

LEARNING TO MAKE RISK NEUTRAL CHOICES IN A SYMMETRIC WORLD*

STEFANO DELLA VIGNA
Department of Economics
Harvard University

MARCO LICALZI
Department of Applied Mathematics
University of Venice

(First version: May 1997. This version: November 1999.)

Abstract

Given their reference point, most people tend to be risk averse over gains and risk seeking over losses. Therefore, they exhibit a dual risk attitude which is reference dependent. This paper considers an adaptive process for choice under risk such that, in spite of a permanent short-run dual risk attitude, the agent eventually learns to make risk neutral choices. The adaptive process is based on an aspiration level, endogenously adjusted over time in the direction of the actually experienced payoff.

Journal of Economic Literature Classification Numbers: D81, D83.

Keywords: risk attitude, reference point, aspiration level, target, prospect theory, reflection effect.

Corresponding author:

Marco LiCalzi Department of Applied Mathematics, University of Venice,
Dorsoduro 3825/E, I-30123 Venezia, Italy.

Fax: [++39] (041) 522-1756.

E-mail: licalzi@unive.it

* We thank Drew Fudenberg, Bruno Jullien, Mark Machina, two very perceptive referees and seminar audiences at the universities of Alicante, CalTech, London (UCL), San Diego (UCSD), Toulouse and Trento for helpful comments. An earlier version of this paper was written while the second author was visiting the Department of Economics at the University of Toronto. We gratefully acknowledge financial support respectively from Bocconi University and MURST.

1 Introduction

Among economists, the most common assumption about choice behavior under risk is that people maximize expected utility. The second most common assumption is that they are risk averse. Serious difficulties in reconciling this second hypothesis with the empirical evidence have been known since Friedman and Savage (1948). As a way to deal with those, Markowitz (1952) suggested that the utility function might shift depending on the current wealth.

This long forgotten intuition was revived by Kahneman and Tversky (1979) and Fishburn and Kochenberger (1979). They showed how the empirical evidence supports a model where the utility function of an agent is based on a reference point such as her current level of wealth. While people may focus on different reference points — see Heath, Huddart, and Lang (1999), most agents tend to exhibit what Kahneman and Tversky (1979) called the reflection effect: they are risk averse in the gains and risk seeking in the losses with respect to their reference point. A variety of studies and psychological explanations surveyed in Kahneman, Knetsch and Thaler (1991) has established that there is a dual risk attitude which depends on the reference point.

The risk attitude, therefore, is not an intrinsic feature of the choice behavior of an agent. Whatever affects the reference point may determine the risk attitude. Depending on the context and her past experiences, the same agent might seek or avert risk and henceforth make different choices over the same set of lotteries. This compels economics to confront two questions: i) whether her choice behavior may converge (and hence become predictable) in the long run; ii) which kind of risk attitude would eventually emerge?

March (1996) approaches the second question showing that a stimulus-response model of experiential learning tends to generate a dual risk attitude in a simple bandit problem. While he considers possible generalizations, he explicitly warns that “adaptive aspirations present a more complicated problem” (p. 317). This paper studies a model of aspiration-based learning over a known domain of symmetrically distributed payoffs where, while maintaining reference-dependent preferences in the short run, the agent eventually learns to choose a lottery which maximizes expected value. Our results answer the first question in the positive and suggest an important distinction in the second one: even if the agent has a dual risk attitude, this does not rule out her making risk-neutral choices in the long run.

The following outline gives the essential features of our model. The agent repeatedly faces a choice problem over monetary lotteries that are symmetrically distributed. Her reference point is based on an aspiration level which acts as a target. At each period, the agent picks a lottery that maximizes the probability of meeting her current target. Right after, she plays the lottery and adjusts her reference point for the next period in the direction of the actually experienced payoff. In the long-run, the reference point settles down to a specific value and her choice behavior converges to maximization of the expected value. However, at each period the agent maintains a dual risk attitude which never withers away: therefore, learning to make risk neutral choices takes place without

the agent learning to be risk neutral.

The crucial idea for this result is simple: since preferences (and hence choices) depend on the reference point, long-run convergence of the choice behavior follows from the convergence of the reference point. Our model, however, has other features as well. We endogenize the reference point and make it depend on the outcome of the past choices. A simple rule of aspiration adaptation provides the link between the evolution of long-run preferences and the short-run choice behavior. Moreover, we model short-run preferences using an approach which is independent of the expected utility hypothesis, while still consistent with it. This provides a simple generalization of the reflection effect to the case of lotteries which simultaneously involve gains and losses.

There is not yet a theory of the learnable in economics — see Marimon (1997), but our framework is similar in many ways to the adaptive learning models for equilibrium behavior developed in game theory. These models postulate a plausible behavior rule for the players. Combined with their current beliefs about opponent’s actions, this rule selects some strategies resulting in an outcome that affects players’ beliefs in the next period. Against a fixed behavior rule, beliefs evolve endogenously. Thus, convergence to equilibrium behavior follows from convergence to equilibrium beliefs.

Here, we postulate that the behavior rule is target-based maximization. Combined with the agent’s current reference point, this rule selects a lottery resulting in an outcome that affects her target for the next period. Against a fixed behavior rule, the reference point evolves endogenously. Convergence to a risk neutral choice follows from convergence of the reference point to the “right” level.

The importance of maintaining the “right” aspiration level is stressed in Gilboa and Schmeidler (1996). They show that a case-based decision maker who follows an appropriate adjustment rule for his aspiration level learns to maximize expected utility in a bandit problem. As in our model, they postulate that in each period the agent follows a deterministic rule of maximization. Börgers and Sarin (1996), instead, study a model where the agent’s choice rule is stochastic. At each period, the agent raises (or lowers) the probability to choose the lottery just selected according to whether its outcome met (or not) her current target. As in our model, the aspiration level is adjusted in the direction of the experienced payoffs. However, since the choice rule introduces probability matching effects, choice behavior usually fails to converge to expected value maximization.

Both of these papers are concerned with choice behavior. Rubin and Paul (1979) and Robson (1996), instead, discuss the formation of individual risk attitudes from an evolutionary viewpoint. Measuring success by the expected number of offsprings, they investigate which kinds of risk attitudes are more likely to emerge. While they suggest reasons for risk seeking over losses by younger males, this approach has not yet been able to fully account for reference-dependent preferences. Dekel, Ely and Yilankaya (1998) go further and study the endogenous determination of preferences in evolutionary games, but provide no specific results on the evolution of risk attitudes.

2 The short-run model

This section describes our target-based model for the short-run choice behavior. The long-run dynamics of our model is studied in the next section. For the sake of readability, we relegate most proofs and a few auxiliary results to the appendix.

