gamble.

The fact that the preferences in (24)-(25) predict first-order risk aversion over both immediate and delayed gambles means that they should have little trouble explaining the rejection of $\underline{G}_S$, the \$110/100 bet, whether in its immediate or delayed form. This stands in stark contrast to the preferences in (16)-(17): in that case, we saw that while the agent is first-order risk averse over an immediate small gamble, he is second-order risk averse over a delayed gamble, thereby predicting the rejection of the former but the *acceptance* of the latter.

We now check that the preferences in (24)-(25) can indeed explain the attitudes in observations I-III without difficulty, whether the gambles are immediate or delayed. We consider the same context as earlier in this section. The investor's only investment opportunity is a risky asset, with a gross return $\widetilde{R}$ of 1.2 or 0.98 each period with equal probability. Barberis and Huang (2002) show that in this case, the investor's value function is given by

$$J(W_t) = \Gamma W_t.$$

The agent will therefore accept an immediate gamble $\widetilde{x}$ iff the utility from not taking it on,

$$W(0, \Gamma\mu(W_t))$$

is less than the utility from taking it,

$$W\left[0, \Gamma\mu(W_t + \widetilde{x}) + b_0 E_t(\overline{v}(\widetilde{x}))\right].$$

Analogously, he will take a delayed gamble $\widetilde{x}$ iff the utility from not taking it,

$$W(C_t, \Gamma\mu(W_{t+1}))$$

is less than the utility from taking it,

$$W\left[\widehat{C}_t, \Gamma\mu(W_{t+1} + \widetilde{x}) + b_0 E_t(\overline{v}(\widetilde{x}))\right],$$

with the hat over $\widehat{C}_t$ again indicating that if the investor takes the gamble, his consumption choice will be different from what it is in the case where he doesn't take on the gamble.

We set β, which has little direct influence on attitudes to risk, to 0.98, and b_0, which controls the degree of narrow framing, to 1. The top panel in Figure 3 shows the range of values of γ and λ consistent with observations I-III when the gambles are played out immediately, while the bottom panel shows the range of values when the gambles are delayed.

The results in the top panel appear sensible. So long as the investor engages in some narrow framing, in other words, so long as b_0 is sufficiently positive, her attitude to small gambles is largely determined by λ, which controls sensitivity to losses relative to sensitivity to gains. In order to explain why an investor rejects $(110, \frac{1}{2}; -100, \frac{1}{2})$ but accepts $(400, \frac{1}{2}; -100, \frac{1}{2})$, we need λ to range between 1.1 and 4, and this is roughly what the figure shows.

The bottom panel shows that in the presence of narrow framing, attitudes to delayed gambles are very similar to those for immediate gambles. It is therefore clear that there is a wide range of parameter values for which the preferences in (24)-(25) can explain observations I-III for *both* immediate and delayed gambles. This contrasts with R-FORA preferences, which lack narrow framing, and which predict very different attitudes to immediate and delayed gambles.

5 Discussion

5.1 Narrow Framing and the Equity Premium

In the previous section, we argued that preferences with first-order risk aversion and narrow framing are attractive because they are easily able to explain observations I-III, whether the underlying gambles are played out immediately or with delay.

One issue that we did not address is whether the parameter values consistent with observations I-III can also generate a substantial equity premium. Put differently, it may be that the kinds of parameter values required to generate a large equity premium are no longer consistent with observations I-III.

We now show that in a simple economy with a risky stock market, it is easy to find parameter values for which the representative agent with the preferences in (24)-(25) would, in equilibrium, demand a high equity premium and would also act consistently with observations I-III.[9]

[9] In assigning our preference specifications to the representative agent, we are clearly assuming that properties of individual preferences survive under aggregation. We do not prove any results about aggregation here – nor indeed, have aggregation results in the literature to date been proven in anything but the simplest

We consider an endowment economy with an infinite number of identical investors, and two assets: a risk-free asset in zero net supply, with gross return $R_{f,t}$ between time t and $t+1$, and a risky asset – the stock market – in fixed positive supply, with gross return R_{t+1} between time t and $t+1$. The stock market is a claim to a perishable stream of dividends $\{D_t\}$, where

$$\frac{D_{t+1}}{D_t} = e^{g_D + \sigma_D \varepsilon_{t+1}}, \tag{26}$$

and where each period's dividend can be thought of as one component of a consumption endowment C_t, where

$$\frac{C_{t+1}}{C_t} = e^{g_C + \sigma_C \eta_{t+1}}, \tag{27}$$

and

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix} \right), \text{ i.i.d. over time.} \tag{28}$$

In our quantitative analysis, we use the endowment process parameters listed in Table 1. These parameters are estimated from U.S. data spanning the 20th century and are standard in the literature.

Suppose that the investors in the economy have the preferences in (24)-(25), exhibiting both first-order risk aversion and narrow framing. At this point, we need to be precise as to which risks these investors frame narrowly; put differently, which of their risks do the G_i in equation (24) refer to?

We assume that they frame their stock market investments narrowly. In other words, they get direct utility from fluctuations in the value of their stock market holdings even though these holdings are only one part of their total wealth; in mathematical terms, $G_{1,t+1} = \widetilde{R}_{t+1} - 1$. We discuss possible interpretations of this assumption more fully in Section 5.3. For now, note that given this assumption, utility is

$$V_t = W\left[C_t, \mu(V_{t+1}) + b_0 E_t(\overline{v}(R_{t+1} - 1))\right]. \tag{29}$$

When we check whether these investors accept or reject simple monetary gambles, we assume that they frame those narrowly too. In other words, if they take on a simple gamble $\widetilde{x}$, their utility is given by

$$V_t = W\left[C_t, \mu(V_{t+1}) + b_0 E_t \overline{v}(R_{i,t+1} - 1) + b_0 E(v(\widetilde{x}))\right] \tag{30}$$

cases. What results there are, though, suggest that the representative agent's risk aversion is a weighted average of individuals' risk aversion. If all agents in our economy display aversion to small risks, then, it seems plausible that the representative agent will also, thereby ensuring that our results go through even after accounting for heterogeneity.

