THEORETICAL NOTE

Adaptive Learning and Risk Taking

Jerker Denrell
Stanford University

Humans and animals learn from experience by reducing the probability of sampling alternatives with poor past outcomes. Using simulations, J. G. March (1996) illustrated how such adaptive sampling could lead to risk-averse as well as risk-seeking behavior. In this article, the author develops a formal theory of how adaptive sampling influences risk taking. He shows that a risk-neutral decision maker may learn to prefer a sure thing to an uncertain alternative with identical expected value and a symmetric distribution, even if the decision maker follows an optimal policy of learning. If the distribution of the uncertain alternative is negatively skewed, risk-seeking behavior can emerge. Consistent with recent experiments, the model implies that information about foregone payoffs increases risk taking.

Keywords: decisions from experience, risk taking, decisions under uncertainty, adaptive behavior, learning from outcome feedback

Much research on decisions under uncertainty has focused on risk attitudes of individuals, defined by the curvature of the utility or value function, and how they vary with the context and past outcomes (Gonzáles-Vallejo, Reid, & Schiltz, 2003; Kahneman & Tversky, 1979; Lopes & Oden, 1998). But decisions under uncertainty are influenced by many other factors that may mask the relationship between risk attitudes and observed decisions. One important influence is the information decision makers have about alternatives. Outside the laboratory, individuals seldom know the outcome distributions of the alternatives between which they can choose. As emphasized in the recent literature on decision from experience (Barron & Erev, 2003; Erev & Barron, 2005; Hertwig, Weber, Barron, & Erev, 2004), decision makers may not have access to descriptions of outcome distributions, as in experiments, but only to samples of past outcomes. The resulting sampling variability weakens the association between risk attitudes and observed choices. It is not clear whether individuals who choose an uncertain alternative instead of a sure thing are risk seeking or only have experienced a favorable sequence of outcomes when sampling the uncertain alternative.

Such sampling variability can also have systematic effects on decision under uncertainty, producing apparent risk-averse and risk-seeking behavior (Denrell & March, 2001; March, 1996). To illustrate this, March (1996) examined how decision makers following different learning rules allocate their choices between a certain alternative and an uncertain alternative with an unknown payoff. In his setup, the certain alternative paid k and the uncertain alternative paid k / r with probability r and zero with probability $1 - r$. For example, if $k = 1$ and $r = .01$, the uncertain alternative pays 100 with probability .01 and zero otherwise, whereas the certain alternative pays 1. Decision makers could only observe the payoff of the chosen alternative. Using simulations of several models of learning, such as the Bush–Mosteller algorithm (Bush & Mosteller, 1955), March (1996) showed that if k was positive and $r < .5$, the proportion of choices of the uncertain alternative quickly fell below 50%.[1]

Two mechanisms contribute to this result. First, if r is small and k is positive, the favorable outcome (k / r) is a rare event. The uncertain alternative then generates a payoff of zero most of the time, and the certain alternative will seem to be superior most of the time. If decision makers rely on recent outcomes when estimating the value of the uncertain alternative, most will underestimate the uncertain alternative. Second, most learning models assume that decision makers revise the probability of sampling the uncertain alternative in response to outcome feedback. To avoid unfavorable future outcomes, decision makers reduce the probability of sampling the uncertain alternative if past outcomes have been poor. Such an adaptive sampling rule implies that decision makers who have a low estimate of the expected value of the uncertain alternative are unlikely to choose it. If information is only available about the chosen alternative, they are unlikely to experience the occasional high payoff (k / r) and to revise their low estimate.

---

[1] Burgos (2002) found similar results using a different learning algorithm. Denrell and March (2001) found similar results using a learning model in which payoffs are compared with a changing aspiration level.

Several researchers have recently examined the first of these mechanisms, the bias in estimates generated by rare events (Erev & Barron, 2005; Hertwig, Barron, Weber, & Erev, 2006; Hertwig et al., 2004; Weber, Shafir, & Blais, 2004). These studies show that decision makers may underestimate the expected value of an uncertain alternative that usually generates a poor payoff but occasionally generates a very favorable payoff. The second mechanism, adaptive sampling, has received less attention (although, see Barron & Erev, 2003). However, adaptive sampling can also lead to a systematic bias in estimates (Denrell, 2005; Denrell & March, 2001; Fazio, Eiser, & Shook, 2004). If decision makers overestimate the expected payoff of an uncertain alternative, they are likely to sample it again and can correct the error. If decision makers underestimate the expected payoff of an alternative, they may avoid it and thus cannot correct the error. As a result, most individuals may end up underestimating an uncertain alternative, even if it has a symmetric distribution (Denrell, 2005).

This article clarifies the implications of adaptive sampling for risk taking in decisions from experience. The purpose is not to suggest a comprehensive descriptive model of risk taking in decisions from experience (for such a model, see Erev & Baron, 2005), but to examine the theoretical and empirical implications of the mechanism of adaptive sampling. I demonstrate formally that common assumptions about learning imply that even a risk-neutral decision maker would learn to prefer a certain alternative to an uncertain alternative with identical expected value and a *symmetric* distribution. I first illustrate this in the next section for the case in which the decision between the two alternatives follows a logit choice rule based on an estimate of the uncertain alternative. In the subsequent section, I show that the basic conclusion holds, even if a risk-neutral decision maker would follow the decision rule that maximizes total expected value in a repeated choice between an uncertain and unknown alternative and a known and certain alternative (the so-called "one-armed bandit problem"). This demonstrates that apparent risk-averse behavior can be the result of a normatively appropriate policy for managing the trade-off between exploration and exploitation in sequential choice under uncertainty, as suggested by March (1996), Denrell & March (2001), and Niv, Joel, Meilijson, and Ruppin (2002). I next show that a model of adaptive sampling also has distinct empirical implications. Consistent with recent experiments, a model of adaptive sampling implies that information about foregone payoffs will increase the tendency to choose risky alternatives. Finally, I illustrate that learning from adaptive samples could also generate apparent risk-seeking behavior if the uncertain alternative has a negatively skewed distribution. In the concluding section, I discuss the significance of adaptive sampling for understanding decision making under uncertainty.

Adaptive Sampling Is Sufficient to Generate Risk-Averse Behavior

An Illustrative Model

Consider a decision maker faced with a repeated choice between a certain alternative and an unknown and uncertain alternative. The decision maker knows that the certain alternative always generates a payoff of zero. The payoff generated by the uncertain alternative is a normally distributed random variable, X, with mean zero and variance σ^2. To demonstrate that adaptive sampling is sufficient to generate risk-averse behavior, I assume in this section that the decision maker cares only about the expected values of the two alternatives.