2.1 Preliminaries

The agent faces a finite (nonempty) set \mathcal{L} of monetary lotteries out of which she must choose one. With no loss of generality, we assume that probabilities are exogenously given. All the lotteries in \mathcal{L} have finite variance. We denote by μ_i and σ_i respectively the expected value and the standard deviation of a lottery X_i in \mathcal{L} . We write $X \sim F$ to denote that F is the cumulative distribution function (for short, c.d.f.) of X .

Given a lottery $X \sim F$, we say that X is symmetrically distributed (and F is symmetric around m) if there exists some m in \mathbb{R} such that $F(m - x) + F(m + x) = 1$ for almost all x in \mathbb{R} . When F is continuous, taking limits for $x \rightarrow 0$ in the previous equality gives $F(m) = 1/2$. Therefore, we define m as the median of a symmetrically distributed X . Note that our definition implies that m is unique even if the set of x 's such that $F(x) = 1/2$ is not a singleton. Moreover, if the expected value of X is finite, this is equal to m . Therefore, for a symmetrically distributed lottery X , we denote both the median and the expected value by μ .

We suppose that all the lotteries in \mathcal{L} are symmetrically distributed and that their c.d.f.'s are absolutely continuous with full support on \mathbb{R} . This restrictive assumption is maintained throughout the paper. For the mere purpose of exemplification, we will occasionally assume that the lotteries in \mathcal{L} are normally distributed.

2.2 Target-based ranking

In order to make her choice, the agent must rank the lotteries. We assume that she follows the target-based procedure described in Bordley and LiCalzi (1998). Let V denotes the agent's target, representing her aspiration level. The agent is interested in maximizing the probability of meeting her target. Therefore, she ranks a lottery X in \mathcal{L} by the index $w(X) = P(X \geq V)$; that is, by the probability that X meets her target. To allow for the case where the agent does not know exactly her aspiration level, we assume that V is a random variable with a subjective distribution stochastically independent of \mathcal{L} . Therefore, the agent ranks $X \sim F$ by the index

$$w(X) = P(X \geq V) = \int P(x \geq V) dF(x). \quad (1)$$

Given that the agent relies on an uncertain target to rank the available lotteries, we need to formulate some assumptions on the distribution of her aspiration level. In a world where all lotteries are symmetrically distributed, the most natural hypothesis is that the agent's target V is also symmetrically distributed. We refer to the situation where both

the lotteries in \mathcal{L} and the target are symmetrically distributed as a “symmetric world”. When lotteries and target are taken to be normally distributed, we speak of a “normal world”.

An agent using the target-based procedure may not be sure about what her aspiration level V should be. Unless this target is precisely known to her, it is not clear how the agent should decide to code a specific outcome as a gain or as a loss. According to prospect theory, the agent sets a reference point ν in \mathbb{R} and classifies an outcome as a gain or as a loss according to its being higher or lower than ν .

Therefore, we need some assumptions to relate the target V with the reference point ν . If the target is degenerate, it is reasonable to expect it to coincide with the reference point. More generally, a natural assumption might be to let ν be the mode of the target. This matches the intuition that in many cases the agent is likely to anchor her perceptions to the modal outcome. However, the mode of V may not be unique and there may be reasons to prefer other statistics. For instance, one might suggest to let the reference point be some central statistics like the expected value or the median of V . Note though that these two statistics usually differ for non-symmetric distributions.

In a symmetric world, most of these concerns disappear. When the target V is symmetrically distributed, its expected value and its median have the same value. Moreover, if the distribution of the target is unimodal, the mode is unique and is also equal to the expected value. Given that any of these three statistics can claim to be a prominent anchor and that they all coincide, we let the reference point be the expected value ν of the aspiration level V . Note that the assumption of a unimodal target is not part of our definition of a symmetric world: except for the next section, this assumption is not used in the rest of the paper.

2.3 Dual risk attitude

This section shows that, in a symmetric world where the distribution of the target is unimodal, the agent exhibits a dual risk attitude centered around her reference point. This matches the compelling evidence that decision makers tend to be risk averse with respect to gains over their reference point and risk seeking with respect to losses.

We illustrate the point with an example based on a “normal” world. Assume temporarily that all the lotteries in \mathcal{L} are normally distributed with strictly positive variances and, by analogy, that the target is also normally distributed with expected value ν and standard deviation $s > 0$; for short, we write $V \sim N(\nu, s)$. By Theorem A.1 in the Appendix,

$$P(X \geq V) = \Phi\left(\frac{\mu - \nu}{\sqrt{\sigma^2 + s^2}}\right),$$

where Φ is the c.d.f. of a standard normal.

Given a lottery X with expected value μ , denote by $c(X)$ the certainty equivalent of X . An agent is said to be risk averse over X if her certainty equivalent $c(X) \leq \mu$; correspondingly, she is risk seeking if $c(X) \geq \mu$ and risk neutral if equality holds. In a normal world, the certainty equivalent exists and is unique for any lottery. If the agent

follows the target-based ranking procedure, Theorem A.2 in the Appendix shows that her certainty equivalent for X can be written as a convex combination of μ and ν :

$$c(X) = \alpha\mu + (1 - \alpha)\nu, \tag{2}$$

where of course α is in $[0, 1]$. Therefore, the agent is risk averse (respectively, neutral or seeking) over all lotteries that have a higher (equal or lower) expected value μ than her reference point ν . To put it differently, she tends to be more cautious when she expects (on average) to meet her target.

We think that this result fits the spirit of prospect theory. However, we must point out that this is not the reflection effect studied by Kahneman and Tversky (1979). They confronted subjects with pairs of lotteries yielding only positive payoffs and most choices exhibited risk aversion; then they changed positive with negative payoffs and several people switched to risk seeking behavior. Each lottery they examined was on just one side of the reference point. In our normal world, instead, the lotteries in \mathcal{L} have full support and therefore none of them offers only gains (or losses).

On the other hand, since normal distributions are symmetric, a lottery X with expected value $\mu > \nu$ is more likely to yield outcomes above the reference point ν than not. Accordingly, the agent tends to view X as a positive lottery and exhibits risk aversion. As $\mu \downarrow \nu$, her “positive” perception of X becomes weaker and weaker until $\mu = \nu$. At this point, the lottery X is as likely to yield outcomes above the reference point as below and the agent perceives it as “balanced”, so that she exhibits risk neutrality. For $\mu < \nu$, the lottery becomes mostly “negative” and the agent turns to a risk seeking behavior.