Barberis and Huang (2002) show that the Euler equations corresponding to the preferences in (29) are given by

$$\beta^{\frac{1}{1-\gamma}}(1-\alpha)^{-\frac{\gamma}{1-\gamma}} R_f E((\frac{C_{t+1}}{C_t})^{-\gamma})(E((\frac{C_{t+1}}{C_t})^{-\gamma}))^{\frac{\gamma}{1-\gamma}} = 1 \quad (31)$$

$$\frac{E((\frac{C_{t+1}}{C_t})^{-\gamma}(R_{S,t+1}-R_f))}{E((\frac{C_{t+1}}{C_t})^{-\gamma})} + b_0 R_f^{-1}(\frac{\beta}{1-\beta})^{\frac{1}{1-\gamma}}(\frac{1-\alpha}{\alpha})^{-\frac{\gamma}{1-\gamma}} E(\overline{v}(R_{S,t+1}-R_f)) = 0 \quad (32)$$

where α is the fraction of wealth consumed by the investor. Given a risk-free rate R_f, α is obtained from (31); with α in hand, (32) can then be used to compute the equity premium.

We again fix β and b_0 to be 0.98 and 1, respectively. The top panel in Figure 4 shows the range of values of γ and λ for which the agent charges an equity premium in excess of *four* percent, a very substantial amount. (The risk-free rate in this economy is 2 percent). The bottom panel shows the values for which, given her equilibrium holdings of stocks, she acts consistently with observations I-III when evaluating gambles $\underline{G}_S$, $\overline{G}_S$ and G_L, whether these gambles are played out immediately or with delay.[10]

The figure confirms that preferences with first-order risk aversion and narrow framing do indeed offer financial economists some attractive features, in that they are able to generate a very substantial equity premium and at the same time, make sensible predictions about attitudes to simple monetary gambles. In particular, the bottom panel shows that λ needs to range between 1.1 and 4 to explain attitudes to $\underline{G}_S$, $\overline{G}_S$ and G_L. The top panel shows that $\lambda > 2.3$ is necessary to produce a sizeable equity premium. Any parameterization with λ between 2.3 and 4 can therefore generate a large equity premium as well as plausible attitudes to laboratory gambles.

The first-order risk aversion and narrow framing that are helpful in explaining attitudes to monetary gambles are precisely the features that help generate such large equity premia. Our investors worry directly about gains and losses in the stock market and because $\lambda > 1$, they are more sensitive to losses than to gains. They fear large losses in the stock market and therefore charge a high equity premium as compensation. Precisely the same mechanism is suggested in Benartzi and Thaler (1995) and Barberis, Huang and Santos (2001) as a possible source of the premium.

5.2 The Importance of Small Gambles

One apparent weakness of our approach is that in discrediting a vast array of standard utility functions, we seem to rely heavily on the fact that people typically reject small gambles, even

[10] The bottom panel is not the same as the intersection of the two regions in Figure 3 because the agent's initial holdings of risky assets are different. In the one case, she is holding a prior risk with two possible outcomes, in the other, a log-normally distributed risk, equity, with average return given by equations (31)-(32).

if they are actuarially favorable. To some people, such gambles might seem "unimportant"; it is good enough, they might say, to choose preference specifications that can satisfactorily explain attitudes to large gambles: small gambles can be safely ignored.

In our opinion, this critique is a weak one, for at least two reasons. First, we have stressed the rejection of small gambles of the \$110/100 variety because they are the most difficult for the standard theories to explain, but similar arguments apply to gambles with larger stakes, such as $(1100, \frac{1}{2}; -1000, \frac{1}{2})$, which are harder to dismiss as unimportant. This gamble is also typically rejected, and yet most utility functions have trouble explaining this, for precisely the reasons laid out in Sections 3 and 4.

Second, for financial economists at least, small gambles are important for the simple reason that stocks themselves are a small gamble. In a moment, we provide evidence for this claim; suppose for now that it is true, and consider its implications. If stocks are a small gamble, a researcher who wants to address the equity premium via investor preferences – in other words, a researcher who wants to generate aversion to stock risk – must find preferences which do a good job explaining aversion to small gambles. And as we have seen in this paper, aversion to small gambles is best explained through first-order risk aversion and narrow framing. We can state this idea in more specific terms. Our results in Sections 3 and 4 have shown that when most preference specifications are calibrated to explain aversion to small gambles, they make counterfactual predictions about attitudes to gambles like G_L, the \$20,000,000/10,000 gamble, and $\overline{G}_S$, the \$400/100 gamble. If stocks are themselves a small gamble, it is likely that when preferences are calibrated to generate aversion to stocks and hence a large equity premium, they will again make implausible predictions about attitudes to G_L and $\overline{G}_S$.

But why are stocks a small gamble? There are a number of ways of seeing this. One is to look at the asset holdings of the typical investor, and to note that stock market holdings are indeed a small portion of their wealth. For example, according to the Investment Company Institute, the average equity owner in the U.S. is 47 years old, makes \$62,500 per year, has \$50,000 in equities and another \$50,000 in financial assets *excluding* home equity. Even the roughest estimates of this individual's human capital wealth or housing wealth suggest that the risk associated with equity holdings is small relative to total wealth.

