

# **Evolution of Risk Aversion: The "Having Descendants Forever" Approach**

Moshe (Shiki) Levy\*

## **Abstract**

Recent models of the evolution of preferences have provided profound new insights into the origins of risk attitudes. In most of these models the evolutionary "objective function" is the maximization of the expected number of offspring, or alternatively, the maximization of the geometric-mean growth rate. We suggest that careful consideration should also be given to the objective of maximizing the probability of Having Descendants Forever,  $p(\text{HDF})$ , first suggested by Meginniss (1977). We show that the  $p(\text{HDF})$  criterion implies risk aversion. Moreover, it leads to preferences that are very closely approximated by the power CRRA preferences. Thus, constant relative risk aversion can be viewed as an evolutionary-developed heuristic aimed to maximize the probability of having descendant forever.

*Keywords: evolution, preference, having descendants forever, constant relative risk aversion, loss aversion.*

JEL classification: D81, D03.

\*Jerusalem School of Business, the Hebrew University. [mslm@huji.ac.il](mailto:mسلم@huji.ac.il).

*Your descendants would have been like the sand, your children like its numberless grains; their name would never be cut off nor destroyed from before me.*

Isaiah 48:19

## **Introduction**

Most economic models take preferences as exogenously given. A large body of economic and psychological literature is devoted to the experimental and empirical estimation of preferences. However, until recently much less attention has been given to the question of *why* preferences are what they are. This question is interesting and important not only because it yields insights into understanding why we behave as we do, but also because it provides a theoretical foundation for the experimentally/empirically focused investigation of preferences. Evolutionary economics is a framework in which this question is addressed.

Attitudes towards risk are most likely affected by education, age and life experience. However, more and more evidence is accumulating about the central role of genetics in determining preferences. Empirical studies have shown that the asset allocation of identical twins are much more correlated than those of twins who are not identical. Furthermore, the asset allocations of identical twins who were raised apart are also highly correlated (see Cesarini et. al. 2009, 2010, and Barnea, Cronqvist, and Siegel 2010). In recent years biologists and economists have identified specific “risk-attitude” genes (see Kuhnen and Chiao 2009, and Zhong et. al. 2009). If preferences are (at least partially) genetically determined, one can view preferences as the result of an evolutionary process selecting for the evolutionary most advantageous risk attitude. This is the approach taken by several studies that have yielded profound

insights into the evolution of preferences<sup>1</sup>, and it is also the approach taken in the present paper.

In an evolutionary context, organisms can be viewed as vessels for carrying their genes (Dawkins 1989). Genes are “successful” if they manage to perpetuate from one generation to the next. Thus, one possible evolutionary “objective function” that can be considered is the maximization of the expected number of offspring – the more offspring, the more copies of the organism’s genes are transmitted to the next generation. While this is a simple and intuitively appealing objective function, it may lead to unreasonable results. To illustrate, consider the following simplified example. Suppose that there are only two possible reproduction prospects or “gambles” to choose from. Gamble A yields 0 offspring with probability 0.3 and 2 offspring with probability 0.7. Gamble B yields 0 offspring with probability 0.8 and 9 offspring with probability 0.2. Suppose also that there are two types of preference genes: gene A that implies the preference of gamble A, and gene B that implies the preference of gamble B.<sup>2</sup> Assume that the gamble realizations are independent across individuals.<sup>3</sup> Then, after  $T$  generations the expected number of individuals with gene A will be  $(0.3 \cdot 0 + 0.7 \cdot 2)^T = 1.4^T$ , while the expected number of individuals with gene B will be  $1.8^T$ . Clearly, after some time the expected number of individuals with gene B becomes much larger than that of gene A, and the ratio of the expected numbers goes to infinity as  $T \rightarrow \infty$ . This may be interpreted as “B dominates the population in the long run”, and it is the motivation for the expected number of offspring criterion.

---

<sup>1</sup> See, for example, Meginniss (1977), Sinn and Weichenrieder (1993), Rogers (1994), Robson (1996, 2001a, 2001b), Samuelson (2001), Curry (2001), Sinn (2003), Schlesinger (2003), and Robson and Samuelson (2007). For an excellent recent review see Robson and Samuelson (2010).

<sup>2</sup> As is typical in this simplified framework, reproduction is assumed to be asexual, i.e. the offspring have the same preferences as their parent (the relaxation of this assumption is discussed in section V).

<sup>3</sup> In addition to this idiosyncratic randomness, individuals and species may also be exposed to “macro” or aggregate environmental risk. Robson (1996) discusses the relation between idiosyncratic risk and aggregate risk in detail. This point is also discussed below.

However, it is far from obvious that this criterion is evolutionary advantageous. Note that in the above example population B has a much larger probability than population A of becoming completely extinct. Let us elaborate. A direct calculation of the probability that A's line becomes extinct is quite cumbersome, because there are infinitely many possible realizations that lead to extinction.<sup>4</sup> Fortunately, it is much easier to solve this problem recursively. Let us denote the probability that the line of descendants of an individual with gene A will become extinct by  $x_A$ . It is possible, with probability 0.3, that the initial individual will have no offspring, and this will imply the end of his line of descendants. However, even if he survives to have 2 offspring, (and this occurs with probability 0.7), it is possible that the lines of both of these offspring will eventually become extinct. As the offspring carry gene A, for each one of them the probability that his line of descendants will eventually become extinct is also by definition  $x_A$ . As the gambles are assumed to be independent, the extinction of one offspring is independent of the extinction of the other, and therefore the probability that both lines eventually become extinct is  $x_A^2$ . Thus,  $x_A$  is the solution to:

$$x_A = 0.3 + 0.7x_A^2, \tag{1}$$

which yields  $x_A = 0.428$ . This value captures all of the possible scenarios leading to eventual extinction. The probability of A Having Descendants Forever,  $p_A( HDF )$ , is given by  $p_A( HDF ) = 1 - x_A = 0.572$ . Similarly, the probability that individual B's line of descendants eventually becomes extinct is given by the solution to:

$$x_B = 0.8 + 0.2x_B^9. \tag{2}$$

---

<sup>4</sup> For example, one possible scenario for extinction is that the original parent has two offspring, each one of these offspring has two offspring, but in the third generation all four offspring die. Of course, there are infinitely many such scenarios leading to extinction, and the probability of extinction,  $x_A$ , is the sum of probabilities for all these events.

Solving eq.(2) numerically yields  $x_B = 0.843$ , which implies  $p_B(HDF) = 1 - x_B = 0.157$ . Thus, while the ratio of the expected population of A to the expected population of B converges to zero as  $T \rightarrow \infty$ , type A has a much higher probability of surviving forever. How can these two facts be reconciled? Note that as  $T$  becomes large the probability distribution of the number of B descendants becomes very skewed – there is a large probability that B will become extinct, but there is a small probability that B will have a very large number of descendants. This extreme low-probability event drives the high expected value of B descendants.

This situation is very reminiscent of the issue of comparing investments for the long run. While the investment with the highest one-period expected return yields the highest expected dollar payoff for any date  $T$ , as  $T \rightarrow \infty$  the investment with the highest geometric mean return almost surely yields a higher dollar payoff than any other investment (including that with the higher expected return). Again, this is due to the skewness of the distribution (see Latané 1959, Samuelson 1971, Hakansson 1971, and Markowitz 1976). Indeed, several authors have suggested the geometric mean growth rate as an evolutionary objective function. Note, however, that if there is some probability of extinction (a total return of 0 in the investment context), the geometric mean is 0, and the comparison of gambles becomes meaningless.<sup>5</sup>

Extinction plays an obvious central role in the evolutionary dynamics. One may suspect, though, that once the population of a given type reaches a certain size the probability of extinction in the i.i.d. reproduction framework is negligible. However, the  $p(HDF)$  criterion may be very important even when the number of individuals in each preference type is large. Consider, for example, the types A and B discussed above. Suppose that there are 100 individuals with risk preference gene A,

---

<sup>5</sup> One way to address this issue is to assume some probability of switching between types. See Robson 1996.

and 100 individuals with gene B. Furthermore, assume that individuals carry many other genes other than the risk preference gene, i.e. each individual represents a unique combination of genes (or alternatively, each “individual” actually represents a sub-type). After many generations, out of the 100 sub-types carrying gene A, on average 57.2 will survive (recall that  $p_A(HDF) = 0.572$ ). In contrast, out of the initial 100 sub-types carrying gene B, on average only 15.7 will survive. Thus, the preference for higher  $p(HDF)$  maintains more genetic diversity, which is an obvious evolutionary advantage. For example, environmental conditions may drastically change, making only a very small number of sub-types viable. The type with more genetic diversity has a higher probability to survive such a change.

Each of the different evolutionary objective functions employed in the literature has its advantages. We do not presume to determine that maximizing  $p(HDF)$  is the “correct” objective function, nor that it is the only one that should be considered. Our standpoint is that the  $p(HDF)$  criterion likely plays an important role in the evolutionary process, and should therefore be given careful consideration. The viewpoint that *both* the number of descendants and the  $p(HDF)$  are important is beautifully captured by the quotation from Isaiah 48:19, given at the beginning of the paper. The blessing is composed of two parts: the first is the promise of many descendants (like the number of grains of sand); the second is the promise of Having Descendants Forever (their name would never be cut off). The two parts have different meaning, and both parts are important.