There is another subtler difference between our approach and Kahneman and Tversky (1979). We postulate target-based ranking as a behavior rule and use this assumption to deduce that the agent will exhibit a dual risk attitude. Kahneman and Tversky (1979), instead, start off from the empirical evidence about dual risk attitudes and show that it may be accounted for by a reflection effect in the utility function. In a sense, while we are trying to suggest an explanatory mechanism, they simply strive to obtain a robust description of the empirical evidence; see Gigerenzer (1996). Moreover, they build their descriptive model by amending the expected utility model (to a significant extent) while we abstract from it.

Our model for a dual risk attitude essentially carries over to a symmetric world, provided that the distribution of the target is unimodal. The whole argument rests on the validity of Equation (2) in a normal world. When trying to extend the equation to a symmetric world, one runs into the technical difficulty that the existence of the certainty equivalent for any lottery X cannot be guaranteed without further regularity assumptions. Theorem A.4 in the Appendix proves that Equation (2) holds whenever the certainty equivalent of X exists and V has a unimodal distribution. Thus, if we base the definition of risk attitude on the certainty equivalent, unimodality of the target implies that the dual risk attitude holds in a symmetric world as well.

2.4 The relationship with expected utility

It is instructive to consider how our agent would look like to someone trying to interpret her choice behavior in view of the expected utility model. Maybe surprisingly, her target-based ranking procedure is both mathematically and observationally equivalent to the expected utility model. This is shown by Castagnoli and LiCalzi (1996) in von Neumann and Morgenstern's setting and by Bordley and LiCalzi (1998) in Savage's framework. For our purposes here, it suffices to note that the choice behavior associated with (1) is compatible with the expected utility model provided that we interpret $P(x \geq V)$ as the utility of the outcome x ; that is, if we let $U(x) = P(x \geq V)$. Accordingly, we denote the c.d.f. of the uncertain target V by U and write $V \sim U$.

To illustrate this point, we assume a normal world where the target V is normally distributed with expected value ν and standard deviation $s > 0$; then $P(x \geq V) = \Phi[(x - \nu)/s]$. An agent that maximizes the probability to meet the target V would rank all lotteries *as if* she is maximizing the expected value of the utility function $U(x) = \Phi[(x - \nu)/s]$. Either approach makes exactly the same predictions.

For instance, note that $\Phi[(x - \nu)/s]$ is concave for $x \geq \nu$ and convex otherwise. In the language of expected utility, we say that the utility function $U(x)$ is S -shaped around ν . In the target-based language, we say that the distribution of the target is unimodal. In either case, this characteristic curvature generates a reflection effect around the reference point ν .

The equivalence between the expected utility model and the target-based procedure raises an important modelling issue: why the second and not the first one? There are two independent reasons for this choice. First, the target-based procedure offers a more plausible account of how reference dependence may come into place in the short run. Second, it suggests a natural way to aggregate many short-run choices into a single long-run process. For other modelling advantages of a target-based language see LiCalzi (1999). We argue here for the first claim and defer the discussion of the second one to Section 3.4.

There is a tendency to confound the assumption that the expected utility model represents the agent's preferences with the claim that the agent actually uses expected utility to find out what she likes better. The first viewpoint is descriptive: the purpose of a theory is to describe behavior, not to explain it. Following this viewpoint, it is immaterial which model we adopt provided that it shares the same descriptive accuracy of its competitors. Both an S -shaped utility function and a unimodal target generate the reflection effect. Therefore, they are equally good and equally bad.

The second viewpoint is closer to looking for an explanation of the observed behavior. An S -shaped utility function does not explain much. The target-based procedure suggests a mechanism which may be the cause behind the reflection effect. Suppose that the agent is not sure about her target and must estimate a probability distribution for it. One possibility is that she focuses on the most likely value ν and then assesses decreasing probabilities as she moves away from ν . The probability that x is the "true" target is a decreasing function of its distance from ν which, accordingly, acts as a reference point. This would lead to the unimodal distribution which generates the reflection effect.

3 The long-run model

3.1 The evolution of the reference point

Consider an agent who deals with the following choice problem under risk. At times $t = 1, 2, \dots$, the agent faces a fixed set \mathcal{L} of lotteries which satisfy the assumptions listed in Section 2.1. We assume a symmetric world where, at each time, any lottery in \mathcal{L} is stochastically independent of the lotteries that were available in previous periods. At each time t , the agent has a target V^t which represents her current (and possibly uncertain) aspiration level. The distribution of the target at time t is stochastically independent of any set \mathcal{L} which was available up to (and including) time t . The agent ranks the lotteries in \mathcal{L} by the probability that they meet the target, following the procedure described in Section 2.2. We denote by X^t the lottery chosen in t and by x^t the corresponding outcome of X^t , which becomes known to the agent right after she selects X^t .

The distribution of the target V^t may change over time. We assume that the agent's aspiration level is closely related to the average outcome that the agent has obtained in the past. Let X^0 be the agent's (arbitrary, and possibly random) initial aspiration level in $t = 0$. At the beginning of period t , the outcome x^τ received for playing lottery X^τ is known for any $\tau = 0, 1, \dots, t - 1$. The agent sets her current reference point to

$$\nu^t = \bar{x}^t = \left(\frac{1}{t}\right) \sum_{\tau=0}^{t-1} x^\tau. \quad (3)$$

The target V^t used to choose a lottery at time t is a (possibly degenerate and) symmetrically distributed random variable with expected value $\nu^t = \bar{x}^t$. Note that we do not place any restriction on the variance or other moments of V^t . The only restriction on the updating rule for V^t is that its expected value should obey (3). Therefore, there is a substantial degree of freedom about the process by which the agent sets her aspiration level.

Intuitively, we are assuming that the reference point \bar{x}^t adapts to how things have been going. If the outcomes experienced in the recent past have been relatively unsatisfactory, the reference point decreases and the agent lowers the expected value of her target (or raises it in the opposite case). Hence, the reference point evolves endogenously as a consequence both of the agent's choices (which lottery X she plays) and of chance (which outcome x actually occurs).

Chance is what makes interesting this simple rule for setting the reference point. Although at the beginning of period t the reference point \bar{x}^t is known with certainty, its evolution over time is stochastic because it depends on which outcome obtains from the lottery that the agent selects. Today we may know exactly her reference point, but what it will become tomorrow depends on how the lottery X^t played today will turn out.

More formally, given the information available at the beginning of period t , the reference point for the next period is the random variable

$$\bar{X}^{t+1} = \left(\frac{t}{t+1}\right) \bar{x}^t + \left(\frac{1}{t+1}\right) X^t. \quad (4)$$

The (already known) reference point \bar{x}^t for the previous period combines with the (yet to be realized) lottery X^t . The randomness of the X^t term makes the evolution of the reference point stochastic. In turn, since the agent's choice depends on her target, this makes the evolution of her choice behavior stochastic as well.