The decision maker does not know the expected value of X but has to rely on past outcome information to estimate it and choose between the two alternatives. To illustrate the effect of adaptive sampling, I first consider a simple model that assumes that the decision in each period between the two alternatives follows a logit choice rule based on an estimate of the uncertain alternative (Denrell, 2005). I also initially abstract away from other influences on decision making from experience, including loss aversion and the tendency for more variable outcomes to produce more random choices (the "payoff variability effect," Erev & Barron, 2005).

Following Busemeyer and Myung (1992), March (1996), Sarin and Vahid (1999), Busemeyer and Stout (2002), and Barron and Erev (2003), I assume that the decision maker's estimate of the expected value of the uncertain alternative in period t, denoted $\hat{x}_t$, is a weighted average of the previous estimate and the observed payoff, if any. Specifically, $\hat{x}_t = (1 - b)\hat{x}_{t-1} + bx_t$, if the uncertain alternative is chosen in period t, and $\hat{x}_t = \hat{x}_{t-1}$ otherwise. Here $\hat{x}_0 = 0$ and b is a positive fraction regulating the weight of the new observation. Because information is only available about the payoff of the chosen alternative, the decision maker only updates the estimate of the uncertain alternative when he or she chooses the uncertain alternative.

I assume that the decision maker mainly samples the uncertain alternative if its estimated value is positive, but sometimes explores it even if its estimated value is negative. Specifically, the probability of choosing the uncertain alternative in period $t + 1$ is assumed to follow the exponential version of the Luce choice rule (Luce, 1959): $P_{t+1} = 1/(1 + e^{-S\hat{x}_t})$.[2] Here S is a parameter regulating how sensitive the choice probability is to the value of the estimate. As $S \to \infty$, the uncertain alternative is only chosen if its estimated value is positive. Experiments on repeated choices between uncertain alternatives have shown that this logit choice model provides a good fit to choice data (Busemeyer & Stout, 2002; Yechiam & Busemeyer, 2005). Moreover, Gans, Knox, and Croson (2004) and Busemeyer and Stout (2002) showed that a logit choice model combined with a weighted-average model for belief updating provided the best overall fit to data from experiments on repeated choices between uncertain alternatives (Yechiam & Busemeyer, 2005, however, concluded that a different model of belief updating provided the best fit).

Given these assumptions, what is the probability of choosing the uncertain alternative? Figure 1 shows how the probability of choosing the uncertain alternative changes over time for different values of S when $b = .5$. The probability of choosing the uncertain alternative is always below 50% after the first period and declines over time. It is possible to derive an explicit formula for the expected probability that the uncertain alternative is chosen as

---

[2] It is also assumed that the uncertain alternative is chosen in the first period. The results do not change if this assumption is changed.
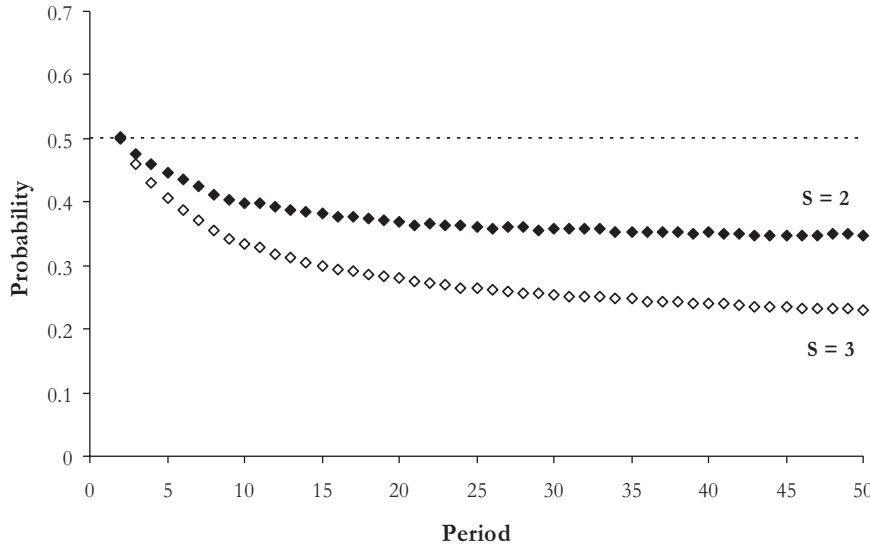
Figure 1. The average probability of choosing the uncertain alternative for two different values of *S*. Each entry is based on averages from 100,000 simulations in which $b = .5$ and $\sigma^2 = 1$.

$t \rightarrow \infty$. As demonstrated in Appendix A, the expected (asymptotic) probability of choosing the uncertain alternative, denoted $E(C_U)$, is

$$E(C_U) = \frac{1}{1 + e^{\frac{S^2b}{2(2-b)}\sigma^2}} < 0.5. \quad (1)$$

Thus, the decision maker behaves as if he or she is risk averse. Simulations show that the prediction of this asymptotic solution is very close to the average probability of choosing the uncertain alternative in, say, Period 50.[3] The asymptotic solution implies that the probability of choosing the uncertain alternative is a decreasing function of both *S* and *b*. Only if they are both large will the bias be large. To illustrate the size of the bias, suppose that $S = 4.81$ and $b = .25$. These are median estimates from an experiment by Gans et al. (2004) in which participants repeatedly choose between two uncertain alternatives with Bernoulli distributions.[4] If the uncertain alternative pays off $+.5$ or $-.5$ with equal probability and the certain alternative has a payoff of zero, these estimates imply that the probability of choosing the uncertain alternative is .40 after 50 periods (based on 10,000 simulations).

The bias against the uncertain alternative is due to the asymmetry in the probability of over- and underestimation. As demonstrated in Denrell (2005), the above assumptions imply that a majority of all decision makers will have a negative estimate of the uncertain alternative. This negativity bias occurs only when the probability of sampling the uncertain alternative depends on the value of the estimate and when information is available only about the payoff of the chosen alternative. If the probability of sampling was exogenous (i.e., independent of the value of the estimate) or if information was always available about the payoff of the uncertain alternative, the probability of choosing the uncertain alternative would be .5.

Suppose next that the expected value of *X* was not zero but $m \neq 0$. In this case, the asymptotic probability of choosing the uncertain alternative under adaptive sampling (as $t \rightarrow \infty$) is

$$E(C_U) = \frac{1}{1 + e^{-Sm + \frac{S^2b}{2(2-b)}\sigma^2}}. \quad (2)$$

Equation 2 implies that the probability of choosing the uncertain alternative will be a strictly increasing function of the expected value (*m*) and a strictly decreasing function of the variance (σ^2) of the uncertain alternative. The probability of choosing the uncertain alternative will only be 50% if the expected value is positive. If $\sigma^2 = 1$, $S = 3$, and $b = .5$, the expected value (*m*) has to be .5 for probability of choosing the uncertain alternative, $E(C_U)$, to be .5.