To the best of our knowledge, the first to introduce the concept of HDF to economics was Meginniss (1977), who analyzed  $p(HDF)$  in the framework of a constant birth probability per unit time and a constant death probability per unit time. Perhaps surprisingly, in the long time that has passed since Meginniss’s innovative

work, the  $p(\text{HDF})$  criterion has not received much attention<sup>6</sup>. The purpose of this paper is to develop the concept of  $p(\text{HDF})$  in a general setting, and to examine the implications of this criterion for the evolutionary foundations of preference.

The structure of this paper is as follows. Section I provides the general framework. In section II we show that maximizing  $p(\text{HDF})$  implies risk averse behavior – if reproduction gamble A dominates gamble B by Second degree Stochastic Dominance (SSD) then  $p_A(\text{HDF}) > p_B(\text{HDF})$ . In other words, individuals who maximize  $p(\text{HDF})$  always act as risk averters. In section III we investigate which of the commonly employed risk-averse utility functions best approximates the goal of maximizing  $p(\text{HDF})$ . We find a very close fit between CRRA preferences and the behavior implied by  $p(\text{HDF})$  maximization. This implies that CRRA preferences can be viewed as a heuristic evolutionary developed to maximize  $p(\text{HDF})$ . Section IV discusses the relationship between the  $p(\text{HDF})$  criterion and loss aversion. Section V extends the analysis to sexual reproduction and mutation. Section VI concludes.

## I. Framework of Analysis

A reproduction “gamble” is given by a set of probabilities corresponding to each possible number of offspring. The gamble  $G$  denoted by:

$$G = (p_0, 0; p_1, 1; \dots, p_N, N) \quad (3)$$

implies probability  $p_0$  for 0 offspring, probability  $p_1$  for 1 offspring, etc. Such a gamble can be viewed as representing a composite lottery resulting from all the decisions faced by the individual throughout reproductive life. A lifelong gamble  $G$  can result from different “local” rules or different dynamic strategies. For example, a local rule could be: “accept a 50-50 chance of losing one offspring or gaining three

---

<sup>6</sup> A few exceptions are Lesourne (1977) and Rubin and Paul (1979).

offspring”. A dynamic strategy could be of the type “be very risk-averse until you have 2 offspring, but once you have 2 offspring become risk-neutral”. Any such local rules or dynamics strategies (and their combinations) can be summarized by a lifelong gamble  $G$ . The relationship between the local rules and dynamic strategies and the resulting lifelong gamble is a potentially complex issue. In most of this paper we focus on the set of lifelong gambles, and the preference among them. In Section IV we discuss the implications of the  $p(\text{HDF})$  criterion to the local concept of loss aversion.

Assume asexual reproduction and that the realizations of gambles are independent across individuals. Consider a set of gambles offered by Nature. The preference of the individual (or his risk-attitude gene) dictates which of these gambles he will choose. Given the choice of a gamble  $G$ , the probability that the individual’s line of descendants will become extinct,  $x$ , is given by the solution to:

$$x = p_0 + p_1x + p_2x^2 + \dots + p_Nx^N, \quad (4)$$

which is a straightforward generalization of the examples described by eq.(1) and (2).<sup>7</sup>

The probability of Having Descendant Forever given the choice of the gamble is given by  $p(\text{HDF}) = 1 - x$ , where  $x$  is the solution to eq.(4). In the general case, there is no analytical solution to eq.(4). However, some general properties of the solution can be determined. First, note that  $x = 1$  is one possible solution (recall that  $p_i$  are probabilities, i.e.  $\sum_{i=0}^N p_i = 1$ ; hence with  $x = 1$  both sides of eq.(4) are equal to 1).

This solution of  $x = 1$  is degenerate – it is the uninteresting case of extinction with certainty. The economically interesting solutions are those for which  $0 < x < 1$ . There is at most a single solution in this range. Figure 1 illustrates why: while the left hand

---

<sup>7</sup> The process (3) is known as a Galton-Watson (1875) branching process. Eq.(4) is the solution to the extinction probability in this process, see for example, Harris (1989) and Athreya and Ney (1972).



side of eq.(4) is the diagonal of the unit square, the right hand side  $f(x) \equiv p_0 + p_1x + p_2x^2 + \dots + p_Nx^N$  is a convex function crossing the y axis at  $p_0$ . In addition,  $f(1) = 1$ , as explained above. Thus,  $f(x)$  crosses the diagonal at most once, as shown in Figure 1. This crossing point, if it exists, is the economically interesting solution (and if it does not,  $x=1$  is the only solution and the gamble leads to extinction with certainty). Such a non-trivial solution exists if and only if the expected number of offspring is greater than 1.<sup>8</sup>

(Please insert Figure 1 about here)

In the next sections we investigate the properties and implications of the solution to eq.(4) in more depth. In Section II we show that maximization of p(HDF) implies risk aversion. In section III we compare the maximization of p(HDF) with the maximization of expected utility with various standard risk-averse utility functions by a numerical examination of the solutions to eq.(4).

## II. p(HDF) and Risk Aversion

From an evolutionary standpoint, the objective function is defined in terms of the number of offspring. Standard economics defines utility in terms of consumption. Of course, there is a close relation between consumption and the number of offspring – the more resources at an individual's disposal, the more offspring he can raise. While the exact relation between the level of consumption and the number of offspring is not obvious, here we adopt the standard, and the most simple, assumption that raising each offspring requires a certain level of consumption,  $C$ , and hence the

---

<sup>8</sup> The crossing point of  $f(x)$  and the diagonal exists iff  $f'(1) > 1$ , i.e.  $f(x)$  approaches the diagonal from below at 1 (see Figure 1).  $f'(1) = p_1 + 2p_2 + \dots + Np_N$ , i.e. it is the expected number of offspring.

number of offspring is proportional to consumption<sup>9</sup>. Thus, the reproduction gamble  $G = (p_0, 0; p_1, 1; p_2, 2; \dots, p_N, N)$  translates into a consumption gamble  $G = (p_0, 0; p_1, C; p_2, 2C; \dots, p_N, NC)$  which yields an expected utility of :

$$EU_G = \sum_{i=0}^N p_i U(i \cdot C). \quad (5)$$

The evolutionary economics framework views the utility function  $U$  as a decision-making tool developed to accommodate the deeper evolutionary objective function of genetic survival. Our purpose here is to link the evolutionary objective of p(HDF) maximization with the properties of the utility function  $U$ .

Theorem 1 :

*If Gamble A dominates Gamble B by Second degree Stochastic Dominance (SSD), then  $p_A(\text{HDF}) > p_B(\text{HDF})$ . Thus, a p(HDF) maximizing agent always acts as a risk-averter.*

Proof:

The proof makes use of the concept of a Probability Preserving Spread (PPS), defined below. Consider the following general reproduction gamble:

$$(p_0, 0; \dots; p_{j-1}, j-1; p_j, j; p_{j+1}, j+1; \dots, p_N, N).$$

Applying a PPS  $\Delta$ -spread to this gamble at  $j$  changes the gamble to:

$$(p_0, 0; \dots; p_{j-1} + \Delta, j-1; p_j - 2\Delta, j; p_{j+1} + \Delta, j+1; \dots, p_N, N).$$

i.e., the PPS shifts probability from  $p_j$  symmetrically to  $p_{j-1}$  and  $p_{j+1}$ .  $\Delta > 0$  is always chosen such that the probabilities remain non-negative after the spread. This concept of a PPS is in the spirit of the Rothschild and Stiglitz (1970) Mean Preserving Spread,

---

<sup>9</sup> This assumption clearly does not hold in modern human societies, where the relationship between wealth and number of offspring is often reverse. However, it is probably a reasonable approximation for the era during which our preferences have evolved, and indeed, this is the assumption made in most studies. For example, Sinn (2003) writes: "... it is assumed that the number of children a parent has is proportional to the amount of resources he commands".

with the difference that they apply the spread to the outcomes, keeping the mean outcome invariant, whereas here we apply the spread to the probabilities, keeping the total probability 1 (but the mean outcome may change).

Theorem 1 is based on the following two lemmas, proven in the Appendix:

Lemma 1: If reproduction gamble A dominates gamble B by SSD, then A can be transformed into B by a series of PPS, plus a shift of probability from the outcome of having  $N$  offspring to the outcome of having only  $N-1$  offspring.

Lemma 2: A PPS decreases the probability of having descendants forever.

Thus, starting with gamble A we reach gamble B by a series of PPS, where each PPS decreases  $p(\text{HDF})$ , and finally by a shift of probability from the outcome of having  $N$  offspring to the outcome of having only  $N-1$  offspring, which again also decreases the  $p(\text{HDF})$ . Therefore, if A dominates B by SSD it follows that  $p_A(\text{HDF}) > p_B(\text{HDF})$ .