3.2 Convergence of the reference point

This section shows that, in spite of this randomness, the agents' reference point converges (almost surely) to a level that induces her to pick a lottery that maximizes expected value, so that she learns to make risk neutral choices.

Let μ_* denote the greatest expected value attainable in the (finite) set \mathcal{L} . The set \mathcal{L} may contain more than one lottery with expected value μ_* . However, we show below in Theorem 1 that if a lottery X_* with expected value μ_* is optimal for the agent, then all (and only all) the lotteries with expected value μ_* are optimal as well. Therefore, for the sake of simplicity, we pretend that there exists only one such lottery X_* and we say that X_* is optimal to mean that any lottery with expected value μ_* is optimal. Note that by definition X_* maximizes expected value and hence it represents the risk neutral choice.

We wish to prove that eventually the agent learns to choose X_* . How might this come about? Suppose that the current reference point is ν . Our first result shows that the reference point $\nu = \mu_*$ has a crucial fixed-point property.

Theorem 1 *At the reference point $\nu = \mu_*$, the optimal choice has expected value μ_* . If $\nu > \mu_*$ (respectively, $\nu < \mu_*$), then any optimal choice has an expected value strictly lower (respectively, higher) than ν .*

PROOF: There are three cases to consider, depending on the sign of $(\mu_* - \nu)$. First, suppose that $\nu = \mu_*$. By Theorem A.3 in the appendix, $P(X_* \geq V) = 1/2$. Since X_* has the greatest expected value from \mathcal{L} , any other lottery X in \mathcal{L} has an expected value $\mu < \mu_* = \nu$. By Theorem A.3 again, $P(X \geq V) < 1/2$ and therefore X cannot be an optimal choice. Second, suppose that $\nu > \mu_*$. Since any lottery X in \mathcal{L} has an expected value $\mu \leq \mu_* = \nu$, the result follows immediately. Third, suppose that $\nu < \mu_*$. By Theorem A.3, $P(X_* \geq V) > 1/2$. By Theorem A.4 again, any lottery X with expected value $\mu < \nu$ gives $P(X \geq V) < 1/2$ and therefore cannot be optimal. Hence, the optimal choice must have an expected value $\mu \geq \nu$. \square

Note that the theorem does not rule out the possibility that X_* be an optimal choice when the reference point $\nu \neq \mu_*$. The most important consequences of Theorem 1 are three. First, when her reference point ν is lower than μ_* , the agent is ambitious in the sense that she goes after lotteries with an expected value higher than ν . Second, when her reference point ν is greater than μ_* , the agent must necessarily pick lotteries with an expected value lower than ν . Third, μ_* is the only reference point which induces the agent to choose a lottery with an expected value equal to the reference point.

Therefore, a sufficient condition for the agent to pick X_* as her optimal choice in \mathcal{L} is that her reference point is $\nu = \mu_*$. The rest of this section is devoted to show that,

under a slight strengthening of ambition, the stochastic evolution of the reference point settles (almost surely) at μ_* regardless of its initial position. Thus, it is by adjusting her reference point until it matches the highest available expected value that the agent learns to make a risk neutral choice.

Consider the stochastic evolution of the reference point. Given a target V^t with reference point \bar{x}^t , let $B(V^t)$ be the choice map induced by the ranking index $P(X \geq V^t)$ at time t . At the beginning of period t , the agent selects a lottery in $B(V^t)$. Since our convergence result holds no matter how indifferences in $B(V^t)$ are resolved, we assume without loss of generality that there is a tie-breaking rule which turns this choice map into a choice function. Therefore, we let $X^t = B(V^t)$ and write Equation (4) as

$$\bar{X}^{t+1} = \left(\frac{t}{t+1}\right)\bar{x}^t + \left(\frac{1}{t+1}\right)B(V^t) = \bar{x}^t + \frac{1}{t+1} [B(V^t) - \bar{x}^t]. \quad (5)$$

Equation (5) has a structure similar to the stochastic approximation algorithm introduced by Robbins and Monro (1951). Since Blum (1954) proved a sufficient condition for its almost sure convergence, a vast literature has developed proposing various types of stochastic approximation procedures and proving several sufficient conditions for their convergence. See Ljung et alii (1992) for a mathematical introduction and Sargent (1993) for a partial survey of applications to adaptive learning in economics. Using methods from stochastic approximation theory, Theorem A.6 in the appendix establishes a result of almost sure convergence for this algorithm.

Let \mathcal{F}^t the σ -algebra generated by $\{X^0, X^1, \dots, X^{t-1}\}$. Let $b^t(V^t) = E [B(V^t) | \mathcal{F}^t]$ the conditional expectation of the optimal choice $B(V^t)$, given the current target V^t and the payoffs realized in the past (up to time t). The ambition property from Theorem 1 states that $\mu_* > \bar{x}^t$ implies $b^t(V^t) > \bar{x}^t$. We say that ambition is strict if furthermore $\mu_* > \bar{x}^t + \varepsilon$ with $\varepsilon > 0$ implies $b^t(V^t) \geq \bar{x}^t + \eta$ with $\eta > 0$, where η depends on ε but not on t or on the events in \mathcal{F}^t . The main result of this section is a corollary of Theorem A.6.

Theorem 2 *Under strict ambition, the reference point \bar{x}^t converges almost surely to μ_* .*

The proof is in the appendix, but we offer here some intuition about how this result comes about. Rewrite Equation (5) as

$$\bar{X}^{t+1} = \bar{x}^t + \frac{1}{t+1} [b^t(V^t) - \bar{x}^t] + \frac{1}{t+1} [B(V^t) - b^t(V^t)].$$

Given \mathcal{F}^t , the conditional expectation of $[B(V^t) - b^t(V^t)]$ is zero and therefore the (conditional) expected motion of \bar{X}^{t+1} is

$$E [\bar{X}^{t+1} | \mathcal{F}^t] = \bar{x}^t + \frac{1}{t+1} [b^t(V^t) - \bar{x}^t].$$

This shows that the evolution of the expected value of \bar{X}^{t+1} tends to be driven by the $[b^t(V^t) - \bar{x}^t]$ term which, by Theorem 1, always points towards μ_* .

This expected motion towards μ_* is perturbed by the stochastic term $[B(V^t) - b^t(V^t)]$, which in principle might lead the system elsewhere. However, the cumulated perturbations

$$\sum_{\tau=1}^t \frac{1}{\tau+1} [B(V^\tau) - b^\tau(V^\tau)]$$

form a martingale and converge (almost surely). Hence, after a sufficient time, the cumulated perturbations cannot counter the expected motion and the system settles down on μ_* .