The low correlation of consumption growth and stock returns also suggests that stocks are effectively a small gamble. If stocks were a large part of total wealth, we would normally expect changes in stock prices to be closely related to changes in consumption. This does not appear to be the case.

Finally, we can provide some analytical evidence that attitudes towards stock risk have much in common with attitudes to small gambles like $\underline{G}_S$, the \$110/100 bet. Suppose that we take two preference specifications and calibrate them so that they deliver similar attitudes to large gambles, but different attitudes to small gambles. Then if stock risk is closer to being a small gamble than a large gamble, the two preference specifications should deliver very different equity premia when applied to the representative agent. In particular, the preference specification displaying greater aversion to small gambles should give a higher equity premium.

It is straightforward to confirm this. Consider first the R-FORA preferences in (16)-(17). It is simple to show that when γ and λ are set to 1.5 and 2.5 respectively, these preferences predict the acceptance of the large gamble G_L as well as the acceptance of the small gamble $\underline{G}_S$.[11] Now consider the preferences in (24)-(25), exhibiting both first-order risk aversion and narrow framing. When γ, λ and b_0 are set to 1.5, 2.5 and 1, respectively, it is simple to show that the agent accepts the large gamble G_L, but *rejects* the small gamble $\underline{G}_S$. To a first approximation, then, these two preferences agree on attitudes to large gambles, but have very different predictions for small gambles.

We now examine the equity premia generated by these two preference specifications when they are applied to the representative agent in the simple endowment economy laid out in (26)-(28). The Euler equations for the preferences in (24)-(25), exhibiting both first-order risk aversion and narrow framing, have already been set out in (31)-(32). It is simple to check that for $(\gamma, \lambda, b_0) = (1.5, 2.5, 1)$, and for a risk-free rate of 2 percent, the equity premium is a very considerable 4.53 percent.

Epstein and Zin (1989) show that for the first-order risk averse preferences in (16)-(17), the Euler equations are

$$E_t \left[\phi \left((\frac{\rho}{1-\alpha})^{1/\theta} \frac{C_{t+1}}{C_t} \right) \right] = 0 \tag{33}$$

$$E_t \left[\phi' \left((\frac{\rho}{1-\alpha})^{1/\theta} \frac{C_{t+1}}{C_t} \right) (R_{S,t+1} - R_{f,t}) \right] = 0, \tag{34}$$

where α is the constant fraction of wealth consumed each period, and where

$$\begin{aligned} \phi &= \frac{x^{1-\gamma} - 1}{1-\gamma}, \; x \geq 1 \\ &= \lambda \frac{x^{1-\gamma} - 1}{1-\gamma}, \; x < 1. \end{aligned}$$

When applied to the total wealth portfolio, (34) gives

$$E_t \left[\phi' \left((\frac{\rho}{1-\alpha})^{1/\theta} \frac{C_{t+1}}{C_t} \right) (\frac{1}{1-\alpha} \frac{C_{t+1}}{C_t} - R_{f,t}) \right] = 0. \tag{35}$$

Equation (33) determines α, equation (34) determines the riskfree rate R_f and (35) determine the expected stock return, thereby giving the equity premium.

[11] To be more precise, they actually predict that $\underline{G}_S$ is rejected in the very special case where the gamble is played out immediately, but, as we saw in Section 4, that it is accepted in the more general and realistic case where it is played out with some delay.

It is now simple to check that for $(\gamma, \lambda) = (1.5, 2.5)$, the equity premium is 0.52 percent, almost *ten times lower* than the 4.53 percent delivered by the other preference specification. In other words, by keeping attitudes to large gambles constant but changing attitudes to small gambles, we obtain a substantial shift in the equity premium, suggesting that attitudes to stocks do indeed have much in common with attitudes to small gambles.

5.3 First Order Risk Aversion and Narrow Framing: Evidence and Explanations

In this paper, we have tried to argue that preferences with first-order risk aversion and narrow framing offer financial economists many attractive features. Is there any evidence, though, that people actually display these features in their decision making? And, if so, why? We address the former issue first.

As we have argued in this paper, the commonly observed aversion to small gambles, whether in the form of laboratory gambles like $\underline{G}_S$ or real-life gambles involving automobile collision risks, is in itself evidence of first-order risk aversion and narrow framing. But there is much other evidence for these two phenomena, and in particular, for a kind of first-order risk aversion known as loss aversion, where people are more sensitive to losses than to gains.

For example, there is what Thaler (1980) calls the endowment effect, which is loss aversion in the absence of uncertainty. Kahneman, Knetsch and Thaler (1990) conducted a series of experiments in which subjects were either given some object such as a coffee mug and then asked if they would be willing to sell it, or not given the mug and offered a chance to buy one. They found that mug owners demanded more than twice as much to sell their mugs as non-owners were willing to pay to acquire one.