Equation 2 is also identical to the probability of choosing the uncertain alternative for a risk-averse decision maker who knows the distribution of the uncertain alternative, who only cares about the expected value and the variance of the uncertain alternative, and whose decisions follow the logit choice rule. A model of learning from experience assuming risk neutrality thus generates exactly the same choice probabilities as does a logit random-utility model (McFadden, 1974) that assumes risk aversion and known outcome distributions. This equivalence between the two models implies that estimates in field studies of a negative effect of variability in outcomes on choice probabilities do not necessarily provide evidence for a distaste of dispersion but may be due to learning. A possible way to distinguish these hypotheses is to estimate choice models that include recent experiences as a proxy for the belief ($\hat{x}_t$) of the decision maker. If decisions are mainly influenced by a belief ($\hat{x}_t$) based on recent experiences, the estimated effect of the variance in outcomes should disappear.

---

[3] This holds true unless *S* is very large, which implies that the uncertain alternative may be avoided for many periods. A simulation of 50 periods will underestimate the duration of such avoidance.

[4] I am grateful to the authors for providing me this information.

The probability of choosing the uncertain alternative is a continuously decreasing function of the variance of the uncertain alternative, σ^2, because the choice probability is a continuous function of the value of the estimate. A different choice rule would give different results. Suppose the alternative with the highest estimated value is chosen with probability $1 - q$, where $q < .5$. The asymptotic probability of choosing the uncertain alternative is then $2q(1 - q)$ (see Appendix A), which is less than .5 but not a continuously decreasing function of the variance of the uncertain alternative.

Generalization

The result that adaptive sampling generates risk-averse behavior is not limited to the case in which the payoff of the uncertain alternative is normally distributed and sampling follows the logit choice rule but holds for a somewhat broader set of assumptions about outcome distributions and choice rules (cf. Niv et al., 2002).

Consider again a repeated choice between an uncertain and a certain alternative. The certain alternative generates a payoff of zero. The payoff generated by the uncertain alternative is a random variable, X. Rather than assuming that X is normally distributed, it is sufficient to assume that X is a random variable with mean zero and with a density that is symmetric around zero: $f(x) = f(-x)$. In the section titled "When Can Adaptive Sampling Generate Risk-Seeking Behavior?" I treat the case in which X has a skewed distribution. Rather than assuming that sampling follows the logit choice rule, it is sufficient to assume that the probability of sampling is an increasing function of the estimate. Specifically, $P(x) \geq P(y)$ whenever $x > y$, with strict inequality for at least some values, x and y, and $P(x) > 0$ for all x. To avoid any inbuilt bias for or against the uncertain alternative, I also assume that $P(x)$ is symmetric around zero, that is, $P(x) = 1 - P(-x)$. Thus, decision makers treat gains and losses symmetrically.

Finally, rather than assuming that the weight of the most recent observation is constant, it is sufficient to assume that the estimate at the end of period $t + 1$, $\hat{x}_{t+1}$, is some weighted average of the N most recent observations. Formally, $\hat{x}_{t+1} = b_1 x_1 + b_2 x_2 + \ldots + b_N x_N$, where x_N is the most recent observation, $b_i \in (0,1)$, and

$$\sum_{i=1}^{N} b_i = 1.$$

This assumption includes, or can approximate, a wide range of possible models of belief formation. For example, the above formulation can approximate very well the model used in the previous section (and also in March, 1996) in which the weight of the most recent observation was constant. It also includes, or can approximate, models in which the weight of the most recent observation declines with experience (unless the weight declines to zero). This assumption does not include a model in which the estimate is the average of all observations made so far.

If information is only available about the chosen alternative, these assumptions imply that decision makers will behave as if they were risk averse: The expected probability that the uncertain alternative is chosen (as $t \to \infty$) is below 50% (see Appendix B).[5] The reason for such risk-averse behavior is that sampling is assumed to be adaptive. Suppose, instead, that the decision maker had to base his or her decision on an estimate formed on the basis

of a given sample of N observations of X. If the estimate and the choice rule followed the assumptions outlined in the above paragraphs, the expected probability of choosing the uncertain alternative would be .5 (see Appendix B).

Even Optimal Sequential Sampling Can Generate Risk-Averse Behavior

The bias against the uncertain alternative occurs in the above models because decision makers are likely to avoid the uncertain alternative if recent outcomes were poor. Although such an adaptive sampling rule is reasonable, it is not necessarily optimal. Even if initial outcomes are poor, it could be optimal to continue to sample the uncertain alternative to obtain further information about its payoff distribution. If a decision maker followed such an optimal sampling policy, would the bias against the uncertain alternative still hold? Or do the above results only hold when decision makers are myopic, in the sense that they ignore the value of experimentation?

The bias against the uncertain alternative holds, even if the decision maker follows an optimal policy of experimentation. To illustrate this, consider a risk-neutral decision maker who can choose, in N periods, between a known alternative with a certain payoff of zero and an uncertain alternative. The payoff generated by the uncertain alternative is normally distributed with an unknown mean, m, and known variance σ_o^2. If m is positive, the decision maker would prefer to choose the uncertain alternative. However, the decision maker does not know the value of m characterizing the uncertain alternative with which he or she is faced. The decision maker only knows that the value of m is drawn from a normal distribution with mean zero (i.e., the prior of the decision maker is normal with mean zero). Thus, ex ante, the value of m is equally likely to be positive or negative.

The decision maker has to rely on outcome information to estimate the value of m and to decide which alternative to choose in each period. Rather than assuming that the decision maker follows some heuristic choice rule, I assume the decision maker chooses, at the beginning of period t, the alternative that maximizes the expected total payoff in the remaining periods

$$t, \ldots, N, E\left(\sum_{i=t}^{N} X_i\right),$$

where X_i is the payoff of the alternative chosen in period i. Thus, the decision maker takes into account the value of experimentation. I also assume that the estimate of the decision maker follows Bayes rule.

The optimization problem facing the decision maker is a version of the classical one-armed bandit problem (Berry & Fristedt, 1985). An optimal policy in this problem needs to specify when past outcomes justify that the uncertain alterna-

[5] Niv et al. (2002) derived a similar result, but only under the slightly stronger assumption that $1 / P(x)$ is a convex function. In related articles Börgers, Morales, and Sarin (2004) and Oyarzun and Sarin (2006) examined the assumptions regarding the mapping between the realized payoff and changes in choice probabilities required for a learning rule to be (locally) risk neutral or risk averse.

tive is abandoned, even if this implies that no new information about its expected value will be obtained. Choosing the uncertain alternative provides additional information about its expected value, which is valuable because it makes possible more informed choices in the future. If the estimated expected value of the uncertain alternative is lower than the payoff of the certain alternative, obtaining such information is costly because information can only be obtained by choosing a seemingly inferior alternative. An optimal policy has to find a balance between exploring the uncertain alternative and exploiting the certain alternative, if it is believed to be better.