3.3 An example in a normal world

This section provides a pictorial example set in a normal world with a degenerate target. When \mathcal{L} is a set of normal distributions and the agent's ranking is compatible with expected utility as discussed in Section 2.4, we can simplify the representation of the choice set and replace a lottery X by the pair (μ, σ) of its expected value and standard deviation; see Meyer (1987). For instance, if the target is known for sure and therefore coincides with the reference point ν , Theorem A.1 in the Appendix implies that the agent behaves as if choosing the lottery that maximizes $(\mu - \nu)/\sigma$.

Let \mathcal{C} be the set of (μ, σ) -pairs associated with lotteries in \mathcal{L} . Note that there may be more than one lottery in \mathcal{L} corresponding to the same pair in \mathcal{C} , but the agent will be indifferent among them. We can represent the new choice set \mathcal{C} as points on the plane; see for instance Figure 1. In this setting, the optimal choice of the agent depends

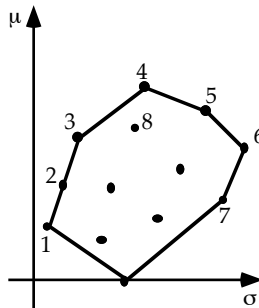


Figure 1: An example of a normal world.

only on her reference level ν . There are lotteries which can never be optimal. For a lottery $X_1 \sim N(\mu_1, \sigma_1)$ to be optimal, there must exist a reference point ν such that $(\mu_1 - \nu)/\sigma_1 = \max_{\mathcal{C}} (\mu - \nu)/\sigma = k$. This occurs if and only if $\mu \leq \bar{x} + k\sigma$ for all $(\sigma, \mu) \in \mathcal{C}$ with equality holding at (σ_1, μ_1) ; or, equivalently, if and only if there exists an hyperplane through (μ_1, σ_1) supporting \mathcal{C} from above (with respect to the μ -axis). Therefore, a lottery $X \sim N(\mu, \sigma)$ can be optimal if and only if (μ, σ) lies on the upper boundary of the convex hull of \mathcal{C} .

To gain some intuition, consider Figure 1. Only the lotteries corresponding to the points from 1 to 6 can be optimal. Lottery 2 can be optimal although it is not an extreme point; but when 2 is optimal, the agent is indifferent between Lotteries 1, 2, and 3. Lottery 7 is an extreme point, but it cannot be optimal because either 5 or 6 will be preferred. Lottery 8 cannot be optimal, although its expected value is higher than most other lotteries.

Order just the lotteries that can be optimal so that they are increasing in the variance. For instance, we have numbered such lotteries in Figure 1 from 1 to 6. Then the actual optimal choice is increasing in the reference point ν ; that is, the higher is ν , the higher is the variance of the optimal choice. In fact, given two reference points $\nu_1 > \nu_2$, let X_1 and X_2 be respectively optimal in ν_1 and ν_2 . By Theorem A.1, it must be

$$\frac{\mu_1 - \nu_1}{\sigma_1} \geq \frac{\mu_2 - \nu_1}{\sigma_2} \quad \text{and} \quad \frac{\mu_2 - \nu_2}{\sigma_2} \geq \frac{\mu_1 - \nu_2}{\sigma_1}.$$

Adding the two inequalities and rearranging terms, we obtain $(\nu_1 - \nu_2)/\sigma_1 \leq (\nu_1 - \nu_2)/\sigma_2$, from which $\sigma_1 \geq \sigma_2$.

Intuitively, the higher is the reference point, the higher the variance of the optimal lottery. If the upper boundary of the convex hull is humped as in Figure 1, an agent with a sufficiently low aspiration level chooses a lottery with low variance (and hence low mean); and if she has a sufficiently high reference point, she picks a lottery with high variance (but again low mean). By well-known results about the risk attitude over normal distributions — see Hanoch and Levy (1969), the agent is risk averse (respectively, risk seeking) when her reference point is sufficiently low (or high).

3.4 Two counterexamples

The main results in this paper are two. First, reference-dependent preferences may converge in the long run. Second, they may lead to risk neutral behavior. This section provides two simple counterexamples, one for each result. Their purpose is to help the reader assess the strengths and the limitations of our approach.

We show first that preferences may fail to converge because the reference point may cycle forever. Recall that we took the reference point to be (essentially) the sample mean of the outcomes received in the past. For simplicity, suppose that the target is degenerate and equal to ν . The choice set \mathcal{L} contains two lotteries: X is normally distributed with expected value $\mu = 0$ and standard deviation $\sigma = 20$; and Y is a (non-symmetric) binary lottery which pays 1 with probability 0.9 and 11 with probability 0.1. The expected value of Y is $2 > \mu$. Let ν be the current reference point.

We claim that the agent strictly prefers X when $\nu > 1$ and strictly prefers Y when $\nu \leq 1$. For $\nu \leq 1$ and $\nu > 11$, this is immediate. For ν in $(1, 11]$, note that X is normally distributed and therefore

$$P(X \geq \nu) > P(X \geq 20) = P(X \geq \mu + \sigma) = 1 - \Phi(1) \approx .1587 > .1 = P(Y \geq \nu).$$

As long as $\nu \leq 1$, the agent chooses Y in each period. Since the expected value of Y is 2, persisting in this choice generates payoffs that drive the sample mean up until $\nu > 1$; at that point, the agent switches to choosing X . But, in turn, the expected value of X is $\mu = 0$ and this must eventually bring the sample mean down so that the agent switches back to Y . The reference point cannot converge and thus the agent's choice keeps going back and forth between X and Y forever, while each cycle stretches (on average) over longer and longer periods.

This first counterexample proves that Theorem 2 is not trivial. Our second counterexample shows that, even when the reference point converges, the long run choice behavior may not be risk neutral. This suggests that our model may be used to explain convergence of long-run behavior to different kinds of risk attitudes in worlds which are not symmetric. For the moment, we must leave this conjecture to future research.

Assume again a degenerate target and a choice set with two lotteries: $X \sim N(0, 20)$ as before and another (non-symmetric) binary lottery Z which pays -1 with probability 0.9 and 10 with probability 0.1. Similarly to the above, the agent strictly prefers X when $\nu > -1$ and strictly prefers Z when $\nu \leq -1$. Thus, as far as $\nu \leq -1$, the agent chooses Z . Since the expected value of Z is 0.1, this must eventually drive the sample mean over -1 and make the agent pick X . Once the agent starts choosing X , which has $\mu = 0$, the sample mean tends to settle around 0. Eventually, the reference point converges to 0 and the agent prefers X although Z has a higher expected value. In other words, the agent learns to make a (strictly) risk averse choice.