Other evidence of loss aversion in combination with narrow framing comes from Thaler, Kahneman, Tversky and Schwarz (1990) and Benartzi and Thaler (1999). Both these papers ask subjects how they would allocate between a risk-free asset and a risky asset over a long time horizon, thirty years, say. The key manipulation is that some subjects are given the distribution of asset returns over short horizons – monthly returns, say – while others are given long-term return distributions – the distribution of 30-year returns, say. In principle, the two groups of subjects should make similar allocation decisions, since they have the same decision problem: those subjects given shorter-term return distributions should simply use them to infer the more directly relevant longer-term distributions. In fact, these subjects allocate substantially less to the risky asset, suggesting that they persist in using the "narrow" frame of short-term returns to make a long-term allocation decision. Since losses occur more frequently over short horizons, their loss aversion leads them to allocate less to risky assets.[12]

[12]Gneezy and Potters (1997) obtain very similar results, while a follow-up study by Gneezy, Kapteyn and

Narrow framing is a very general phenomenon, and need not occur in combination with loss aversion. Shortly after publishing their work on prospect theory, Tversky and Kahneman (1981) introduced the idea of decision framing and demonstrated narrow framing with examples such as the following:

Imagine that you face the following pair of concurrent decisions. First examine both decisions, then indicate the options you prefer:

Choice (I) Choose between:

A. a sure gain of $240

B. 25% chance to gain $1000 and 75% chance to gain nothing

Choice (II) Choose between:

C. a sure loss of $750

D. 75% chance to lose $1000 and 25% chance to lose nothing

A large majority of subjects choose A and D; however, this choice is dominated by B and C. This suggests that subjects are framing their decision narrowly, treating each choice separately as if it had no connection to the other; indeed in situations where subjects are asked to choose *only* between A and B, or *only* between C and D, they typically choose A and D, respectively. Conversely, if they are asked to choose between options E and F, where E = A+D and F=B+C, then every subject chooses F. Subjects do not knowingly choose a dominated alternative, but when the dominance is not transparent, they make choices one at a time.

Now that we have presented a few examples of narrow framing, we can ask why it is that people might have this feature embedded in their preferences, particularly when in some cases, as in the example above, it can lead to violations of invariance and dominance. Our best answer to this question is that narrow framing is a simplifying strategy for dealing with a very complex world. Suppose that someone is presented with clear information about a gamble whose outcome will be resolved at some point in the future and is wondering whether or not to accept it. Even if she knows that the right thing to do is to integrate the gamble with other gambles she is already facing, and then to check whether the resulting overall risk is preferable to the overall risk she previously faced, it may be difficult to do this from a computational standpoint. The individual may not be sure about the probability distribution of the outcomes for her other gambles, nor about the correlation between the gamble under consideration and previously accepted risks. As a result, she may adopt the general rule of deciding each gamble she faces in isolation, as if it is the only risk she faces in the world – in

Potters (2002) finds that manipulating the type of information provided can affect prices in an experimental asset market.

other words, she may use narrow framing. Computational difficulties would certainly seem to underlie the use of narrow framing in Tversky and Kahneman's (1981) example above.

In Section 5.1., we explained the equity premium by saying that agents get utility from the outcome of their stock market investments even if those investments are only one part of their overall wealth; in other words, they frame the stock market narrowly. Does it seem plausible that narrow framing might indeed apply in the case of the stock market?

The bounded rationality view of narrow framing suggests that it might. When people think about how much to invest in the stock market, they typically spend substantial time thinking about the distribution of outcomes for the stock market gamble itself. However, integrating this risk with other risks they face, such as labor income risk, house price risk, and proprietary business risk, is much more difficult. Few people have a good quantitative sense of the size of the other risks, nor of the correlation between them and stock market risk. Given these complications and uncertainties, it is possible that people frame stock market risk narrowly to some extent, ignoring the other risks they face.

All this is not to say that people never take a broader view. If the stakes are very large, and the relation between various contingencies are obvious then a broader frame may be adopted. An example might be when an employee reaches retirement and is considering what to do with the proceeds of a defined contribution pension plan. In such circumstances, other sources of income and wealth are more likely to be considered in making a choice than, say, when the asset allocation decision was last revisited while the employee was still working.

6 Conclusion

Many different preference specifications have been proposed as a way of addressing the equity premium. How should we pick between them? We suggest one possible metric, namely these utility functions' ability to explain *other* evidence on attitudes to risk. We consider some simple observations about attitudes to monetary gambles with just two outcomes and show that the vast majority of utility functions used in asset pricing have difficulty explaining these observations. However, utility functions with two features – first-order risk aversion and narrow framing – can easily explain them. We argue that by this metric at least, such utility functions may be very attractive to financial economists: they can generate substantial equity premia and at the same time, make sensible predictions about attitudes to monetary gambles.

7 Appendix

Proof of Proposition 1. (We prove part (b) here. The argument for part (a) is very similar). With expected utility, and under the assumption that the outcome of a gamble does not affect I_t and is independent of all other economic uncertainty, the individual's preference to the gamble is evaluated through $E_t[J(W_t + \widetilde{v}; I_t, C_{-t})]$, where not taking the gamble corresponds to $\widetilde{v} = 0$. The argument in Rabin (2000), which applies to one-period utility functions defined over wealth, can therefore be applied to $J(W_t; I_t, C_{-t})$, giving the result.

Proof of Proposition 2. Epstein and Zin (1989) propose that an individual with recursive utility preferences evaluates an immediate gamble by inserting an infinitesimal time step Δt at time t and applying the recursive utility calculation over this time step. Under the assumption that the outcome of a gamble does not affect I_t and is independent of all other economic uncertainty, the individual's preference to the gamble is evaluated through

$$W(0, \mu(J(W_t + \widetilde{v}; I_t))) = W(0, h^{-1}[E(h \cdot J(W_t + \widetilde{v}; I_t))])$$

where not taking the gamble corresponds to $\widetilde{v} = 0$. In this case, immediate gambles again are ranked by expected utility over wealth, with utility function given by $h \cdot J(\cdot)$. The argument in Rabin (2000), which applies to one-period utility functions defined over wealth, can therefore be applied to $h \cdot J(\cdot)$.