The above statement is true for *any* two gambles where A SSD dominates B. This implies that if the choices of a  $p(\text{HDF})$  maximizing agent are always consistent with those of a risk-avertter: if A is preferred to B by all risk-avertters, it is guaranteed that the  $p(\text{HDF})$  maximizing agent will also choose A.

Q.E.D.

The intuition of this result is that shifting probability from the outcome of, say, three offspring to the outcome of four offspring increases the  $p(\text{HDF})$ , but not as much as the decrease in  $p(\text{HDF})$  caused by shifting the same probability from the three-offspring outcome to the two-offspring outcome. This resonates with the notion of decreasing marginal utility, but note that in the evolutionary framework things are somewhat more complex, as we can speak in terms of “the utility from an additional

offspring”, only in the statistical sense of the impact of changing the probability for a particular number of offspring on  $p(\text{HDF})$ .

Theorem 1 implies that individuals genetically “programmed” to maximize  $p(\text{HDF})$  will act as risk-averters: they will always choose reproduction gambles from the SSD efficient set. However,  $p(\text{HDF})$  maximization actually gives us more information about preferences. Given a set of gambles, the SSD efficient set will typically include several of these gambles (see, for example, Levy 2006). However, any individual maximizing  $p(\text{HDF})$  will choose a single specific gamble from this efficient set – the gamble that maximizes the  $p(\text{HDF})$ .<sup>10</sup> Is there a risk-averse utility function  $U$  that will always yield the gamble with maximal  $p(\text{HDF})$  as the gamble maximizing expected utility? In other words, other than the general property of risk-aversion, can  $p(\text{HDF})$  maximization give us more information on preference? To this question we turn next.

### **III. $p(\text{HDF})$ and Constant Relative Risk Aversion**

The evolutionary framework of maximizing  $p(\text{HDF})$  is fundamentally different than the standard expected utility framework. In the latter, individuals are endowed with some utility function defined over consumption (or possibly over the number of offspring), and they act to maximize their expected utility. In the former, there is no utility function, only a criterion to choose among a set of possible gambles. Can these two very different frameworks be reconciled? Is there some utility function  $U$  that when plugged into eq.(5) will always imply the same choices among prospects as maximization of  $p(\text{HDF})$ ? Stated differently, suppose that evolution has indeed shaped us to maximize  $p(\text{HDF})$ . Now, when we attempt to describe behavior with the

---

<sup>10</sup> Except in the unlikely event that two gambles yield exactly the same  $p(\text{HDF})$ . In this case the individual will be indifferent between the two gambles.

expected utility framework, what utility function will provide the best description of behavior?

Unfortunately, as there is no general analytical solution to eq.(4), we cannot analytically derive a utility function that will perfectly mimic p(HDF) maximization. However, we can numerically examine the choice implied by p(HDF) maximization and compare it with the choice implied by various utility functions typically employed in the literature. This is the idea of the analysis described below.

We employ two approaches to investigate the relation between various utility functions and the p(HDF) criterion. In the first approach, we randomly draw 1,000 reproduction gambles of type (3), (i.e. 1,000 vectors of p's), and for each gamble we numerically calculate both the p(HDF) and the expected utility for various utility functions. We take the maximal possible number of offspring as  $N=10$ , and we draw each one of the p's independently from a uniform distribution on  $[0,1]$ . After the 11 p's are drawn, they are normalized so that  $\sum_{i=0}^N p_i = 1$ .<sup>11</sup> As the choice is invariant to a linear transformation of the utility function in the expected utility framework, we are not concerned with the absolute EU value, but only with the correlation between the EU value (as given by eq.(5)) and the p(HDF). The higher the correlation, the closer the utility function fits the p(HDF) criterion. In the extreme, a correlation of +1 implies that the utility function perfectly describes the ranking and choice by the p(HDF) criterion.

We examine the utility functions that are most commonly employed in the economic literature: linear preference, the negative exponential, the power utility function, and the quadratic utility function. While the negative exponential is popular because of its analytical tractability, it implies Constant Absolute Risk Aversion

---

<sup>11</sup> E.g. after all the p's are drawn, we divide each one of them by  $\sum_{i=0}^N p_i$ .

(CARA), which is considered by most as unrealistic.<sup>12</sup> The power utility function, which includes the logarithmic utility as a special case, implies Constant Relative Risk Aversion (CRRA). This is the utility function for which there is the most empirical and experimental support (see, for example, Friend and Blume 1975, Kroll and Levy 1992, and Levy 1994), and it is probably considered the most realistic. The quadratic function has the advantage of justifying the mean-variance framework even when the return distributions are not normal, but it has the very undesirable implication of Increasing Absolute Risk Aversion (IARA). A well-known alternative to the EU framework is Prospect Theory (PT). In PT the decisions are based not on absolute values (of money, consumption, or the number of offspring), but rather on *changes* of these values relative to the status quo. Thus, the PT framework can be thought of as a “local” rule, rather than a rule for choosing among lifelong alternatives. We discuss the relation between such local rules and the resulting lifelong distribution in the next section.

Figure 2 describes EU as a function of p(HDF) for the above discussed utility functions.<sup>13</sup> While there is a general positive correlation between EU and p(HDF) for all four utility functions, for the linear, exponential, and quadratic functions (panels a-c) the fit is far from perfect. The results shown here are typical for the results obtained with other parameter values.<sup>14</sup> In contrast to these results, for the CRRA preference

---

<sup>12</sup> See, for example, Markowitz, Reid, and Tew 1994.

<sup>13</sup> For example, consider a prospect G as given by eq.(3). Employing eq.(5) to calculate the Expected Utility for linear preferences yields:  $EU_G = C \sum_{i=0}^N p_i \cdot i$ . For CRRA preferences we obtain:

$EU_G = \frac{C^\gamma}{\gamma} \sum_{i=0}^N p_i \cdot i^\gamma$ , etc. In all cases the value of C does not affect the correlation between EU and p(HDF), and we take this value as C=1.

<sup>14</sup> Note that for the quadratic utility function  $U(W) = W - bW^2$  the parameter  $b$  must not exceed  $1/2W_{max}$ , otherwise utility is declining with wealth. Here the maximum outcome is N=10, and we must have  $b \leq 1/20$ . Figure 1c shows the results for the maximal b value, but similar results are obtained

the correlation is very high. Thus, CRRA preference is an almost perfect proxy for p(HDF) maximization.

In order to check the robustness of this result, we first examine the fit over a larger range of p(HDF) values. In Figure 2, we have  $p(\text{HDF}) > 0.7$  for all gambles. In order to examine prospects with higher extinction probabilities, we extend the range from which  $p_0$  is drawn<sup>15</sup>. The results are shown in Figure 3. With the potential for high  $p_0$  values, we now obtain prospects with p(HDF) over the entire possible range between 0 and 1. As panel a shows, the fit to the CRRA utility function remains very good across the entire range.

The next question is what is the best RRA parameter  $\gamma$  to use. Note that in our framework we must have  $\gamma > 0$ , because one of the possible outcomes is 0, and the CRRA function is not defined in this case if  $\gamma \leq 0$ . Thus, the value of  $\gamma = 0.05$  in panel a represents the lower end of possible  $\gamma$  values. Panels b and c show the results for the values  $\gamma = 0.3$  and  $\gamma = 0.8$  respectively. The fit remains very good for  $\gamma = 0.3$ , but for higher values of  $\gamma$  the fit deteriorates for gambles with high p(HDF). These and similar experiments with other  $\gamma$  values lead us to conclude that CRRA preferences fit p(HDF) maximization very well for RRA values in the range of  $0 < \gamma < 0.5$ .

(Please insert Figures 2 and 3 about here)

In the second approach we employ, we analyze the probability that employing the EU criterion leads to a wrong choice by the p(HDF) criterion. To do so, we randomly draw two gambles, and check whether the EU ranking is the same as the p(HDF) ranking. We report the percentage of times that errors are made by employing

---

with lower values. Note that for  $b \rightarrow 0$  the quadratic preference converges to the linear preference shown in Figure 1a.

<sup>15</sup> We take the range of  $p_0$  as  $[0,8]$  before normalization, see footnote 11.

the EU criterion. Obviously, if the  $p(\text{HDF})$  of one prospect is much higher than that of the other, we can say that one prospect is “much better” than the other, and it is not very likely that the EU criterion will imply a reverse ranking. Generally, we would expect that the closer the prospects are in terms of their  $p(\text{HDF})$ , the more likely that EU will lead to a mistake. In Figure 4 we report the error rates as a function of the difference  $\Delta p$  between the  $p(\text{HDF})$ 's of the two prospects,  $\Delta p \equiv |p_A(\text{HDF}) - p_B(\text{HDF})|$ . As expected, the error rate decreases with  $\Delta p$  for all utility functions. More importantly, the error for the CRRA preference is much lower than the error rates of all the other utility functions, and this holds for all values of  $\Delta p$ . This result is consistent with and reinforces the previous findings – EU maximization with CRRA preferences provides a very good approximation for  $p(\text{HDF})$  maximization.

(Please insert Figure 4 about here)

CRRA preferences are considered by many researchers as the most realistic description of behavior within the expected utility framework, and it is the preference assumed in many cornerstone economic models. Thus, it is very encouraging, and also quite surprising, to find that this preference can be viewed as a result of the evolutionary process of selecting for  $p(\text{HDF})$  maximization.