Even if risk-neutral decision makers followed an optimal policy, fewer than 50% of all decision makers will choose the uncertain alternative in the last period. The proof of this is very simple and follows from a few well-known characteristics of the optimal policy in this finite-horizon one-armed bandit problem (for a characterization of the optimal policy see Burnetas & Katehakis, 1997).[6]

First, because the expected values of the uncertain and the certain alternatives are equal at the beginning of the first period, it is optimal to choose the uncertain alternative in the first period to obtain information about its distribution. Second, if the decision maker ever decides to switch to the certain alternative, it is never optimal to switch back to the uncertain alternative (Berry & Fristedt, 1985), because no further information is available. Third, in the last period, the alternative with the highest expected value, given past outcomes, should be chosen. This implies that the uncertain alternative is only chosen in the last period if

$$\sum_{i=1}^{N-1} x_{i,u} > 0,$$

where $x_{i,u}$ is the payoff of the uncertain alternative in period i. Fourth, in all periods before the last, it is optimal to choose the uncertain alternative in period t if its expected value given past outcomes, $\hat{m}_{t-1}$, is higher than some (negative) threshold value (Burnetas & Katehakis, 1997). This implies that the uncertain alternative is only chosen in period $t = 2, \ldots, N - 1$ if the sum of its past outcomes is higher than some negative threshold (Burnetas & Katehakis, 1997):

$$\sum_{i=1}^{t-1} x_{i,u} > -c_{t-1}, \forall t : c_t > 0.$$

Overall, this implies that the decision maker chooses the uncertain alternative in the final period only if the following conditions hold:

$$x_{1,u} > -c_1 \cap x_{1,u} + x_{2,u} > -c_2 \cap \ldots \cap \sum_{i=1}^{N-2} x_{i,u}$$

$$> -c_{N-2} \cap \sum_{i=1}^{N-1} x_{i,u} > 0. \quad (3)$$

The probability of this combined event is lower than the probability of any one of the events (because they are not identical events). In particular, the probability of the combined event is lower than

$$P(\sum_{i=1}^{N-1} X_{i,u} > 0) = 0.5.$$

Thus, if a decision maker follows an optimal policy, the probability that the uncertain alternative will be chosen in the last period is less than 50%. The same reasoning implies that most decision makers underestimate the expected value of the uncertain alternative. The above argument shows that for more than 50% of all decision makers, the expected value of the uncertain alternative at the beginning of period N, given the past outcomes experienced by the decision maker, denoted $\hat{m}_{N-1}$, is negative. It can also be demonstrated that the probability that $\hat{m}_t$ is negative is larger than 50% for all $t > 1$.[7]

To illustrate how the probability of choosing the uncertain alternative changes over time, the optimal policy needs to be calculated. Although the optimal policy for $N > 3$ is very difficult to calculate for a one-armed bandit with a normal distribution, it can be computed, using stochastic dynamic programming, for a one-armed bandit with a Bernoulli distribution (e.g., Berry & Fristedt, 1985). Suppose, for example, that a decision maker can choose in 10 periods between a known alternative with a certain payoff of 0 and an uncertain alternative, which pays off 1 with probability q and -1 with probability $1 - q$. The decision maker does not know q but has a correct prior about q: q is a random variable drawn from a uniform distribution between zero and 1.

Figure 2 shows how the probability of choosing the uncertain alternative changes over time if the decision maker follows an optimal policy in this one-armed bandit problem. The probability of choosing the uncertain alternative is initially 1, reflecting the value of exploring the uncertain alternative to obtain information about its expected value. The probability of choosing the uncertain alternative then quickly falls below 50%, as more and more decision makers believe that it is not useful to continue to explore the uncertain alternative. In the 10th period only 35.5% of all decision makers choose the uncertain alternative. This also implies that most decision makers underestimate the uncertain alternative at the end of the 10th period. In this case, $\hat{m}_{10}$, the expected value of the uncertain alternative at the end of the 10th period, given the past outcomes experienced by the decision maker, is negative for 64.7% of all decision makers, is equal to zero for 1.6%, and is positive for 33.8% (based on 100,000 simulations).

Although it may seem peculiar that the optimal policy should have these implications, it is more intuitive once one considers the "costs" of correcting errors of under- and overestimation. Suppose the uncertain alternative is mistakenly classified as inferior to the certain alternative, an error of underestimation. Because new information about the uncertain alternative can only be gained if the uncertain alternative is sampled, and the uncertain alternative is believed to be inferior, sampling the uncertain alternative is costly

---

[6] Note that the Gittins (1979) index does not apply to this bandit problem because it is a finite-horizon bandit problem.

[7] To show this, note that it is possible that a decision maker will abandon the uncertain alternative after the first period if the payoff is sufficiently negative, $X_1 < -c_1$. If so, the decision maker will continue to believe that the uncertain alternative has a negative expected value. The probability that $\hat{m}_2$ is positive is thus $P(X_1 > -c_1 \cap X_1 + X_2 > 0)$, which is less than .5. The same argument can be applied to any period.
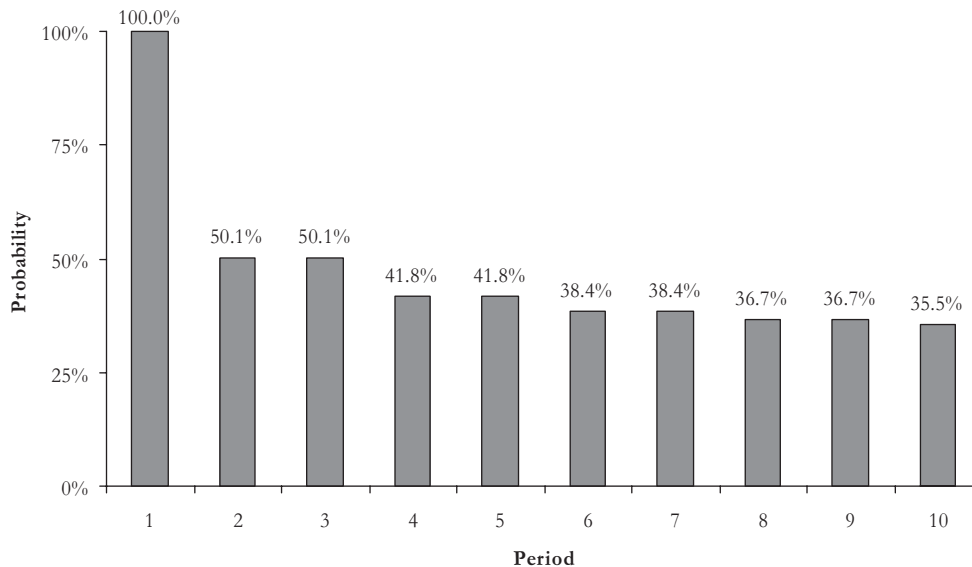
Figure 2. The probability of choosing the uncertain alternative over time, based on 100,000 simulations of the optimal policy in a 10-period one-armed bandit problem in which the uncertain arm has a Bernoulli distribution and the prior has a uniform distribution.

in the sense that it is believed to generate an immediate payoff lower than that of the certain alternative. If the uncertain alternative is mistakenly classified as superior to the certain alternative, however, sampling the uncertain alternative is not costly because it is believed to generate a higher immediate payoff. Thus, it is more costly to correct errors of underestimation than errors of overestimation, and even an optimal policy has a tendency to produce more errors of underestimation than overestimation.