We can offer some intuition about why risk neutrality fails to emerge in this example. The key property is *ambition*, as it appears in Theorem 1: if there is a lottery with expected value strictly higher than ν , the agent does not content herself with any lottery whose expected value is lower than ν . If the agent is ambitious for all values of $\nu < \mu_*$ (and if she prefers X_* at $\nu = \mu_*$), then her choice behavior will converge to risk neutrality. In the counterexample above, ambition fails for ν in $[0, 0.1)$ and this prevents convergence to the risk neutral choice.

We note one last important feature of our model. Establishing Theorem 1 suffices to prove Theorem 2 and therefore to obtain convergence to risk neutrality. We proved Theorem 1 using Theorem A.3, which relies on the assumption of a symmetric world. However, there may be different “worlds” in which Theorem 1 is true. In any of them, the choice behavior would converge to risk neutrality by Theorem 2. This suggests how to go about proving a piece of our conjecture above: risk neutrality may also emerge in worlds that are not symmetric.

3.5 Changing preferences and expected utility

As the reference point changes, so do the short-run preferences of the agent. It is instructive to consider how their dynamics looks like in both our target-based model and the expected utility model. We begin from the first one.

At the beginning of period t , we can write the target

$$V^t = \bar{x}^t + \zeta^t$$

as the sum of the current reference point \bar{x}^t and a random term ζ^t . This allows one to model situations where the agent is uncertain about which reference point she should use. For instance, the agent may feel that the average experienced payoff \bar{x}^t is an undistorted estimate of the ideal reference point, but that there might be some other relevant (yet unavailable) information that she should take into account. This information combines into a probabilistic assessment of the aspiration level at each period t .

Given that convergence in our model only requires convergence of \bar{x}^t , the possible dynamics of these probabilistic assessments includes many possible cases. For instance, we could assume that the standard deviation s^t of the target V^t is decreasing over time as the sample of experienced payoffs becomes longer: the greater the sample, the lower the uncertainty. Alternatively, we might assume that $s^t = 0$ if $\bar{x}^t < \mu_*$ and $s^t > 0$ otherwise; that is, the agent feels confident that \bar{x}^t is her target when there is at least one lottery in \mathcal{L} that has a higher mean, but entertains doubts once she realizes that her reference point may be too high. More generally, provided that (3) holds, our model is compatible with any updating rule for V^t . We do not even require V^t to converge, although \bar{x}^t will.

Hence, in our model based on a random target, the aggregation of short-run preferences into a long-run preferences is compatible with a large class of probabilistic assessments and updating rules. The agent is not required to have a utility function and her preferences change simply because her opinions change.

The expected utility model is compatible with ours. Therefore it is possible to rewrite our model while having the agent maximize expected utility at each period. However, doing so requires *ad hoc* assumptions that betray that the formal equivalence is not substantial. Given V^t , let U^t be its corresponding c.d.f. When V^t is updated to V^{t+1} , U^t becomes U^{t+1} . The equivalence requires that U^t be used as utility function in the expected utility representation of the agent's behavior. Therefore, the same dynamics ensues if we postulate that the utility function U^t changes its shape over time while the reference point independently evolves so that (3) holds. But it would take a quite contrived story to explain why and how U^t shifts over time. The model would manage to describe the dynamics of preferences only by ignoring the issue of why they are changing.

4 Conclusions

The adaptive process described in this paper seems close to the learning experience underlying the formation of expertise. Suppose that an inexperienced agent sets her aspiration level too low (or too high). If she exhibits a dual risk attitude, her choices will lead to outcomes that systematically overshoot (or undershoot) the target. As her experience grows, she learns to set more realistic targets and to make choices matching (on average) her aspiration level. Moreover, this process tends to maximize the expected value

of the chosen lotteries. Experience is embodied in the evolution of the target and in the stabilization of the reference point at the “right” level.

Unfortunately, we are not aware of any experimental evidence which bears on this interpretation. More generally, the literature lacks attempts to study the evolution of risk attitudes over time. It may be time to undertake such an investigation. Behavioral economics is becoming increasingly aware that preferences are often endogenous and that they may be learnt. However, what (if anything at all) can be eventually learnt is not yet clear.

This paper proposes a model which makes testable predictions about the evolution of choice behavior over symmetric lotteries. In particular, it would be interesting to study the risk attitudes of experienced versus inexperienced agents in a symmetric world. We predict that the first group would be more likely to engage in risk neutral choices.

A Appendix

Theorem A.1 *Suppose that $X \sim N(\mu, \sigma)$ and $V \sim N(\nu, s)$ are stochastically independent and normally distributed. If $\sigma^2 + s^2 > 0$, then*

$$P(X \geq V) = \Phi\left(\frac{\mu - \nu}{\sqrt{\sigma^2 + s^2}}\right)$$

where Φ is the distribution function of a standard normal. If instead $\sigma^2 = s^2 = 0$, then $P(X \geq V) = 1$ if $\mu \geq \nu$ and 0 otherwise.

PROOF: The difference of the two normals V and X is $(V - X) \sim N((\nu - \mu), \sqrt{\sigma^2 + s^2})$. If $\sigma^2 + s^2 > 0$, we have that

$$Z = \frac{(V - X) - (\nu - \mu)}{\sqrt{\sigma^2 + s^2}}$$

is a standard normal and therefore

$$P(X \geq V) = P(V - X \leq 0) = P\left(Z \leq \frac{\mu - \nu}{\sqrt{\sigma^2 + s^2}}\right) = \Phi\left(\frac{\mu - \nu}{\sqrt{\sigma^2 + s^2}}\right).$$

If instead $\sigma^2 = s^2 = 0$, both X and V are degenerate and $P(X \geq V) = P(\mu \geq \nu)$. \square

Theorem A.2 *Given a target $V \sim N(\nu, s)$ with $s > 0$ and a lottery $X \sim N(\mu, \sigma)$, there exists α in $[0, 1]$ such that the certainty equivalent $CE(X)$ satisfies*

$$CE(X) = \alpha\mu + (1 - \alpha)\nu.$$

PROOF: By Theorem A.1, $CE(X)$ must satisfy

$$\frac{\mu - \nu}{\sqrt{\sigma^2 + s^2}} = \frac{CE(X) - \nu}{s}.$$

Rearranging, we obtain

$$\text{CE}(X) = \nu + \frac{\mu - \nu}{\sqrt{1 + (\sigma^2/s^2)}} = \left(\frac{1}{\sqrt{1 + (\sigma^2/s^2)}} \right) \mu + \left(1 - \frac{1}{\sqrt{1 + (\sigma^2/s^2)}} \right) \nu$$

and the result follows by taking

$$\alpha = \left(\frac{1}{\sqrt{1 + (\sigma^2/s^2)}} \right).$$

□

Theorem A.3 *Let $X \sim F$ and $V \sim U$ be stochastically independent and symmetrically distributed, with medians respectively μ and ν . Suppose that F is absolutely continuous and has full support on \mathbb{R} . Then*

- (i) $\mu = \nu$ implies $P(X \geq V) = 1/2$;
- (ii) $\mu > \nu$ implies $P(X \geq V) > 1/2$;
- (iii) $\mu < \nu$ implies $P(X \geq V) < 1/2$.