#### **IV. Loss Aversion and $p(\text{HDF})$**

Expected utility is defined in terms of the *absolute* level of consumption (or wealth). Similarly, the  $p(\text{HDF})$  maximization framework is defined in terms of the *absolute* number of offspring. These absolute values are appropriate for the choice between lifelong alternatives. In contrast, the concept of loss-aversion, a key element of the celebrated Prospect Theory, is defined in terms of *changes* relative to a given



status-quo reference point. The loss-aversion framework generally contradicts the expected utility framework. However, as we shall see below, it may be consistent with the p(HDF) framework, and it can even be viewed a “local” heuristic implied by p(HDF) maximization. The key difference between the lifelong framework and the local framework is that in the local framework the gamble is considered as a unique one-time idiosyncratic opportunity that is not expected to repeat, and does not affect the p(HDF) of each of the offspring. The difference between the lifelong framework and the local framework is similar to the difference between the lifelong expected utility maximization framework and the local Prospect Theory heuristic.

The most simple formulation of loss aversion is by a piecewise-linear value function of the type:

$$V(x) = \begin{cases} x & \text{for } x \geq 0 \\ \lambda x & \text{for } x < 0 \end{cases} \quad (6)$$

where  $x$  is the change of wealth relative to the current wealth,  $\lambda > 1$  is the loss-aversion parameter, and the objective is the maximization of  $E[V(x)]$ .<sup>16</sup>

Consider a prospect that yields  $+C$  with probability  $p$  or  $-C$  with probability  $(1-p)$ , where  $C$  is the consumption needed to grow one offspring. Then, according to the loss-averse value function (6), this prospect will be accepted if and only if :

$$\frac{\lambda}{1+\lambda} < p. \quad (7)$$

---

<sup>16</sup> Benartzi and Thaler (1995) employ this streamlined formulation. Prospect Theory (Kahneman and Tversky 1979, Tversky and Kahneman 1992) models  $V(x)$  as concave for  $x > 0$  and convex for  $x < 0$ , and also allows for subjective probability weighting. We abstract from these features here for the sake of simplicity.

<sup>17</sup> The expected value with the gamble is:  $EV = pC - (1-p)\lambda C$ . The gamble is accepted if and only if this  $EV$  is larger than 0 (the expected value of the status quo), i.e. if  $p > \lambda(1-p)$ , or  $\frac{\lambda}{1+\lambda} < p$ .

For any loss-averse individual ( $\lambda > 1$ ) the prospect will not be accepted unless  $1/2 < p$ . The more risk averse the individual, the higher the  $p$  required in order for the gamble to be accepted.

Now consider this prospect in the p(HDF) framework. If the gamble can be repeated many times, it is clear that any gamble with  $p > 1/2$  should be accepted, because of the law of large numbers. Thus, it only makes sense to think of this gamble as a unique one-time opportunity. This is in contrast to the lifelong gamble in eq.(3) and the formulation of eq.(4), where preferences for gambles affect not only the parent, but also all of his descendants, i.e. they all have the same  $x$ . Here, the acceptance/rejection of the one-time gamble is assumed not to affect the extinction probability,  $x$ , of the offspring. These distinctions are at the heart of the difference between the lifelong-decision framework as discussed in the previous section and the expected utility framework, and the “local” one-time gamble framework discussed here.

Consider an individual that has  $n$  offspring, each with a probability of  $1-x$  of HDF and a probability of  $x$  of having his line of descendants eventually becoming extinct. The individual is offered the above idiosyncratic one-time gamble. If the gamble is not taken, the individual has a probability of  $1 - x^n$  of HDF. If the gamble is accepted, with probability  $p$  the individual will have one more offspring, and with probability  $(1-p)$  he will have one less offspring. Thus, if the gamble is taken p(HDF) becomes  $1 - [px^{n+1} + (1-p)x^{n-1}]$ . The gamble is accepted if and only if it increases p(HDF), i.e. if:

$$1 - [px^{n+1} + (1-p)x^{n-1}] > 1 - x^n \quad (8)$$

i.e if

$$x^n > [px^{n+1} + (1-p)x^{n-1}]$$

or:

$$x > px^2 + (1 - p). \quad (9)$$

First, note that the acceptance/rejection of the prospect does not depend on the initial number of offspring,  $n$ . This is a non-trivial point, which conforms with the loss-aversion notion of relative rather than absolute levels. The decision depends only on the probability  $p$  for a favorable outcome in the gamble, and the probability of each offspring eventually becoming extinct,  $x$ . Solving (9) we find that the gamble is accepted if and only if:

$$\frac{1}{1+x} < p. \quad (10)$$

For  $p \leq 1/2$  the gamble is refused for any value of  $0 < x < 1$ , i.e. if the gamble is fair (or worse) it is refused, in line with the notion of loss aversion.

Eq.(10) implies that the smaller  $x$  is, the larger the probability  $p$  required to make the gamble acceptable. The intuition for this is as follows. An addition of one offspring increases  $p(\text{HDF})$  less than the reduction of one offspring diminishes it. This is why a fair gamble is not accepted, and why  $p$  must be larger than  $1/2$  for the gamble to be accepted. If  $x$  is close to 1, the difference between  $x^{n-1}$  and  $x^{n+1}$  is not very large, and therefore  $p$  does not have to exceed  $1/2$  by much. However, if  $x$  is small the above difference is (relatively) large, and  $p$  must be large to compensate. In other words, when most offspring are likely to have descendants forever, an additional

---

<sup>18</sup> Eq.(9) can be written as:

$px^2 - x + (1 - p) < 0$ . The expression on the r.h.s. is an upward facing parabola, so the  $x$ 's solving eq.(9) are those between the two roots of the equation  $px^2 - x + (1 - p) = 0$ . These roots are:

$$x_{1,2} = \frac{1 \pm \sqrt{1 - 4p(1 - p)}}{2p} = \frac{1 \pm \sqrt{(2p - 1)^2}}{2p} = \frac{1 \pm (2p - 1)}{2p} \Rightarrow x_1 = \frac{1 - p}{p} \quad x_2 = 1.$$

Thus, any  $x$  in the relevant range  $0 < x < 1$  that satisfies  $x > (1 - p)/p$  also satisfies eq.(9). Rearranging, we obtain:

$$\frac{1}{1+x} < p.$$

offspring does not contribute much to  $p(\text{HDF})$ , but a reduction of one offspring does have a large effect on  $p(\text{HDF})$ . Thus, a large  $p$  is required to make the gamble acceptable.

Comparing eqs.(7) and (10) reveals that  $\frac{1}{1+x}$  plays the role of  $\frac{\lambda}{1+\lambda}$ .

Simplifying the relation between  $x$  and  $\lambda$ , we obtain:

$$\frac{1}{1+x} = \frac{\lambda}{1+\lambda} \Rightarrow x = \frac{1}{\lambda}. \quad (11)$$

Hence,  $\frac{1}{x}$  plays the role of the loss-aversion parameter  $\lambda$ .

One may view loss-aversion as a “local” heuristic aimed at maximizing  $p(\text{HDF})$ . This may suggest an evolutionary explanation for loss-aversion. Experimental and empirical estimates of the loss-aversion parameter  $\lambda$  typically yield values close to 2.2 (see Kahneman, Knetsch and Thaler (1990), Tversky and Kahneman (1991,1992), Pennings and Smidts (2003), Levy (2009)). This value translates by eq.(11) to an  $x$  value of  $x = 1/2.2 = 0.45$ . This suggests that loss-aversion has evolved in an era where the typical value of  $x$  was approximately 0.45, i.e. the value of  $p(\text{HDF})$  was about 0.55. This value is close to the average value of 0.54 empirically estimated for Mexico and Israel by Keyfitz and Tyree (1967).<sup>19</sup>

## V. Mutation and Sexual Reproduction

Up to this point we have assumed asexual reproduction and no mutations, i.e. we have assumed that the extinction probability,  $x$  (and its complimentary  $p(\text{HDF})=1-x$ ), are the same for the parent and its offspring. Sexual reproduction and mutation

---

<sup>19</sup> They estimate a lower  $p(\text{HDF})$  value of 0.20 for the U.S. In general, the average number of children per family, and thus the  $p(\text{HDF})$ , is higher in developing countries relative to developed countries. The former may be more representative of the era over which preferences have evolved.

complicate the analysis, because they imply that the offspring's  $x$  may be different than that of the parent. These effects can be modeled by:

$$x_{\text{offspring}} = x_{\text{parent}} + \tilde{\varepsilon}, \quad (12)$$

where  $\tilde{\varepsilon}$  represents the random effects of mutation and the effects of the other parent's genes (not only in the present generation, but all the way down the road – recall that  $x$  is the probability of extinction of an individual's entire line of descendants). We assume that  $\tilde{\varepsilon}$  is uncorrelated across offspring. Let us initially also assume that  $\tilde{\varepsilon}$  is unbiased, i.e. that  $E(\tilde{\varepsilon}) = 0$ .