Calculating the optimal policy is very difficult, and experiments on bandit problems show that participants do not follow an optimal policy (Gans et al., 2004; Meyer & Shi, 1995). The value of the above analysis of the consequences of following an optimal policy is that it illustrates a different rationale for risk-averse behavior. The analysis shows why even a risk-neutral decision maker would prefer a policy that results in a tendency to choose a certain alternative instead of an uncertain alternative with identical expected value.

Information About Foregone Payoffs Increases Risk Taking

The results so far illustrate that even a risk-neutral decision maker may learn to prefer a certain alternative to an uncertain alternative with identical expected value. These results suggest that apparent risk-averse behavior should not necessarily be attributed to nonlinear value functions but could also be due to adaptive sampling. In reality, decisions may be influenced both by properties of the value function and adaptive sampling. A decision maker may be both loss averse and sample adaptively, which makes the impact of each difficult to identify. Nevertheless, it is possible to detect the influence of adaptive sampling by examining the effect of information about foregone payoffs. Adaptive sampling may reduce the tendency to choose an uncertain alternative instead of a sure thing only when information about foregone payoffs is not available. If information

about foregone payoffs is available, a decision maker who follows an adaptive sampling rule should thus be more likely to choose an uncertain alternative in a repeated choice between an uncertain alternative and a sure thing. This holds even if the value function of the decision maker is nonlinear.

This effect of information about foregone payoffs on risk taking can be demonstrated formally under quite general assumptions about the value function of the decision maker. Consider a decision maker who repeatedly chooses between a certain alternative and an uncertain alternative with an *arbitrary* (continuous or discrete) distribution. Regarding the value function, it could be assumed that the value of an outcome of X is determined according to a concave value or utility function. Alternatively, the decision maker may evaluate gains and losses differently. For example, it could be assumed that the value of a gain of $x > 0$ is $v(x) = x$, whereas the value of a loss of $-x < 0$ is $v(-x) = -\lambda x$, where $\lambda > 1$ (Kahneman & Tversky, 1979). Introducing loss aversion or a concave utility function does not change the basic setup because these assumptions only imply that one arbitrary probability distribution, of the outcomes, X, is transformed into another, of the values or utilities, $v(X)$. The estimate is formed according to the assumptions outlined under the *Generalization* heading in the section titled "Adaptive Sampling is Sufficient to Generate Risk-Averse Behavior," but the estimate is based on the experienced value or utility, $v(x)$, rather than the outcome, x. The probability of choosing the uncertain alternative in period t depends on its estimated value at the beginning of period t. Regarding the choice rule, I only assume that $P(v_1) \geq P(v_2)$ whenever $v_1 > v_2$, with strict inequality for at least some values, v_1 and v_2, and $P(v) > 0$ for all v.

Two different information conditions are contrasted. In the first, information is only available about the payoff generated by the chosen alternative. This implies that the decision maker only

updates the estimate of the value of the uncertain alternative in periods in which it is chosen. Let $E(C_U)$ denote the expected (asymptotic) probability of choosing the uncertain alternative in this information condition. In the second information condition, information is available in every period about the payoffs generated by both alternatives. This implies that the decision maker updates the estimate of the value of the uncertain alternative in every period. Let $E(C_{U,I})$ denote the expected (asymptotic) probability of choosing the uncertain alternative in this information condition.

As shown in Appendix B, the expected (asymptotic) probability of choosing the uncertain alternative is higher in the second information condition, in which information is available about foregone payoffs: $E(C_{U,I}) > E(C_U)$. This prediction is consistent with findings from recent experiments (see Erev & Barron, 2005, and Yechiam & Busemeyer, 2006, for reviews and discussion). For example, in an experiment by Haruvy and Erev (2001, see also Erev & Barron, 2005) participants repeatedly chose, in 200 trials, between a certain alternative that paid off 10 and an uncertain alternative that paid off 21 or 1 with equal probability. The probability of choosing the uncertain alternative in the last 100 trials was .59 if no information was available about foregone payoffs and .68 if information about foregone payoffs was available (Erev & Barron, 2005, Table 1). A similar but much smaller effect occurred if all payoffs were multiplied by −1.

Yechiam and Busemeyer (2006) showed that information about foregone payoffs can have an even more dramatic effect. In their experiment, participants repeatedly chose, in 400 trials, between two uncertain alternatives with different risk levels. The first paid −2 with probability .995 and −8 with probability .005, whereas the second paid −1 with probability .995 and −300 with probability .005. Although the expected values of the two alternatives are similar (−2.03 vs. −2.50), the variance of the second is substantially higher. Half of the participants were only informed about the outcome for the chosen alternative, whereas the rest were also informed in every second period about the outcome of the other alternative. The average proportion choosing the risky alternative in the last 50 periods was 70% if information about foregone payoffs was available but only 45% otherwise. Similar results were found in an experiment on repeated choices between four uncertain alternatives (the Iowa Gambling task; Yechiam & Busemeyer, 2005) and in an experiment on repeated choices between 100 alternatives (Grosskopf, Erev, & Yechiam, 2006).

When Can Adaptive Sampling Generate Risk-Seeking Behavior?

Adaptive sampling does not always generate apparent risk-averse behavior, in the sense that $E(C_U) < .5$. The above models show that apparent risk-averse behavior can emerge when the distribution of the uncertain alternative is symmetric. If the distribution of the uncertain alternative is not symmetric, the first of the mechanisms mentioned in the introduction—the bias in estimates generated by rare events—also matters and can lead to apparent risk-seeking behavior.