Proof of Proposition 3. For a small change in the period $t+1$ value function $\Delta\widetilde{V}_{t+1} = \Delta\widetilde{V}(\widetilde{W}_{t+1}, I_{t+1})$, the certainty equivalent changes by

$$\Delta\mu = \frac{E(u'(\widetilde{V}_{t+1})\Delta\widetilde{V}_{t+1}) + (\lambda - 1)E(u'(\widetilde{V}_{t+1})\Delta\widetilde{V}_{t+1}1(\widetilde{V}_{t+1} < \mu))}{u'(\mu)(1 + (\lambda - 1)\Pr(V_{t+1} < \mu))} + o(||\Delta V_{t+1}||)$$

where $||x|| = E(|x|)$ and $\lim_{x \to 0}(o(x)/x) = 0$ by definition, and $u(\cdot)$ can be a general smooth and increasing function as in Gul (1991) or can be taken to be $u(x) \equiv x^{1-\gamma}$ for the special case considered in this paper.

Assume for now that the agent does not optimally adjust his optimally chosen consumption and portfolio strategy if he decides to take the gamble. Then we have

$$\Delta\widetilde{V}_{t+1} = V_W(\widetilde{W}_{t+1}, I_{t+1})\widetilde{v} + o(||\widetilde{v}||).$$

which implies

$$\Delta\mu = \frac{E(u'(\widetilde{V}_{t+1})V_W(\widetilde{W}_{t+1}, I_{t+1})\widetilde{v}) + (\lambda - 1)E(u'(\widetilde{V}_{t+1})V_W(\widetilde{W}_{t+1}, I_{t+1})\widetilde{v}1(\widetilde{V}_{t+1} < \mu))}{u'(\mu)(1 + (\lambda - 1)\Pr(V_{t+1} < \mu))} + o(||\widetilde{v}||).$$

Given our assumption that v is independent of other economic uncertainties, we have

$$\Delta\mu = E(\widetilde{v})\frac{E(u'(\widetilde{V}_{t+1})V_W(\widetilde{W}_{t+1}, I_{t+1})) + (\lambda - 1)E(u'(\widetilde{V}_{t+1})V_W(\widetilde{W}_{t+1}, I_{t+1})1(\widetilde{V}_{t+1} < \mu))}{u'(\mu)(1 + (\lambda - 1)\Pr(V_{t+1} < \mu))} + o(||\widetilde{v}||).$$

So to first order, the certainty equivalence value of $\widetilde{V}_{t+1}$ depends only on $E(\widetilde{v})$, not on its standard deviation.

Finally, the aggregator function $W(\cdot,\cdot)$ will not generate any first-order dependence on the standard deviation of the gamble $\widetilde{v}$. In addition, assuming that the agent will adjust his consumption and portfolio choice optimally while taking into account the additional gamble, should he choose to take it on, will only introduce terms of second order of $\widetilde{v}$.

8 References

Abel, A. (1990), “Asset Prices Under Habit Formation and Catching Up with the Joneses,” *American Economic Review Papers and Proceedings* 80, 38-42.

Barberis, N. and M. Huang (2002), “A New Preference Specification for Addressing the Equity Premium,” Working paper, Stanford University.

Barberis, N., Huang M., and T. Santos (2002), “Prospect Theory and Asset Prices,” *Quarterly Journal of Economics* 116, 1-53.

Bekaert, G., Hodrick R., and D. Marshall (1997), “The Implications of First-order Risk Aversion for Asset Market Risk Premiums,” *Journal of Monetary Economics* 40, 3-39.

Benartzi, S. and R. Thaler (1995), “Myopic Loss Aversion and the Equity Premium Puzzle,” *Quarterly Journal of Economics* 110, 75-92.

Benartzi, S. and R. Thaler (1999), “Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments,” *Management Science* 45, 364-381.

Campbell, J.Y. (1996), “Understanding Risk and Return,” *Journal of Political Economy* 104, 298-345.

Campbell, J.Y. and J. Cochrane (1999), “By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior,” Journal of Political Economy,

Campbell, J.Y. and L. Viceira (1999), “Consumption and Portfolio Decisions when Expected Returns are Time Varying,” *Quarterly Journal of Economics* 114, 433-495.

Chew S. (1983), “A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox,” *Econometrica* 51, 1065-1092.

Chew S. and K. MacCrimmon (1979), “Alpha-nu Choice Theory: An Axiomatization of Expected Utility,” Working paper, University of British Columbia.

Constantinides, G. (1990), “Habit Formation: A Resolution of the Equity Premium Puzzle,” *Journal of Political Economy* 98, 519-543.

Epstein, L. (1992), “Behavior Under Risk: Recent Developments in Theory and Applications,” in *Advances in Economic Theory: Sixth World Congress of the Econometric Society*, J. Laffont ed., Cambridge: Cambridge University Press.

Epstein, L. and S. Zin (1989), “Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Perspective,” *Econometrica* 57, 937-969.

Epstein, L. and S. Zin (1990), "First-order Risk Aversion and the Equity Premium Puzzle," *Journal of Monetary Economics* 26, 387-407.

Epstein, L. and S. Zin (1991a), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy* 99, 263-266.

Epstein, L. and S. Zin (1991b), "The Independence Axiom and Asset Returns," Working paper, Carnegie Mellon University.

Gneezy, U. and J. Potters (1997), "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics* 112, 631-645.

Gneezy, U., Kapteyn, A., and J. Potters (2002), "Evaluation Periods and Asset Prices in a Market Experiment," forthcoming *Journal of Finance*.

Grgeta, A., and R. Thaler (2002), "Irrational Behavior in the Collision Insurance Market," Working paper, University of Chicago.

Gul, F. (1991), "A Theory of Disappointment Aversion," *Econometrica* 59, 667-686.