Let us consider first the most simple case where there is at most one offspring, and that  $\tilde{\varepsilon}$  can take the values  $+\varepsilon$  or  $-\varepsilon$  with equal probabilities of 1/2. In this case the parent's probability of extinction is given by:

$$x = \frac{1}{2}[p_0 + p_1(x + \varepsilon)] + \frac{1}{2}[p_0 + p_1(x - \varepsilon)] = p_0 + p_1x.$$

In the general case, the parent's probability of extinction is the solution to:

$$(13)$$

$$x = p_0 + p_1E[x + \tilde{\varepsilon}_{1,1}] + p_2E[x + \tilde{\varepsilon}_{2,1}]E[x + \tilde{\varepsilon}_{2,2}] + \dots + p_NE[x + \tilde{\varepsilon}_{N,1}] \cdot \dots \cdot E[x + \tilde{\varepsilon}_{N,N}]$$

where the subscripts of the  $\varepsilon$ 's denote the different random realization for each offspring. As the  $\varepsilon$ 's are uncorrelated and unbiased ( $E(\tilde{\varepsilon}) = 0$ ) we have:

$$x = p_0 + p_1x + p_2x^2 + \dots + p_Nx^N. \quad (14)$$

This is exactly the same equation as equation (4), obtained with no mutation or sexual reproduction. Thus, unbiased mutation and sexual reproduction does not change the p(HDF) analysis discussed in the preceding sections.

If the errors are biased, i.e.  $E(\tilde{\varepsilon}) \neq 0$ , eq. (13) can be written as:

$$x = p_0^* + p_1^*x + p_2^*x^2 + \dots + p_N^*x^N, \quad (15)$$

where all the  $E(\tilde{\varepsilon})$  terms are absorbed in the new  $p^*$ 's. Eq.(15) again has a very similar structure to eq.(4), however, note that  $\sum_{i=0}^N p_i^*$  may generally be different than 1.

## VI. Conclusion

Evolution is a powerful and profound process that can explain many of the observed characteristics of plants, animals and humans. An innovative strand of the economics literature suggests viewing preferences through the perspective of evolution. Most of the studies in this strand take the evolutionary objective function as the maximization of the expected number of offspring, or alternatively, as the maximization of the geometric mean growth rate. While these objectives are appealing and intuitive, in this paper we suggest that careful consideration should also be given to another evolutionary objective function: maximization of the probability of Having Descendants Forever (p(HDF)). This objective captures elements that are absent in the more standard objective functions, elements that are likely to play an important role in the evolutionary process, especially when extinction is possible. Thus, we view the HDF approach as complementary to the approaches of maximizing the expected number of offspring or geometric mean growth rate. The present paper explores some of the implications of p(HDF) maximization for preferences.

Our main findings are as follows. First, we show that p(HDF) maximization implies risk-aversion. This is different than the risk-aversion resulting from a convex “production function” for offspring as suggested by some studies. Here risk-aversion is implied even when the number of offspring is linear in the consumption. Next, we compare p(HDF) maximization with the standard risk-averse utility functions. We find very close agreement between the p(HDF) criterion and CRRA preferences.

Thus, one may view CRRA preferences as a proxy for an underlying evolutionary goal of maximizing  $p(\text{HDF})$ . These results hold not only in the standard framework of asexual reproduction, but also in the case of unbiased sexual reproduction and mutation.

While the  $p(\text{HDF})$  framework is defined in terms of the lifelong distribution of the number of offspring, one can also examine the implications of this framework to “local” decisions about idiosyncratic one-time prospects. This analysis reveals that  $p(\text{HDF})$  maximization implies loss-aversion. Furthermore, there is a direct link between the  $p(\text{HDF})$  and the loss-aversion parameter  $\lambda$ . The experimental and empirical estimates of  $\lambda \approx 2.2$  imply a  $p(\text{HDF})$  of about 0.55. While this value seems low by modern Western standards, it may be reasonable for the era during which our preferences have evolved.

The concept of risk-aversion has been discussed for hundreds of years. Loss-aversion has been formalized and documented decades ago. While these properties of preference are usually taken as exogenously given, evolutionary economics views them as a result of a deeper evolutionary force. Though the evolutionary economics approach is only in its first stages of development, it seems that it holds great promise of offering a new and different perspective for understanding attitudes towards risk.

## References

- Athreya, K.B., and Ney, P., 1972. *Branching Processes*. New York: Springer Verlag.
- Benartzi, S., and Thaler, R., 1995. "Myopic Loss Aversion and the Equity Premium Puzzle", *Quarterly Journal of Economics*, 110, 73-92.
- Cesarini, D., Dawes, C.T., Johannesson, M., Lichtenstein, P., and Wallace, B., 2009. "Genetic Variation in Preferences for Giving and Risk Taking". *Quarterly Journal of Economics* 124, 809-842.
- Cesarini, D., Johannesson, M., Lichtenstein, P., Sandewall, O., and Wallace, B., 2010. "Genetic Variation in Financial Decision Making". *Journal of Finance*, forthcoming.
- Curry, P.A., 2001. "Decision Making under Uncertainty and the Evolution of Interdependent Preferences", *Journal of Economic Theory* 98, 357-369.
- Dawkins, R., 1989. *The Selfish Gene*. Oxford: Oxford University Press.
- Friend, I., and Blume, M.E., 1975. "The Demand for Risky Assets". *American Economic Review* 65, 900-922.
- Hadar, J., and Russell, W., 1969. "Rules for Ordering Uncertain Prospects". *American Economic Review* 59, 25-34.
- Hakansson, N.H., 1971. "Capital Growth and the Mean-Variance Approach to Portfolio Selection". *Journal of Financial and Quantitative Analysis* 6, 517-557.
- Hanoch, G., and Levy, H., 1969. "The Efficiency Analysis of Choices Involving Risk". *Review of Economic Studies* 36, 335-346.
- Harris, T.E., 1989. *The Theory of Branching Processes*. Toronto: Dover.
- Kahneman, D., Knetsch, J.L., and Thaler, R.H., 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem". *Journal of Political Economy* 98, 1325-1348.
- Kahneman, D., and Tversky, A., 1979. "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47, 263-292.
- Keyfitz, N., and Tyree, A., 1967. "Computerization of the Branching Process". *Behavioral Science* 12, 329-336.
- Kroll, Y., and Levy, H., 1992. "Further Tests of the Separation Theorem and the Capital Asset Pricing Model". *American Economic Review* 82, 664-670.
- Kuhnen, C.M., and Chiao, J., 2009. "Genetic Determinants of Financial Risk Taking". *PLoS ONE* 4, 1-4.



- Latané, H.E., 1959. "Criteria for Choice Among Risky Ventures". *Journal of Political Economy* 57, 144-155.
- Lesourne, J., 1977. *A Theory of the Individual for Economic Analysis*, Amsterdam.
- Levy, H., 1994. "Absolute and Relative Risk Aversion: An Experimental Study," *Journal of Risk and Uncertainty* 8, 289-307.
- Levy H. 2006. *Stochastic Dominance*. Boston: Kluwer Academic Publishers.
- Levy, M., 2009. "Loss Aversion and the Price of Risk". *Quantitative Finance*, 1-14.
- Markowitz, H.M., 1976. "Investment for the Long Run: New Evidence for an Old Rule". *Journal of Finance* 31, 1273-1286.
- Markowitz, H.M., Reid, D.W., and Tew, B.V., 1994. "The Value of a Blank Check". *Journal of Portfolio Management* 20, 82-91.
- Meginniss, J.R. 1977. "Alternatives to the Expected Utility Rule". Unpublished dissertation, University of Chicago.
- Robson, A.J., 1996. "A Biological Basis for Expected and Non-Expected Utility". *Journal of Economic Theory* 68, 397-424.
- Robson, A.J., 2001a. "The Biological Basis of Economic Behavior". *Journal of Economic Literature* 39, 11-33.
- Robson, A.J., 2001b. "Why Would Nature Give Individuals Utility Functions?". *Journal of Political Economy* 109, 900-914.
- Robson, A.J., and Samuelson, L., 2007. "The Evolution of Intertemporal Incentives". *American Economic Review* 97, 492-495.
- Robson, A.J., and Samuelson, L., 2010. "The Evolutionary Foundations of Preferences". In *The Social Economics Handbook*. Benhabib, J., Bisin, A., and Jackson, M., (Eds.). Elsevier Press, forthcoming.
- Rogers, A.R., 1994. "Evolution of Time Preference by Natural Selection". *American Economic Review* 84, 460-481.
- Rothschild, M., and Stiglitz., J., 1970. "Increasing Risk: I. A definition", *Journal of Economic Theory* 2, 225-243.
- Rubin, P.H., and Paul, C.W., 1979. "An Evolutionary Model of Taste for Risk". *Economic Inquiry* 17, 585-596.
- Samuelson, L., 2001. "Introduction to the Evolution of Preferences". *Journal of Economic Theory* 97, 225-230.

Samuelson, P.A., 1971. "The Fallacy of Maximizing the Geometric Mean in Long Sequences in Investing or Gambling". *Proceedings of the National Academy of Science* 68, 2493-2496.