Consider a decision between a certain alternative with a payoff of 9 and an uncertain alternative that pays 10 with probability .9 and zero otherwise. The uncertain alternative has a negatively skewed distribution (the median is above the

mean), and it usually generates a payoff higher than the payoff of the certain alternative. Because its payoff is usually higher, the uncertain alternative may seem superior to the certain alternative if only small samples of the outcomes of each alternative are available. Such a bias in estimates can lead to apparent risk-seeking behavior (Hertwig et al., 2004; Weber et al., 2004). For example, in the experiment by Weber et al. (2004), participants were asked to sample each of the above alternatives for as many times as they wanted to and then make a decision (thus, sampling was exogenous). Most participants chose the uncertain alternative (76%).

Suppose that sampling instead is adaptive and information is only available about the chosen alternative. The uncertain alternative may then initially seem superior, but continued sampling of the uncertain alternative will reveal that it sometimes generates a payoff of zero, and this will reduce the probability of choosing the uncertain alternative. But it does not follow that the probability of choosing the uncertain alternative will eventually fall below 50%. In fact, in an experiment by Barron and Erev (2003), in which participants could choose between the two alternatives described in the above paragraph in 400 periods and information was only provided about the payoff of the chosen alternative, 56% of all participants chose the uncertain alternative (averaged over all periods).

A simple model can illustrate when such risk-seeking behavior persists under adaptive sampling. Consider a decision maker repeatedly choosing between a certain alternative with payoff k and an unknown and uncertain alternative that pays k/r with probability r and zero otherwise. Suppose the estimate of the uncertain alternative is $\hat{x}_t = (1 - b)\hat{x}_{t-1} + bx_t$ if it is chosen in period t and $\hat{x}_t = \hat{x}_{t-1}$ otherwise. Moreover, suppose that the alternative with the highest estimated value at the beginning of period t is chosen in period t with probability $1 - q$, and the other alternative is chosen with probability $q < .5$. Thus, q is the probability of exploring the seemingly inferior alternative.

Let Z be the random variable to which the estimate of the uncertain alternative converge if it is sampled infinitely often and sampling is exogenous. The probability of choosing the uncertain alternative, if sampling is adaptive, converges to (see Appendix A):

$$E(C_{U,q}) = \frac{q(1 - q)}{1 - q - P(Z > k)(1 - 2q)}. \qquad (4)$$

This probability is increasing in $P(Z > k)$, the probability of overestimating the uncertain alternative if sampling is exogenous. If $r \leq .5$ and the uncertain alternative most often generates a payoff lower than k, then $P(Z > k) \leq .5$, which implies that $E(C_{U,q}) < .5$, and the model predicts risk-averse behavior. But if $r > .5$ and $P(Z > k)$ and q are sufficiently large, $E(C_{U,q})$ can be larger than .5, and the model predicts risk-seeking behavior. For example, if $k = 9$ and $r = .9$ (the case described above) and b (the weight of the most recent observation) is .5, then $P(Z > k)$ is approximately .7235 (based on 50,000 simulations of 5,000 periods), and $E(C_{U,q})$ is higher than .5 whenever $.28 < q < .5$ and is at a maximum of 52.8% when $q = .38$. $E(C_{U,q})$ will be higher if r, and thus also $P(Z > k)$, is higher.

This model illustrates that persistent risk-seeking behavior (as $t \to \infty$) under adaptive sampling requires (a) that the distribution of the uncertain alternative is negatively skewed (i.e., $r > .5$), which implies $P(Z > k) > .5$, and (b) that decision makers are quite likely

to choose the uncertain alternative even after an unfavorable recent outcome (i.e., q is high but less than .5, which implies random choice). These implications of the model are consistent with findings from experiments in which participants have to choose repeatedly between a certain and an uncertain alternative without knowing the payoff distribution of the uncertain alternative. When risk-seeking behavior occurred in these experiments, (a) the distribution of the uncertain alternative was negatively skewed and (b) although the probability of choosing the uncertain alternative declined after an unfavorable outcome, it still remained at a relatively high level (Erev & Barron, 2005).

The model also illustrates that adaptive sampling can lead to risk-seeking behavior in the domain of gains and not only in the domain of losses, in contrast to the simulation results of March (1996), which showed that risk-seeking behavior only occurs in the domain of losses. In the above models, however, the domain of the payoffs does not influence the results. Consider a choice between a certain alternative with a payoff of -1 and an uncertain alternative that pays off zero with probability .9 and -10 otherwise. Adding 10 to each outcome, we get the decision problem introduced in the beginning of this section. For any choice rule that only depends on the difference in the estimates of the uncertain alternative and the certain, such as the logit choice rule and the choice rule that chooses the alternative with the highest estimated value with probability $1 - q$, the two decision problems are identical. March (1996) could not generate risk-seeking behavior in the domain of gains because of the choice rule he assumed. He assumed that the probability of choosing the uncertain alternative was $\hat{x}_{t,u}/(\hat{x}_{t,u} + \hat{x}_{t,c})$, where $\hat{x}_{t,u}$ is the estimate of the uncertain alternative and $\hat{x}_{t,c}$ the estimate of the certain. As discussed in the appendix to March (1996), this choice rule implies that if the uncertain alternative pays 10 with probability .9 and zero otherwise and the certain alternative pays 9, the probability of choosing the uncertain alternative can never be above $10/(9 + 10) \approx 0.53$, whereas the lower limit is zero.

Overall, the model introduced in this section suggests that whether learning from experience can lead to risk-seeking behavior in a repeated choice between a sure thing and an unknown uncertain alternative depends on whether the uncertain alternative has a negative skew and does not depend on the domain of the payoffs. This conclusion is consistent with recent experiments on decision from experience, which have found risk-seeking behavior both in the domain of gains and losses but mainly when the distribution of the uncertain alternative was negatively skewed (Barron & Erev, 2003; Erev & Barron, 2005).

Although the model introduced in this section shows that a negatively skewed uncertain alternative can lead to risk-seeking behavior, the model also implies that whether such risk-seeking behavior occurs depends on the choice rule. Risk-seeking behavior will not emerge in the model introduced in this section if q is low. More generally, the effect of a negative skew depends on how sensitive the choice rule is to recent outcomes. The model introduced in this section, in which the alternative with the highest estimate is chosen with probability $1 - q$, implies that decision makers are more risk seeking when choosing between a sure thing and a negatively skewed uncertain alternative, consistent with experiments on decisions from experience (Barron & Erev, 2003). However, simulations show

that if the choice rule is the logit, $P_{t+1} = 1/(1 + e^{-S\hat{x}_t})$, and S is large, decision makers can be more risk seeking when choosing between a sure thing and a positively skewed uncertain alternative. Such a preference for positively skewed uncertain alternatives is consistent with experiments with animals (Shafir, Bechar, & Weber, 2003) and with experiments in which participants do not have to learn but are told about the distributions of the alternatives between which they can choose (Weber et al., 2004). It is thus possible that the different results from experiments on skewness preferences could be explained, in part, by differences in sensitivity to recent outcomes.