Kahneman, D., Knetsch A. and R. Thaler (1990), "Experiment Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy* 98, 1325-1348.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263-291.

Kandel, S. and R. Stambaugh (1991), "Asset Returns and Intertemporal Preferences," *Journal of Monetary Economics* 27, 39-71.

Rabin, M. (2000), "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica* 68, 1281-1292.

Rabin, M. and R. Thaler (2000), "Anomalies: Risk Aversion," *Journal of Economic Perspectives* 15, 219-232.

Read, D., Loewenstein G., and M. Rabin (1999), "Choice Bracketing," *Journal of Risk and Uncertainty* 19, 171-197.

Rubinstein (2001), "Comments on the Risk and Time Preferences in Economics," Working paper, Princeton University.

Segal, U. and A. Spivak (1990), "First order versus Second Order Risk Aversion," *Journal of Economic Theory* 51, 111-125.

Sundaresan, S. (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies* 2, 73-88.

Thaler R. (1980), "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization* 1, 39-60.

Thaler, R., Tversky A., Kahneman D., and A. Schwartz (1997), "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test," *Quarterly Journal of Economics* 112, 647-661.

Table 1: Parameter values for a simple consumption-based model.

Parameter	
g_C	1.84%
σ_C	3.79%
g_D	1.5%
σ_D	12.0%
ω	0.15

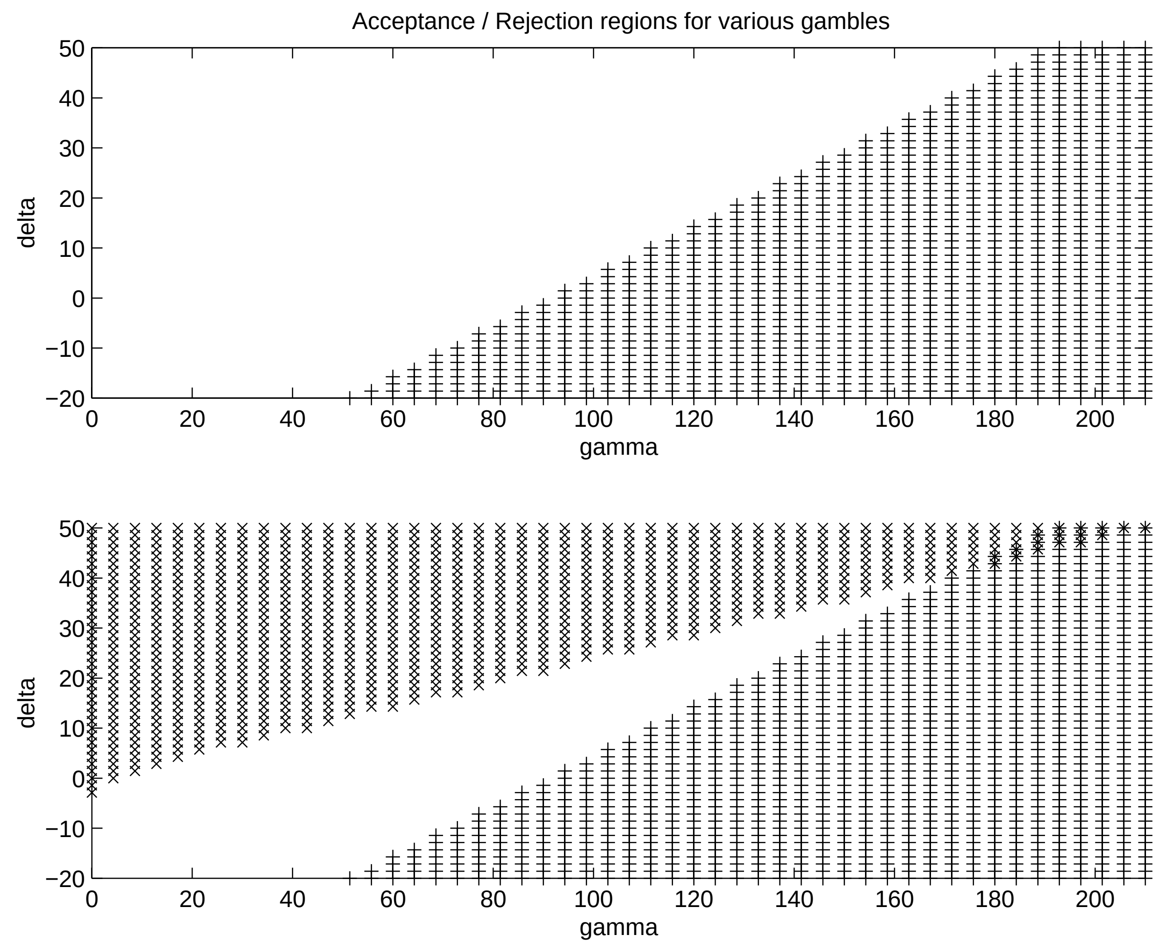

Figure 1. The + signs show the range of values of γ and δ for which a recursive utility function with Chew (1983) certainty equivalent rejects the gamble $(110, \frac{1}{2}; -100, \frac{1}{2})$. The x signs show where the same preferences accept the gamble $(400, \frac{1}{2}; -100, \frac{1}{2})$.