Schlesinger, H., 2003. "Some Remarks on the Evolution of Risk Preferences". *Geneva Papers on Risk and Insurance Theory* 28, 101-104.

Sinn, H.W., 2003. "Weber's Law and the Biological Evolution of Risk Preferences". *Geneva Papers on Risk and Insurance Theory* 28, 87-100.

Sinn, H.W., and Weichenrieder, A., 1993. "Biological Selection of Risk Preferences". In *Risk as a Construct*. Bayerrische Ruck (Ed.), Munich: Knesbeck.

Tversky, A. and Kahneman, D., 1991. "Loss Aversion in Riskless Choice: A Reference Dependent Model", *Quarterly Journal of Economics* 106, 1039-1061.

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

Watson, H.W., and Galton, F., 1875. "On the Probability of the Extinction of Families", *Journal of the Anthropological Institute of Great Britain* 4, 138-144.

Zhong, S., Israel, S., Xue, H., Ebstein, R.P., and Chew, S.H., 2009. "Monoamine Oxidase A Gene (MAOA) Associated with Attitude Towards Longshot Risks". *PLoS ONE* .

## Appendix

### Proof of Lemma 1:

Consider the following two prospects:

$$F=(p_0,0; p_1,1; \dots, p_N, N) \quad \text{and} \quad G=(q_0,0; q_1,1; \dots, q_N, N)$$

(if  $N$  is not the same for the two prospects, one can always take the maximal  $N$ , and assign probabilities of zero where necessary). Assume that  $F$  dominates  $G$  by Second degree Stochastic Dominance (SSD). This implies that:

$$\int_0^j (G(t) - F(t)) dt \geq 0 \quad (A1)$$

for all  $j$ , where  $F(t)$  and  $G(t)$  are the cumulative distribution functions of  $F$  and  $G$ , respectively (see, Hanoch and Levy 1969, Rothschild and Stiglitz 1970, Hadar and Russel 1969, and Levy 2006). For the discrete prospects considered here, we have:

$$\int_0^j F(t) dt = jp_0 + (j-1)p_1 + (j-2)p_2 + \dots + 2p_{j-2} + p_{j-1} \quad (A2)$$

(see Figure A1 for a graphical illustration of eq.(A2)). Thus, the SSD condition for the dominance of  $F$  over  $G$  translates to: (A3)

$$jp_0 + (j-1)p_1 + (j-2)p_2 + \dots + 2p_{j-2} + p_{j-1} \leq jq_0 + (j-1)q_1 + (j-2)q_2 + \dots + 2q_{j-2} + q_{j-1}$$

for all  $j=1, \dots, N$ . We can write these conditions as:

$$\begin{aligned} p_0 &= q_0 - \Delta_0 \\ 2p_0 + p_1 &= 2q_0 + q_1 - \Delta_1 \\ 3p_0 + 2p_1 + p_2 &= 3q_0 + 2q_1 + q_2 - \Delta_2 \\ &\vdots \\ Np_0 + (N-1)p_1 + \dots + p_{N-1} &= Nq_0 + (N-1)q_1 + \dots + q_{N-1} - \Delta_{N-1} \end{aligned} \quad (A4)$$

The SSD condition as written in eq.(A3) ensures that all of the above  $\Delta$ 's are positive.

We prove that one can get from  $G$  to  $F$  by a series of PPS by constructing the required PPS's. First, note that from the first equation in (A4) we have  $p_0 = q_0 - \Delta_0$ . Plugging

this expression into the second equation in (A4) and rearranging yields:

$p_1 = q_1 + 2\Delta_0 - \Delta_1$ . Plugging this expression for  $p_1$  in the third equation in (A4)

yields:  $p_2 = q_2 - \Delta_0 + 2\Delta_1 - \Delta_2$ . Repeating this for the fourth equation in (A4) we

obtain:  $p_3 = q_3 - \Delta_1 + 2\Delta_2 - \Delta_3$ . Similarly, it is easy to show by induction that

$p_j = q_j - \Delta_{j-2} + 2\Delta_{j-1} - \Delta_j$  for any  $2 \leq j \leq N-1$ .<sup>20</sup> Thus, prospect  $F$  can be written

as:

$$F = \begin{cases} p_0 = q_0 - \Delta_0 \\ p_1 = q_1 + 2\Delta_0 - \Delta_1 \\ p_2 = q_2 - \Delta_0 + 2\Delta_1 - \Delta_2 \\ p_3 = q_3 - \Delta_1 + 2\Delta_2 - \Delta_3 \\ \vdots \\ p_{N-1} = q_{N-1} - \Delta_{N-3} + 2\Delta_{N-2} - \Delta_{N-1} \\ p_N = q_N - \Delta_{N-2} + \Delta_{N-1} \end{cases}$$

(Note that unlike the other  $p$ 's,  $p_N$  is not given explicitly as a function of  $q_N$ . Instead,

from the equality  $\sum_{i=1}^N p_i = \sum_{i=1}^N q_i = 1$  and the preceding equations we obtain

$p_N = q_N - \Delta_{N-2} + \Delta_{N-1}$ ). Let us demonstrate how  $G$  can be obtained by applying a

series of PPS to  $F$ . The first PPS to be applied is subtracting  $2\Delta_0$  from  $p_1$  and adding

$\Delta_0$  to  $p_0$  and to  $p_2$ . This first PPS yields the following prospect, denoted by the

superscript 1:

$$F^1 = \begin{cases} p_0^1 = q_0 \\ p_1^1 = q_1 - \Delta_1 \\ p_2^1 = q_2 + 2\Delta_1 - \Delta_2 \\ p_3^1 = q_3 - \Delta_1 + 2\Delta_2 - \Delta_3 \\ \vdots \\ p_N^1 = q_N - \Delta_{N-2} + \Delta_{N-1} \end{cases}$$

---

<sup>20</sup> Assume that  $p_i = q_i - \Delta_{i-2} + 2\Delta_{i-1} - \Delta_i$  for all  $1 < i < j$ , and that  $p_0 = q_0 - \Delta_0$  and  $p_1 = q_1 + 2\Delta_0 - \Delta_1$ . Plugging these expressions in the  $j+1^{\text{th}}$  equation in (A4) yields  $p_j = q_j - \Delta_{j-2} + 2\Delta_{j-1} - \Delta_j$ , as required.

The next PPS to apply is subtracting  $2\Delta_1$  from  $p_2^1$  and adding  $\Delta_1$  to  $p_1^1$  and to  $p_3^1$ .

This yields the new prospect  $F^2$ :

$$F^2 = \begin{cases} p_0^2 = q_0 \\ p_1^2 = q_1 \\ p_2^2 = q_2 - \Delta_2 \\ p_3^2 = q_3 + 2\Delta_2 - \Delta_3 \\ \vdots \\ p_N^2 = q_N - \Delta_{N-2} + \Delta_{N-1} \end{cases}$$

Similarly, subtracting  $2\Delta_2$  from  $p_3^2$  and adding  $\Delta_2$  to  $p_2^2$  and to  $p_4^2$  yields:

$$F^3 = \begin{cases} p_0^3 = q_0 \\ p_1^3 = q_1 \\ p_2^3 = q_2 \\ p_3^3 = q_3 - \Delta_3 \\ p_4^3 = q_4 + 2\Delta_3 - \Delta_4 \\ \vdots \\ p_N^3 = q_N - \Delta_{N-2} + \Delta_{N-1} \end{cases}$$

This is repeated until the last PPS at which stage we have:

$$F^{N-1} = \begin{cases} p_0^{N-1} = q_0 \\ p_1^{N-1} = q_1 \\ p_2^{N-1} = q_2 \\ \vdots \\ p_{N-2}^{N-1} = q_{N-2} \\ p_{N-1}^{N-1} = q_{N-1} - \Delta_{N-1} \\ p_N^{N-1} = q_N + \Delta_{N-1} \end{cases}$$

Finally, the last transformation required to obtain prospect  $G$  is a shift of probability of  $\Delta_{N-1}$  from  $p_N^{N-1}$  to  $p_{N-1}^{N-1}$ , i.e. from the outcome of having  $N$  offspring to the outcome of having only  $N-1$  offspring.