PROOF: Recall that $P(X \geq V) = \int U(x) dF(x)$. Case (i) follows immediately by the symmetry of U and the absolute continuity of F . Consider Case (ii). Let $\delta = \mu - \nu > 0$. Define the auxiliary function

$$A(x) = \begin{cases} U(x) & \text{if } x \geq \mu \\ U(x - 2\delta) & \text{if } x < \mu \end{cases}$$

Note that $A(x)$ is bounded and symmetric around μ . Moreover, $U(x) \geq A(x)$ for all x . Therefore,

$$P(X \geq V) = \int U(x) dF(x) \geq \int A(x) dF(x) = \frac{1}{2},$$

where the last equality follows by the symmetry of A and the absolute continuity of F . To strengthen the inequality, note that there must exist an interval $[a, b]$ in $(-\infty, \mu)$ where $U(x) > A(x)$. By the full support assumption, F puts positive probability on I and therefore the strict inequality holds. The proof of Case (iii) follows by reversing the argument just used for Case (ii). □

We say that a random variable V has a unimodal distribution U if there exists a mode ν in \mathbb{R} such that U is convex on $(-\infty, \nu]$ and concave on $[\nu, +\infty)$. If U is unimodal, the set of its modes is a (possibly degenerate) interval. If U is unimodal and has a unique mode ν , it may fail to be continuous at (and only at) ν ; in particular, any degenerate distribution is unimodal.

Theorem A.4 *Suppose that the target has a unimodal distribution. Under the assumptions of Theorem A.3, if there is a certainty equivalent $c(X)$ for X then there exists α in $[0, 1]$ such that $c(X) = \alpha\mu + (1 - \alpha)\nu$.*

PROOF: Recall that $c(X)$ solves the equation $P(X \geq V) = P(c \geq V)$. We distinguish three cases. First, suppose that $\mu = \nu$. By Theorem A.3, $P(X \geq V) = 1/2$. Hence, $P(c \geq V) = 1/2$ and so $c = \nu$. Therefore, $c = \mu = \nu$ and any α will do. Second, suppose that $\mu > \nu$. By Theorem A.3, $P(X \geq V) > 1/2$ and so $c \geq \nu$. To prove the result, we need to show that $\mu \geq c$ or, equivalently, that

$$U(c) - U(\mu) = \int [U(x) - U(\mu)] dF(x) \leq 0.$$

Let $\delta = \mu - \nu > 0$. Define the auxiliary function

$$A(x) = \begin{cases} U(x) - U(\mu) & \text{if } x \geq \mu \\ U(x - 2\delta) - U(\mu - 2\delta) & \text{if } x < \mu \end{cases}$$

Note that $A(x)$ is bounded and symmetric around μ , with $A(\mu + x) + A(\mu - x) = 0$ for almost all x . Moreover, $A(x) \geq U(x) - U(\mu)$ for all x by the unimodality of U . Therefore, by the symmetry of A ,

$$\int [U(x) - U(\mu)] dF(x) \leq \int A(x) dF(x) = 0,$$

as it was to be shown. The third case, when $\mu < \nu$, can be similarly handled. \square

The next theorem is a variation on an almost sure convergence result that is well-known in the literature on stochastic approximation algorithms. Given some arbitrary random variable \bar{X}^0 , consider the stochastic algorithm

$$\bar{X}^{t+1} = \bar{x}^t + \frac{1}{t+1} H(\bar{x}^t, X^t), \quad (6)$$

where $H(x, X)$ is a deterministic function. We provide a sufficient condition for the almost sure convergence of (6) to a point. Our proof uses an elegant argument developed in Robbins and Siegmund (1971), where the following lemma is proved.

Lemma A.5 *Let $\{\mathcal{F}^t\}$ be an increasing sequence of σ -algebras. Suppose that Z^t, A^t, B^t, C^t are nonnegative \mathcal{F}^t -measurable random variables such that*

$$E [Z^{t+1} | \mathcal{F}^t] \leq Z^t(1 + A^t) + B^t - C^t$$

for all $t = 1, 2, \dots$. Then, on the set

$$\left\{ \sum_{t=1}^{+\infty} A^t < \infty, \sum_{t=1}^{+\infty} B^t < \infty \right\},$$

it almost surely occurs that Z^t converges to a finite random variable Z and $\sum_{t=1}^{+\infty} C^t$ is finite.

Theorem A.6 Let \mathcal{F}^t be the σ -algebra generated by $\{X^0, X^1, \dots, X^{t-1}\}$. Suppose that (6) satisfies the following three assumptions:

(i) the term $H(\bar{x}^t, X^t)$ is \mathcal{F}^t -measurable for all t ;

(ii) there exists a constant $k > 0$ such that

$$E \left[(H(\bar{x}, X))^2 \mid \mathcal{F}^t \right] \leq k (1 + |\bar{x}|^2)$$

for all t and all \bar{x} ;

(iii) for

$$h^t(\bar{x}) = E \left[H(\bar{x}, X) \mid \mathcal{F}^t \right],$$

which is well defined by (ii), there exists μ_* such that $\varepsilon > 0$ implies

$$\sup_{\varepsilon \leq |\bar{x} - \mu_*| \leq (1/\varepsilon)} (\bar{x} - \mu_*) h^t(\bar{x}) \leq \gamma$$

for some $\gamma < 0$ which depends on ε but not on t or on the events in \mathcal{F}^t .

Then \bar{X}^t converges almost surely to μ_* .