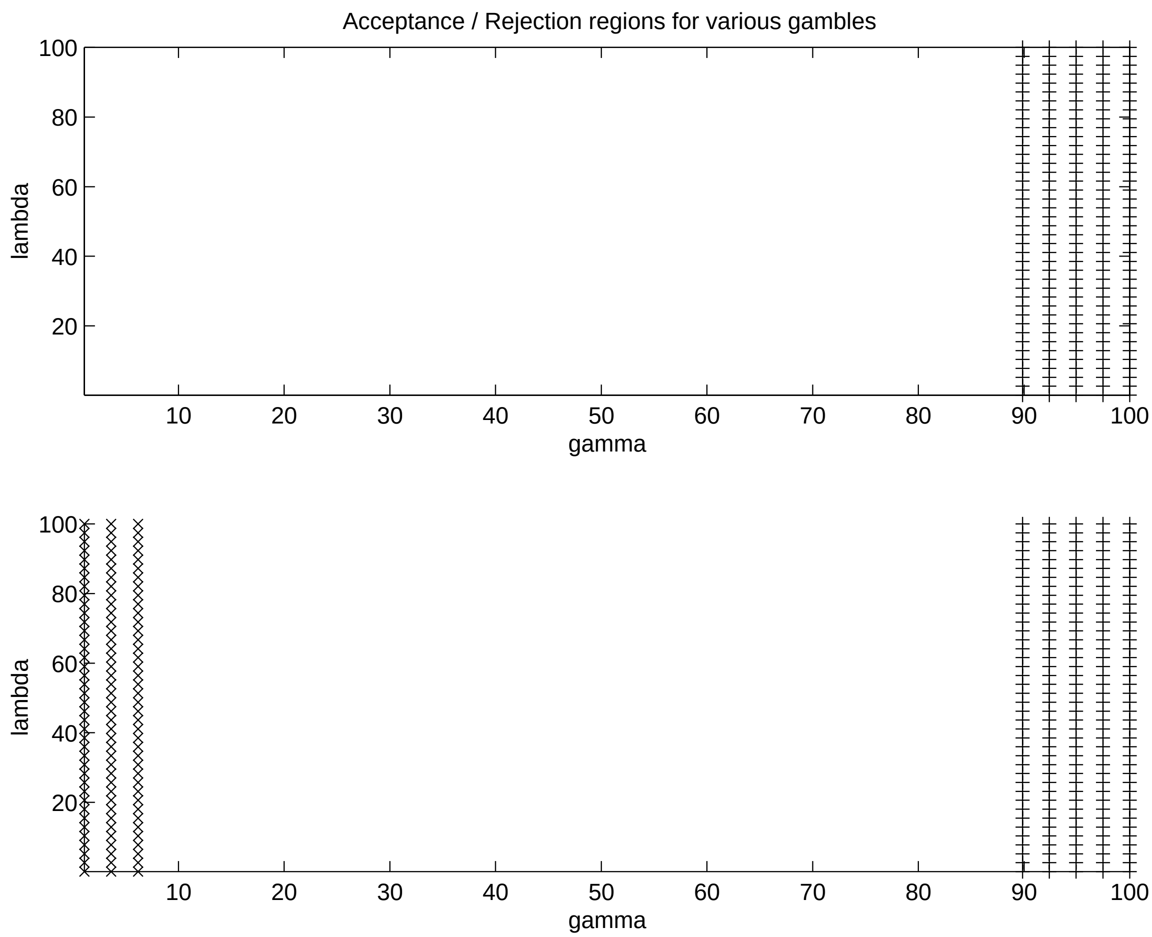

Figure 2. The + signs show the range of values of γ and λ for which a recursive utility function with Gul (1991) certainty equivalent rejects the delayed gamble $(110, \frac{1}{2}; -100, \frac{1}{2})$. The x signs show where the same preferences accept the delayed gamble $(400, \frac{1}{2}; -100, \frac{1}{2})$.