The Empirical Significance of Adaptive Sampling

How important is adaptive sampling in explaining risk-taking behavior? Clearly, it is only a piece of the puzzle. Properties of the value function, such as loss aversion, have an important impact on decisions between uncertain alternatives, even in decisions from experience (Erev & Barron, 2005). But recent experiments have shown that adaptive sampling is also important in modeling risk-taking behavior in decisions from experience. Based on a review of several experiments, Erev and Barron (2005) concluded that descriptive models of risk taking in decisions from experience need to take into account the effect of adaptive sampling (what they call the "stickiness effect"). Models that ignore adaptive sampling exaggerate the probability that decision makers will choose alternatives with high expected values and high variability. As discussed in the section on the effect of information about foregone payoffs, adaptive sampling also explains the tendency for information about foregone payoffs to increase risk taking. Yechiam and Busemeyer (2006) also showed that a simple learning model, similar to the illustrative model introduced in the second section, predicts well this effect of foregone information. Finally, adaptive sampling implies that the probability of choosing a sure thing, instead of an uncertain alternative with identical expected value, should increase with experience, as observed in recent experiments (Barron & Erev, 2003, Experiment 3b; Munichor, Erev, & Lotem, 2006).

Although these results show that adaptive sampling does influence risk taking in predictable ways, one should not exaggerate its importance. First, adaptive sampling only matters when decision makers can learn only from personal experience. Whether sampling is adaptive or not does not influence risk taking when decision makers can observe foregone payoffs or obtain information about outcome distributions from others. Second, the magnitude of the effect depends on the sampling rule of the decision maker. A large effect requires limited experimentation with seemingly inferior alternatives: In the logit choice model, S has to be large or rise quickly. Experimental studies show that limited experimentation is more likely if the time horizon (N) is short or the discount rate is high (Banks, Olson, & Porter, 1997, p. 68). If the horizon (N) is long or the discount rate low, extensive experimentation can be motivated and the effect of adaptive sampling will be small. Third, the magnitude of the effect also depends on the assumptions made about how decision makers form an estimate about the value of the uncertain alternative. The bias will be smaller if the estimate of the value of the uncertain alternative changes, even if the uncertain alternative is not sampled (Yechiam & Busemeyer, 2005).

Mechanisms outside the scope of the present models also influence risk taking and will moderate the effect of adaptive sampling. Two of the most important are changing risk attitudes and the gambler's fallacy. Suppose decision makers take more risks if their status quo is below a reference point (Kahneman & Tversky, 1979). A decision maker who has chosen the uncertain alternative several times with poor results might then become more risk prone. To have a chance of obtaining a satisfactory total payoff, he or she might continue to choose the uncertain alternative rather than avoid it. This would reduce the bias against the uncertain alternative and could possibly even reverse it. A gambler's fallacy would also reduce the bias against the uncertain alternative. If decision makers believe that good outcomes follow poor outcomes, decision makers will be less likely to avoid alternatives with poor recent outcomes. The tendency for more variable outcomes to produce more random choices (the payoff variability effect) will also reduce the tendency to avoid alternatives with poor past outcomes.

Finally, the present analysis is limited to choices between a sure thing and an uncertain alternative. Suppose decision makers can choose between several uncertain alternatives. Simulations show that a decision maker may still end up favoring the alternative with the lowest variance in a repeated choice between several normally distributed uncertain alternatives with identical expected values. A formal treatment of this problem is more difficult, especially an analysis of the implications of following an optimal policy in a multiarm bandit problem.[8]

Despite these limitations, the present results illustrate how the adaptive tendency to reduce the probability of sampling alternatives with poor past outcomes is sufficient to generate apparent risk-averse behavior. Even a normatively appropriate policy for managing the trade-off between exploration and exploitation in sequential choice under uncertainty can lead to apparent risk-averse behavior.

---

[8] Sarin and Vahid (1999), however, examine the consequences of one particular learning model.

References

Banks, J., Olson, M., & Porter, D. (1997). An experimental analysis of the bandit problem. *Economic Theory, 10,* 55–77.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16,* 215–233.

Berry, D. A., & Fristedt, B. (1985). *Bandit problems.* London: Chapman & Hall/CRC.

Börgers, T., Morales, A. J., & Sarin, R. (2004). Expedient and monotone learning rules. *Econometrica, 72,* 383–405.

Burgos, A. (2002). Learning to deal with risk: What does reinforcement learning tell us about risk attitudes? *Economics Bulletin, 4,* 1–13.

Burnetas, A. N., & Katehakis, M. N. (1997). On the finite horizon one-armed bandit problem. *Stochastic Analysis and Applications, 16,* 845–859.

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General, 121,* 177–194.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara Gambling Task. *Psychological Assessment, 14,* 253–262.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning.* New York: Wiley.

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review, 112,* 951–978.

Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect. *Organization Science, 12,* 523–538.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112,* 912–931.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87,* 293–311.

Gans, N., Knox, G., & Croson, R. (2004). *Simple models of discrete choice and their performance in bandit experiments.* Manuscript in preparation, The Wharton School, University of Pennsylvania.

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society, Series B, 41,* 148–164.

Gonzáles-Vallejo, C., Reid, A. A., & Schiltz, J. (2003). Context effects: The proportional difference model and reflection of preference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 942–953.

Grosskopf, B., Erev, I., & Yechiam, E. (2006). Foregone with the wind: Indirect payoff information and its implications for choice. *International Journal of Game Theory, 36,* 285–302.

Haruvy, E., & Erev, I. (2001). Interpreting parameters in learning models. In R. Zwick & A. Rapoport (Eds.), *Advances in experimental business research* (pp. 285–300). Dordrecht, The Netherlands: Kluwer Academic.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). Risky prospects: When valued through a window of sampled experiences. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 72–91). Cambridge, U.K.: Cambridge University Press.

Hertwig, R., Weber, E. U., Barron, G., & Erev, I. (2004). Decisions from experience and the effects of rare events in risky choices. *Psychological Science, 15,* 534–539.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica, 47,* 263–291.

Lopes, L. L., & Oden, G. C. (1998). The role of aspiration level in risky choice. *Journal of Mathematical Psychology, 42,* 478–488.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis.* New York: Wiley.

March, J. G. (1996). Learning to be risk averse. *Psychological Review, 103,* 309–319.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). New York: Academic Press.

Meyer, R. J., & Shi, Y. (1995). Sequential choice under ambiguity: Intuitive solutions to the armed-bandit problem. *Management Science, 41,* 817–834.

Munichor, N., Erev, I., & Lotem, A. (2006). Risk attitudes in small timesaving decisions. *Journal of Experimental Psychology: Applied, 12,* 129–141.

Niv, Y., Joel, D., Meilijson, I., & Ruppin, E. (2002). Evolution of reinforcement learning in uncertain environments: A simple explanation for complex foraging behaviors. *Adaptive Behavior, 10,* 5–24.