PROOF: Define $Z^t = (\bar{x}^t - \mu_*)^2$ and, for notational simplicity, let $a_t = 1/(t+1)$. Then

$$Z^{t+1} = Z^t + 2a_t(\bar{x}^t - \mu_*)H(\bar{x}^t, X^t) + a_t^2 \left[H(\bar{x}^t, X^t) \right]^2.$$

Using (ii) and the inequality $(\bar{x}^t)^2 \leq 2Z^t + 2\mu_*^2$, this implies

$$\begin{aligned} E \left[Z^{t+1} \mid \mathcal{F}^t \right] &= Z^t + 2a_t(\bar{x}^t - \mu_*)E \left[H(\bar{x}^t, X^t) \mid \mathcal{F}^t \right] + a_t^2 E \left[\left(H(\bar{x}^t, X^t) \right)^2 \mid \mathcal{F}^t \right] \\ &\leq Z^t + 2a_t(\bar{x}^t - \mu_*)h^t(\bar{x}^t) + ka_t^2 (1 + |\bar{x}^t|^2) \\ &\leq Z^t (1 + 2ka_t^2) + ka_t^2(1 + 2\mu_*^2) + 2a_t(\bar{x}^t - \mu_*)h^t(\bar{x}^t). \end{aligned}$$

Let $A^t = 2ka_t^2$, $B^t = ka_t^2(1 + 2\mu_*^2)$ and $C^t = -2a_t(\bar{x}^t - \mu_*)h^t(\bar{x}^t)$. Note that $\sum A^t < \infty$, $\sum B^t < \infty$ and that C^t is nonnegative by (iii). Therefore, Lemma A.5 implies that it almost surely occurs that Z^t converges to Z and that $\sum a_t(\bar{x}^t - \mu_*)h^t(\bar{x}^t)$ is finite.

It remains to show that $Z = 0$ almost surely. Proceed by contradiction and suppose that $Z(\omega) \neq 0$. Then there exist some $\varepsilon > 0$ and some T such that $\varepsilon \leq |\bar{x}^t - \mu_*| \leq (1/\varepsilon)$ for all $t \geq T$. By (iii), $\sup_{t \geq T} (\bar{x}^t - \mu_*)h^t(\bar{x}^t) \leq \gamma < 0$. But $\sum a_t$ diverges and therefore $\sum a_t(\bar{x}^t - \mu_*)h^t(\bar{x}^t)$ cannot be finite. \square

PROOF OF **Theorem 2**: Equation (5) is a special case of (6) where $H(\bar{x}^t, X^t) = X^t - \bar{x}^t$ and the lottery $X^t = B(V^t)$ is the optimal choice at time t . Therefore we only need

to check that the three assumptions of Theorem A.6 are satisfied. Assumption (i) holds because the optimal choice $X^t = B(V^t)$ is a stochastically independent lottery. Let σ_M be the greatest standard deviation attainable in the (finite) set \mathcal{L} . Then Assumption (ii) holds because

$$E \left[(X^t - \bar{x}^t)^2 \mid \mathcal{F}^t \right] = E \left[(B(V^t) - b(V^t))^2 \mid \mathcal{F}^t \right] + [b(V^t) - \bar{x}^t]^2 \leq \sigma_M^2 + 2[\mu_*^2 + (\bar{x}^t)^2].$$

Finally, since $h^t(\bar{x}) = b^t(V^t) - \bar{x}^t$, Assumption (iii) holds because $|\bar{x} - \mu_*| \geq \varepsilon$ implies $(\bar{x} - \mu_*) h^t(\bar{x}) \leq -\varepsilon \cdot \min\{\varepsilon, \eta\} < 0$. \square

References

- [1] J.R. Blum (1954), “Approximation Methods which Converge with Probability One,” *Annals of Mathematical Statistics* **25**, 382–386.
- [2] R.F. Bordley and M. LiCalzi (1998), “Decision Analysis using Targets instead of Utility Functions,” *mimeo*, December.
- [3] T. Börgers and R. Sarin (1996), “Naive Reinforcement Learning with Endogenous Aspirations,” *mimeo*, February.
- [4] E. Castagnoli and M. LiCalzi (1996), “Expected Utility without Utility,” *Theory and Decision* **41**, 281–301.
- [5] E. Dekel, J.C. Ely, and O. Yilankaya (1998), “Evolution of Preferences,” *mimeo*, July.
- [6] P.C. Fishburn and G.A. Kochenberger (1979), “Two-Piece von Neumann–Morgenstern Utility Functions,” *Decision Sciences* **10**, 503–518.
- [7] M. Friedman and J.L. Savage (1948), “The Utility Analysis of Choices Involving Risk,” *Journal of Political Economy* **56**, 279–304.
- [8] G. Gigerenzer (1996), “On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky (1996),” *Psychological Review* **103**, 592–596.
- [9] I. Gilboa and D. Schmeidler (1996), “Case-Based Optimization,” *Games and Economic Behavior* **15**, 1–26.
- [10] G. Hanoch and H. Levy (1969), “The Efficiency Analysis of Choices Involving Risk,” *Review of Economic Studies* **36**, 335–346.
- [11] C. Heath, S. Huddart, and M. Lang (1999), “Psychological Factors and Stock Option Exercise,” *Quarterly Journal of Economics* **114**, 601–627.
- [12] D. Kahneman, J.L. Knetsch, and R.H. Thaler (1991), “The Endowment Effect, Loss Aversion, and the Status Quo Bias,” *Journal of Economic Perspectives* **5**, 193–206.

- [13] D. Kahneman and A. Tversky (1979), "Prospect Theory: an Analysis of Decision under Risk," *Econometrica* **47**, 263–291.
- [14] M. LiCalzi (1999), "A Language for the Construction of Preferences under Uncertainty," *mimeo*, January.
- [15] L. Ljung, G. Pflug and H. Walk (1992), *Stochastic Approximation and Optimization of Random Systems*, Boston: Birkhäuser.
- [16] R. Marimon (1997), "Learning from Learning in Economics," in: D.M. Kreps and K.F. Wallis (eds.), *Advances in Economics and Econometrics: Theory and Applications*, Volume I, Cambridge, England: Cambridge University Press, 278–315.
- [17] J.G. March (1996), "Learning to Be Risk Averse," *Psychological Review* **103**, 309–319.
- [18] H. Markowitz (1952), "The Utility of Wealth," *Journal of Political Economy* **60**, 151–158.
- [19] J. Meyer (1987), "Two-Moment Decision Models and Expected Utility Maximization," *American Economic Review* **77**, 421–430.
- [20] H. Robbins and S. Monro (1951), "A Stochastic Approximation Method," *Annals of Mathematical Statistics* **22**, 400–407.
- [21] H. Robbins and D. Siegmund (1971), "A Convergence Theorem for Non Negative Almost Supermartingales and Some Applications," in: J.S. Rustagi, *Optimizing Methods in Statistics*, New York: Academic Press, 233–257.
- [22] A.J. Robson (1996), "The Evolution of Attitudes to Risk: Lottery Tickets and Relative Wealth," *Games and Economic Behavior* **14**, 190–207.
- [23] P.H. Rubin and C.W. Paul, II (1979), "An Evolutionary Model of Taste for Risk," *Economic Inquiry* **17**, 585–596.
- [24] T.J. Sargent (1993), *Bounded Rationality in Macroeconomics*, Oxford: Clarendon Press.