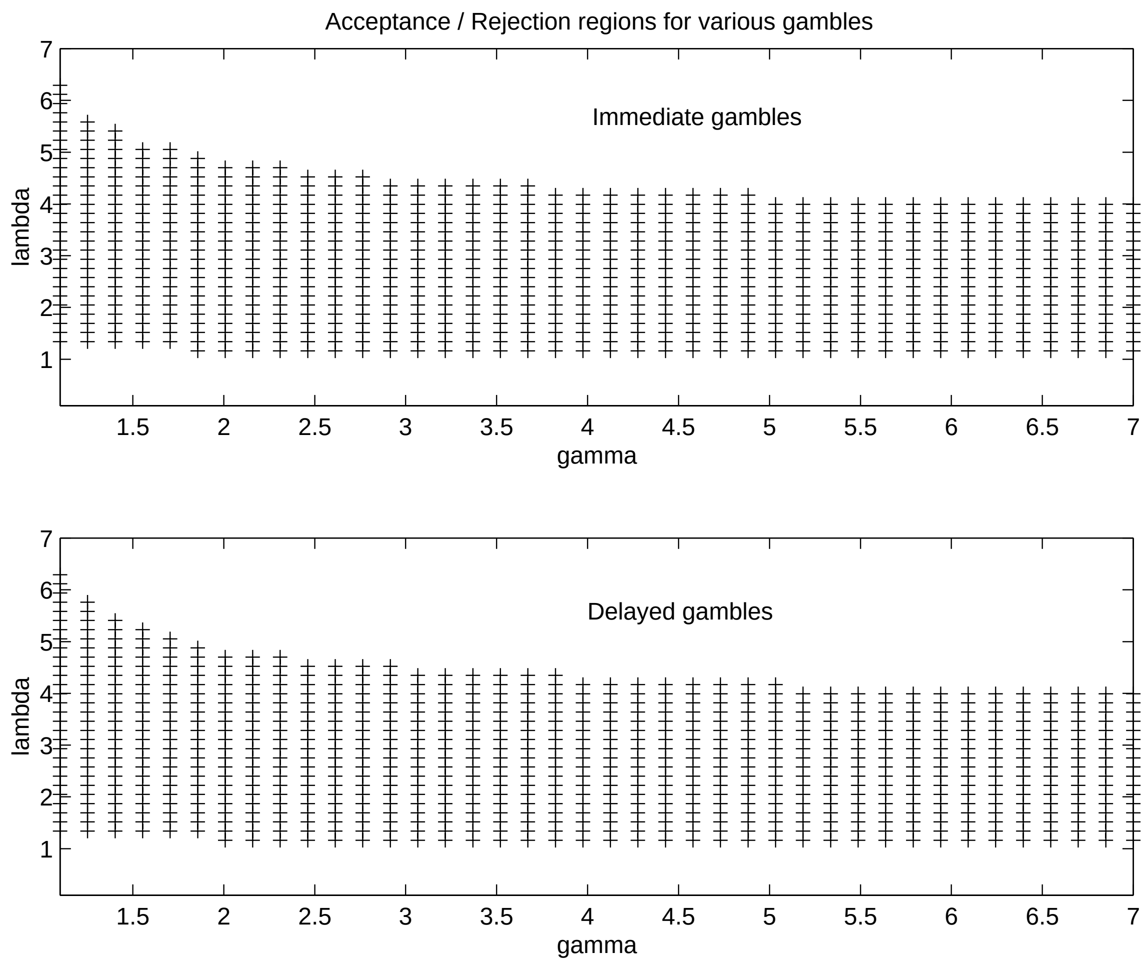

Figure 3. The shaded area in the top panel shows the range of parameter values for which a recursive utility function with both first-order risk aversion and narrow framing rejects certain immediate gambles and accepts others. The bottom panel repeats the same analysis for delayed gambles.

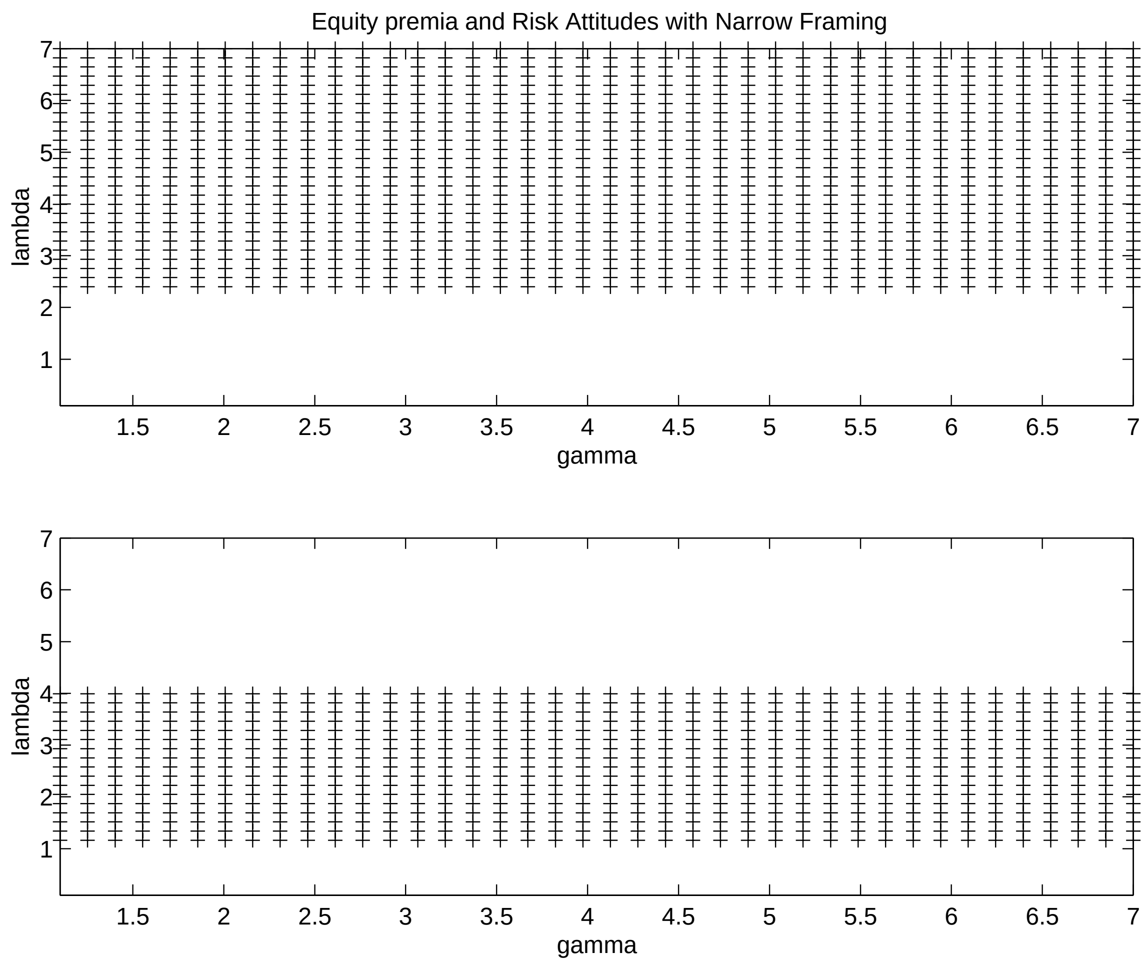

Figure 4. The top panel shows the range of parameter values for which a recursive utility function with both first-order risk aversion and narrow framing produces an equity premium in excess of four percent. The bottom panel shows the range of parameter values for which the same preferences reject certain laboratory gambles and accept others.