Proof of Lemma 2:

Consider the effect on  $x$  of slightly increasing  $p_j$  and decreasing  $p_0$  by the same amount. Expressing  $p_0$  as  $1 - \sum_{i=1}^N p_i$ , equation (4) can be written as:

$$x = 1 - \sum_{i=1}^N p_i + p_1 x + p_2 x^2 + \dots + p_N x^N. \quad (\text{A5})$$

Differentiating (A5) with respect to  $p_j$  we obtain:

$$\frac{\partial x}{\partial p_j} = -1 + p_1 \frac{\partial x}{\partial p_j} + 2p_2 x \frac{\partial x}{\partial p_j} + 3p_3 x^2 \frac{\partial x}{\partial p_j} + \dots + Np_N x^{N-1} \frac{\partial x}{\partial p_j} + x^j.$$

Rearranging yields:

$$\frac{\partial x}{\partial p_j} = \frac{x^j - 1}{\left[1 - (p_1 + 2p_2 x + 3p_3 x^2 + \dots + Np_N x^{N-1})\right]}. \quad (\text{A6})$$

Note that the denominator in eq.(A6) does not depend on  $j$  (i.e. it is the same regardless of the  $p_j$  with respect to which  $x$  is differentiated). Furthermore, this expression is positive – it is equal to  $1 - f'(x)$ , where  $f(x) \equiv p_0 + p_1 x + p_2 x^2 + \dots + p_N x^N$ ; i.e., it is the difference between the slope of the diagonal and the slope of  $f(x)$ . As  $f(x)$  crosses the diagonal from above (see Figure 1a), this difference is positive at the intersection point. Thus, we have:

$$\frac{\partial x}{\partial p_j} = \frac{x^j - 1}{\left[1 - f'(x)\right]}, \quad (\text{A7})$$

and as  $x < 1$  this derivative is negative. This makes sense: if we increase  $p_j$  and decrease  $p_0$  we decrease the extinction probability,  $x$  (and increase  $p(\text{HDF})$ ).

Now, consider the effect on  $x$  of an infinitesimal PPS:

$$(p_0, 0; \dots; p_{j-1} + \Delta, j-1; p_j - 2\Delta, j; p_{j+1} + \Delta, j+1; \dots, p_N, N),$$

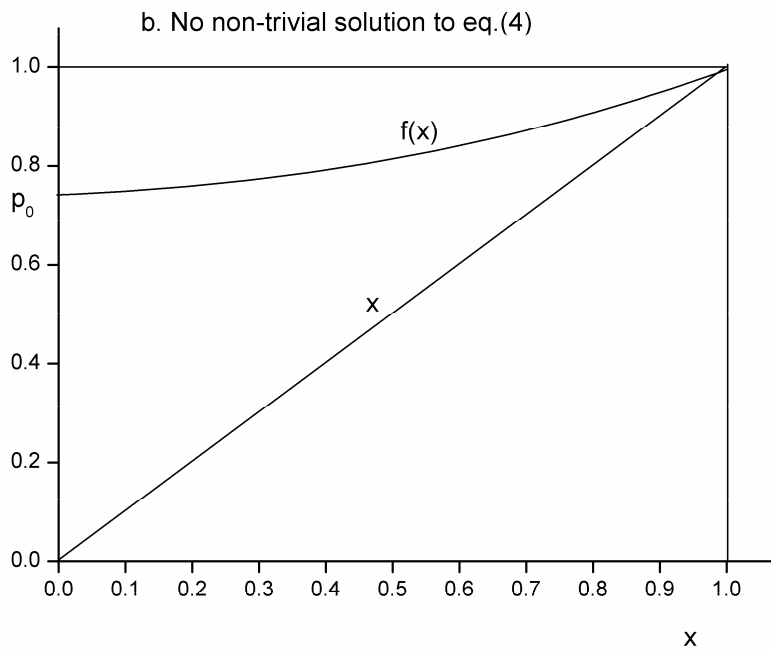
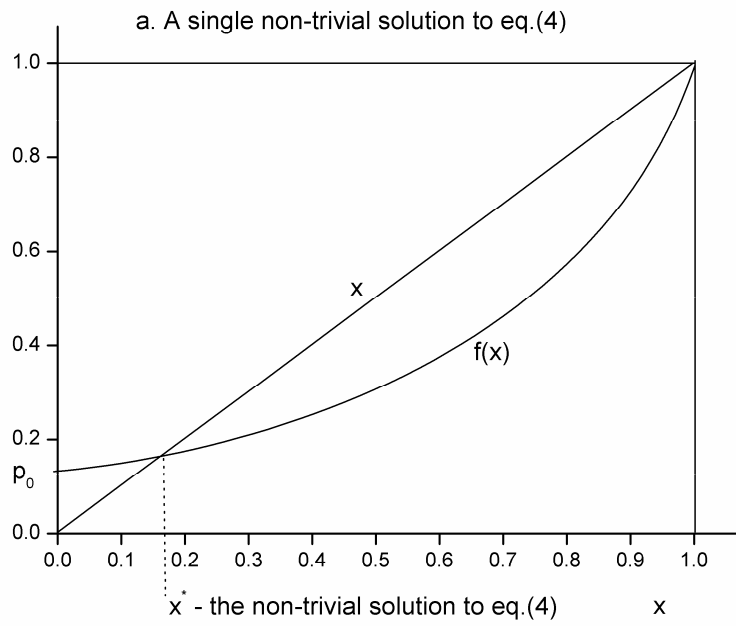
where  $\Delta$  is very small.

The effect of this PPS on  $x$  is:

$$\Delta \frac{\partial x}{\partial p_{j-1}} - 2\Delta \frac{\partial x}{\partial p_j} + \Delta \frac{\partial x}{\partial p_{j+1}} = \frac{\Delta}{[1 - f'(x)]} (x^{j-1} - 2x^j + x^{j+1}). \quad (\text{A8})$$

Note that this expression is positive:  $\Delta$  is positive by definition,  $[1 - f'(x)]$  is positive as explained above, and the term  $(x^{j-1} - 2x^j + x^{j+1})$  is positive because  $x^j$  is a convex function of  $j$ . Thus the infinitesimal PPS increases  $x$ , i.e. it reduces  $p(\text{HDF})=1-x$ . Any PPS that is not infinitesimal can be achieved by a series of infinitesimal PPS's, thus, any PPS increases  $x$  and reduced  $p(\text{HDF})$ . Q.E.D.

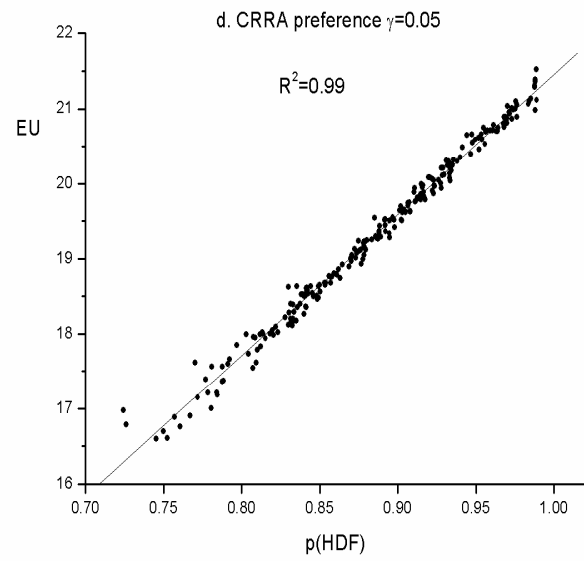
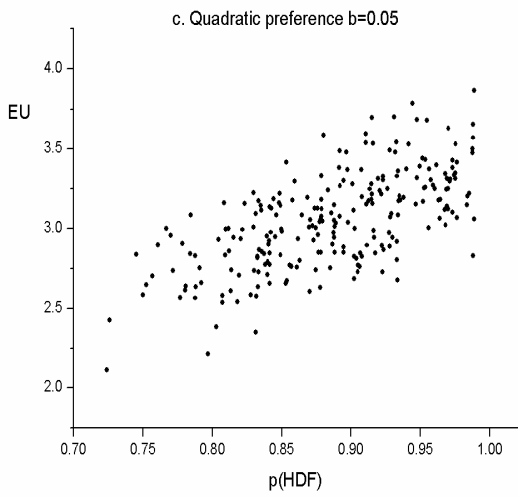
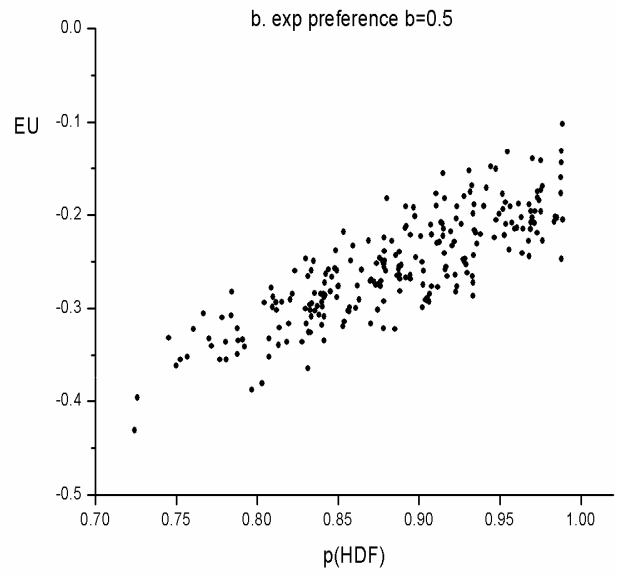
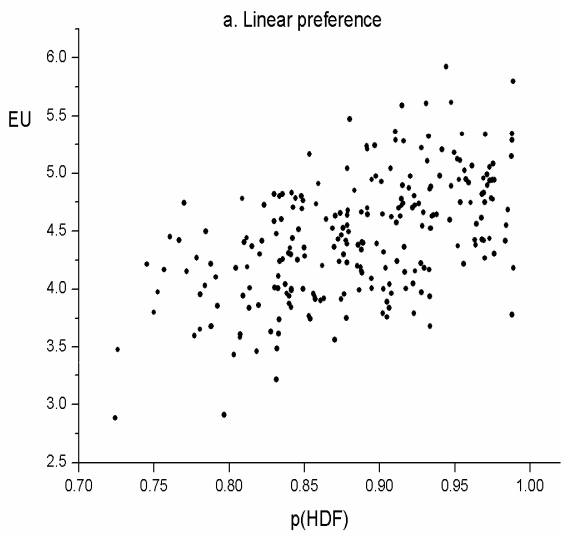
Note also that a shift of probability from the outcome of  $N$  offspring to the outcome of  $N-1$  offspring trivially increases  $x$  (i.e. it makes the gamble less desirable). This is the shift employed in the last stage of Lemma 1.