Oyarzun, C., & Sarin, R. (2006). *Learning and risk aversion.* Manuscript in preparation, Texas A&M University.

Ross, S. M. (2000). *Introduction to probability models* (7th ed.). San Diego, CA: Harcourt Academic Press.

Sarin, R., & Vahid, F. (1999). Payoff assessments without probabilities: A simple dynamic model of choice. *Games and Economic Behavior, 28,* 294–309.

Shafir, S., Bechar, A., & Weber, E. U. (2003). Cognition-mediated coevolution: Context-dependent evaluations and sensitivity of pollinators to variability in nectar rewards. *Plant Systematics and Evolution, 238,* 195–209.

Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review, 11,* 430–445.

Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin and Review, 12,* 387–402.

Yechiam, E., & Busemeyer, J. R. (2006). The effect of forgone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making, 19,* 1–16.

Appendix A

The Asymptotic Probability of Choosing the Uncertain Alternative

The asymptotic density of the estimate for the illustrative model is (Denrell, 2005, p. 975)

$$\frac{(1 + e^{-Sy})f(y)}{\int_{y=-\infty}^{\infty} (1 + e^{-Sy})f(y)dy},$$ (1A)

where $f(y)$ is the normal density, with mean zero and variance $v^2 = b\sigma^2 / (2 - b)$, to which the estimate would converge if the uncertain alternative was sampled (exogenously) infinitely often. Thus,

$$E(C_U) = \frac{\int_{y=-\infty}^{\infty} \frac{1}{1 + e^{-Sy}}(1 + e^{-Sy})\frac{1}{\sqrt{2\pi v^2}}e^{-\frac{y^2}{2v^2}}dy}{\int_{y=-\infty}^{\infty} (1 + e^{-Sy})\frac{1}{\sqrt{2\pi v^2}}e^{-\frac{y^2}{2v^2}}dy}$$

$$= \frac{1}{1 + e^{\frac{S^2 v^2}{2}}} = \frac{1}{1 + e^{\frac{S^2 b}{2(2 - b)}\sigma^2}}.$$ (2A)

When $E(X) = m \neq 0$, the same reasoning gives Equation 2.

If the alternative with the highest estimate is chosen with probability $1 - q$, similar reasoning implies that the expected asymptotic probability of choosing the uncertain alternative is

$$\frac{\int_{y=-\infty}^{0} \frac{q}{q}\frac{1}{\sqrt{2\pi v^2}}e^{-\frac{y^2}{2v^2}}dy + \int_{y=0}^{\infty}\frac{1-q}{1-q}\frac{1}{\sqrt{2\pi v^2}}e^{-\frac{y^2}{2v^2}}dy}{\int_{y=-\infty}^{0}\frac{1}{q}\frac{1}{\sqrt{2\pi v^2}}e^{-\frac{y^2}{2v^2}}dy + \int_{y=0}^{\infty}\frac{1}{1-q}\frac{1}{\sqrt{2\pi v^2}}e^{-\frac{y^2}{2v^2}}dy}$$

$$= \frac{1}{\frac{1}{q}0.5 + \frac{1}{1-q}0.5} = 2q(1 - q).$$ (3A)

If the uncertain alternative pays off k / r with probability r and zero otherwise and the choice rule is to choose the alternative with the highest estimate with probability $1 - q$, then (as $t \to \infty$)

$$E(C_{U,q}) = \frac{1}{\int_{z=-\infty}^{k}\frac{1}{q}g(z)dz + \int_{z=k}^{\infty}\frac{1}{1-q}g(z)dz}$$

$$= \frac{1}{\frac{1}{q}P(Z < k) + \frac{1}{1-q}P(Z > k)},$$ (4A)

where $g(\cdot)$ is the density of random variable that the estimate of the uncertain alternative would converge to if the decision maker could sample (exogenously) this alternative infinitely often.

Appendix B

The Effect of Adaptive Sampling and Information About Foregone Payoffs

The asymptotic density of the estimate (assuming a continuous density of the outcome) is

$$h(y) = \frac{\frac{1}{p(y)}g(y)}{\int_{y=-\infty}^{\infty}\frac{1}{p(y)}g(y)dy},$$ (1B)

(Denrell, 2005, p. 975), where $g(\cdot)$ is the density of the random variable, Z, to which the estimate of the value (or utility) of the uncertain alternative would converge if it was sampled (exog-

enously) infinitely often. The expected probability that the uncertain alternative is chosen is

$$E(C_U) = \int_{y=-\infty}^{\infty} P(y)h(y)dy = \frac{\int_{y=-\infty}^{\infty} P(y)\frac{1}{P(y)}g(y)dy}{\int_{y=-\infty}^{\infty}\frac{1}{P(y)}g(y)dy}$$

$$= \frac{E_g\left(P(Y)\frac{1}{P(Y)}\right)}{E_g\left(\frac{1}{P(Y)}\right)},$$ (2B)

where the index denotes that the expectation is with respect to the density $g(\cdot)$ (for a similar derivation of the expected choice probability, see Niv et al., 2002). Thus, $E(C_U)$ equals

$$\frac{E_g(P(Y))E_g\left(\dfrac{1}{P(Y)}\right) + Cov_g\left(P(Y), \dfrac{1}{P(Y)}\right)}{E_g\left(\dfrac{1}{P(Y)}\right)}$$

$$= E_g(P(Y)) + \frac{Cov_g\left(P(Y), \dfrac{1}{P(Y)}\right)}{E_g\left(\dfrac{1}{P(Y)}\right)} . \quad (3B)$$

Because $1 / P(Y)$ is a decreasing function of Y and $P(Y)$ an increasing, $Cov_1[P(Y), 1 / P(Y)] < 0$ (Ross, 2000, p. 626). Because $E(C_{U,I}) = E_g(P(Y))$, $E(C_{U,I}) > E(C_U)$ (a similar proof holds when the distribution of the outcome of the uncertain alternative is discrete). To show that $E(C_U) < .5$, if $f(\cdot)$ is symmetric around zero, note that then $g(\cdot)$ is also symmetric around zero (a distribu-

tion is symmetric around zero if and only if its characteristic function is real valued, and the characteristic function of Z is real valued if the characteristic function of X is real valued). The assumption of $P(-y) = 1 - P(y)$ and the variable substitution $t = -y$ gives

$$E_g(P(Y)) = \int_{y=-\infty}^{\infty} P(y)g(y)dy$$

$$= \int_{t=0}^{\infty} (1 - (P(t))g(t)dt + \int_{y=0}^{\infty} P(y)g(y)dy$$

$$= \int_{0}^{\infty} g(y)dy = 0.5. \quad (4B)$$