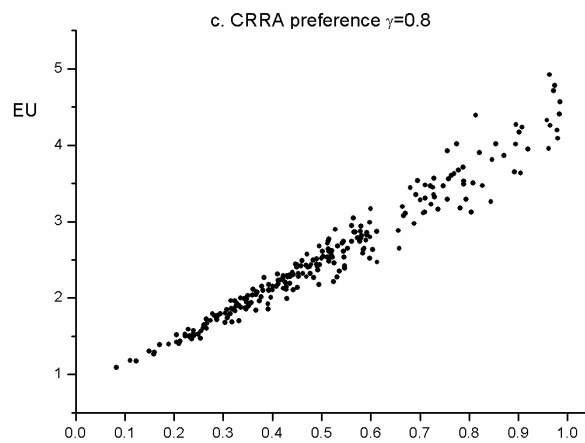
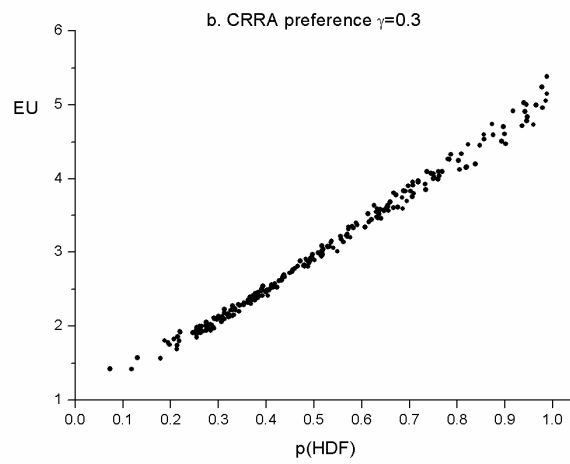
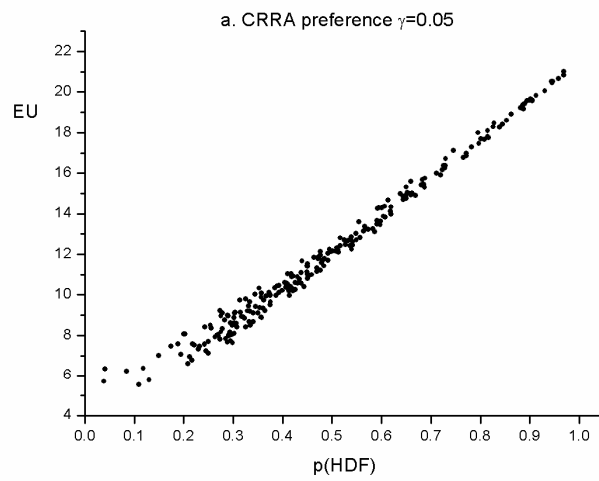
**Figure 1**

There is either a unique solution, or no solution, to eq.(4), as shown by panels a and b, respectively.

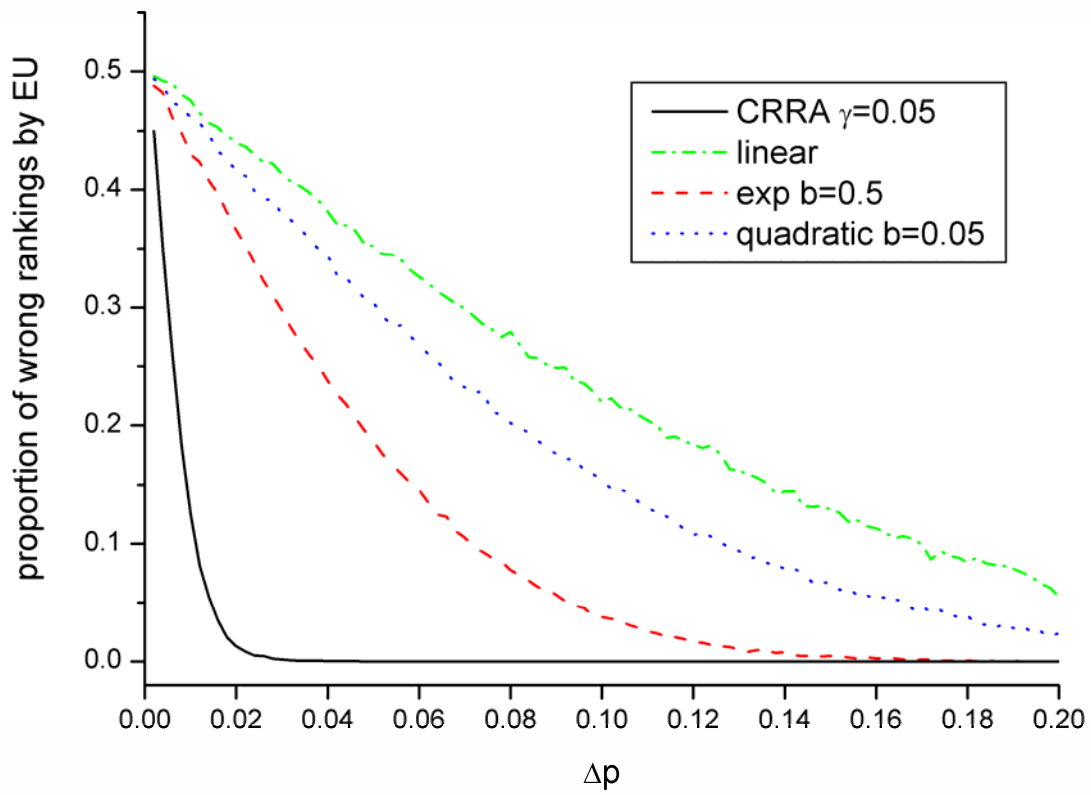




**Figure 2**  
 Expected Utility and  $p(\text{HDF})$  for 1,000 randomly drawn prospects  
 and for various risk-averse utility functions.

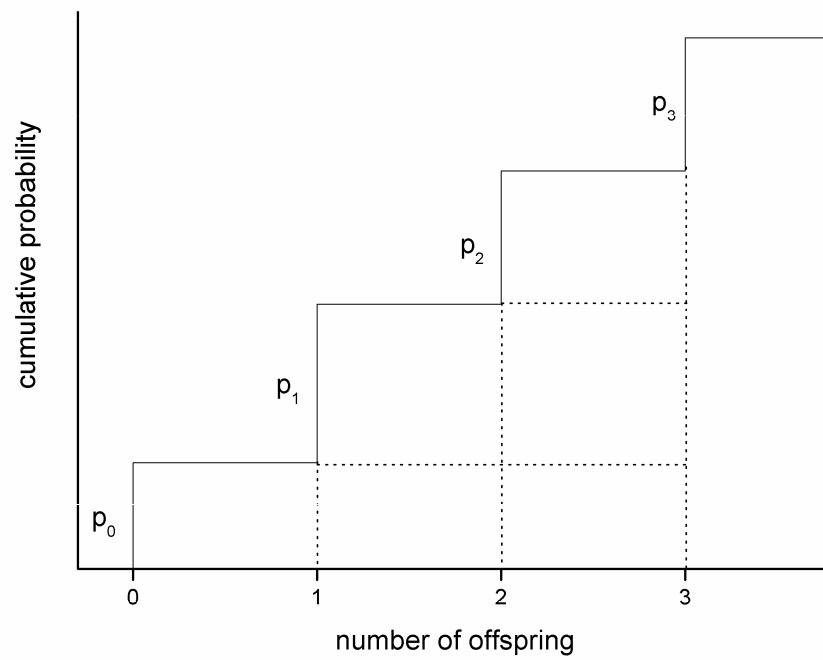


**Figure 3**  
 CRRA Expected Utility and  $p(\text{HDF})$  for various values of the RRA parameter  $\gamma$ .



**Figure 4**

The proportion of errors made by using the EU criterion instead of the p(HDF) criterion as a function of the p(HDF) difference between the two prospects. CRRΑ preference dominates the other utility functions.



**Figure A1**

The cumulative distribution of a prospect up to the outcome of 3 offspring. The area under this function, or integral of the cumulative function, up to  $j=3$  is given by  $(3 \cdot p_0) + (2 \cdot p_1) + p_2$ . In general, the area up to point  $j$  is given by:  $jp_0 + (j-1)p_1 + (j-2)p_2 + \dots + 2p_{j-2} + p_{j-1}$ .