

# When can it make sense to throw away payoffs?

**Felix Höfler**

WHU – Otto Beisheim  
School of Management  
felix.hoeffler@whu.edu

Excellence in  
Management  
Education

WHU – Working Paper  
Series in Economics

**WP 09/01**

# When can it make sense to throw away payoffs?

Felix Höffler<sup>\*†</sup>

July 21, 2009.

## Abstract

We investigate when destroying payoff can be beneficial in a learning process of a boundedly rational, risk averse agent. Destroying payoff reduces average returns but can – as a "self punishment" – accelerate learning. This smoothens payoffs and can thereby increase expected utility.

JEL-Classification: C73, D81, D83

Keywords: Bounded rationality, Learning, Sunk costs

## 1 Introduction

Consider a situation where you have to choose between two different actions,  $\mathfrak{A}$  and  $\mathfrak{B}$ . Your payoffs depend on the state of the world, which can be  $\theta_A$  (with probability  $\alpha$ ) and  $\theta_B$  (with probability  $1 - \alpha$ ). Now imagine that you can choose between the following two payoff matrices:

	Matrix 1			Matrix 2	
	$\mathfrak{A}$	$\mathfrak{B}$		$\mathfrak{A}$	$\mathfrak{B}$
$\theta_A$	$A$	$D$	$\theta_A$	$A$	$D - \delta$
$\theta_B$	$C$	$B$	$\theta_B$	$C - \delta$	$B$

Assume  $A > C \geq D > B$ . Clearly, knowing the payoff matrices, you in either matrix choose  $\mathfrak{A}$ , since in both states of the world this yields higher payoffs.<sup>1</sup> Furthermore,  $\delta > 0$ , which implies that you prefer Matrix 1. However, is this also true

---

<sup>\*</sup>WHU - Otto Beisheim School of Management, Burgplatz 2, 56179 Vallendar, Germany. ++49(0)261 6509 220, felix.hoeffler@whu.edu.

<sup>†</sup>Max Planck Institute for Research on Collective Goods, Bonn, Germany.

<sup>1</sup>That  $\mathfrak{A}$  dominates in both states is not essential. The results remain qualitatively similar if  $\mathfrak{A}$  is superior to  $\mathfrak{B}$  only in expected terms. Details are available from the author upon request.

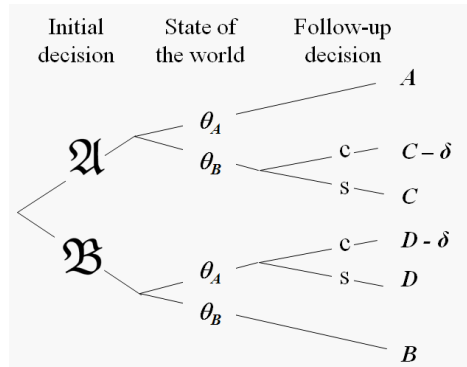


Figure 1: Sunk Cost Example

if you have first to learn the correct action? In this note, we show that in a standard payoff monotonic learning process, it might be better to choose Matrix 2.

The reason is that "throwing away payoffs" (the " $-\delta$ " in Matrix 2) can act like a self punishment which, while reducing expected payoffs, speeds up the learning process, and by so doing, reduces the variance of payoffs over time. In case of risk aversion, the latter has a value which can compensate the former.

An important application for "self punishment" is the well known "sunk cost fallacy". If you have bought expensive opera tickets, but, at the evening of the performance learn that there is an interesting football match on TV which you actually prefer to watch, you nevertheless feel obliged to go to the opera to avoid "wasting the opera tickets" (see Arkes and Ayton (1999) for an overview of experimental results).

More generally: You carry on with a project, even though you have learnt that it would be better to stop it because you invested a lot in the past. For instance, we could interpret the Matrix 2 as a representation of the following problem (see Figure 2): After starting project  $\mathfrak{A}$  ( $\mathfrak{B}$ ) you learn the state of the world and, in state  $\theta_B$  ( $\theta_A$ ) it becomes clear that it is better to stop ( $s$ ) the project  $\mathfrak{A}$  ( $\mathfrak{B}$ ), instead of continuing ( $c$ ) with it. Continuing means that you commit the sunk cost fallacy and throw away payoffs of  $\delta$ .

Evolutionary stability of non profit maximizing behavior has long been studied in games. For instance, Schaffer (1989) shows the evolutionary stability of non-profit maximization as a commitment device to deter entry. This is similar to the idea of "spite". Dufwenberg and G uth (2000) discuss evolutionary stability of behavior

that prefers outcomes with lower "material payoffs" if this leads also to a reduction of the payoff of opponents.

Another line of research deals with our key example, the sunk cost fallacy, where payoff is thrown away by caring about sunk cost. It has frequently been argued (e.g., Tröger (2002), Ellingsen and Robles (2002), or Carmichael and MacLeod (2003)) that sunk investment may influence the bargaining outcome in hold-up problems, thereby leading to more efficient outcomes. Our contribution is complementary to the existing literature by - instead of looking at strategic interaction - focusing on a simple individual decision problem.

The remainder is organized as follows. First we describe the learning process. Then, we derive conditions when choosing Matrix 2 is favorable, and then conclude.

## 2 A simple learning mechanism

We compare the expected payoffs of two identical agents. The first agent chooses Matrix 1, the second agent Matrix 2, and each plays the decision problem for a finite period of times,  $T$ . Both are boundedly rational in the sense that they use a payoff monotonic (reinforcement) learning rule when choosing between  $\mathfrak{A}$  and  $\mathfrak{B}$ . Both are risk averse and value a payoff  $\zeta$  at period  $t$  according to the following utility function:

$$u(\xi) = -e^{-\gamma\xi}, \quad (1)$$

with a constant absolute risk aversion of  $\gamma > 0$ . There is no discounting.

The agents use a mixed strategy, where  $p(t)$  denotes the probability to select  $\mathfrak{A}$  at time  $t$ . If she played  $\mathfrak{A}$  and received a payoff  $\pi$ , then she would reduce the deviation from selecting  $\mathfrak{A}$  with certainty, which is given by  $1 - p(t)$ , by the factor  $\pi/A$ . For  $\mathfrak{B}$ , vice versa. This learning process can be described by a standard replicator equation:<sup>2</sup>

$$\frac{dp}{dt} = p(t)(1 - p(t))(x - y), \quad (2)$$

where  $x$  is the expected payoff from choosing  $\mathfrak{A}$  and  $y$  is the expected payoff from choosing  $\mathfrak{B}$  (in what follows, we focus on agent 2; to receive the expression for agent

---

<sup>2</sup>Without loss of generality, we normalize  $A = 1$  for what follows.

1, one only needs to set  $\delta = 0$ ) :

$$\begin{aligned}x &= \alpha A + (1 - \alpha)(C - \delta), \\y &= \alpha(D - \delta) + (1 - \alpha)B.\end{aligned}$$

This is a simple payoff monotonic learning rule, where the probability to choose  $\mathfrak{A}$  increases in the difference in expected payoffs, which we denote by  $z = x - y$ , which we assume is positive (which, due to our assumption on the payoffs  $A, B, C, D$ , holds for  $\delta$  not too big). The equation of motion (2) then has two obvious characteristics. First,  $dp/dt > 0$  for  $0 < p < 1$  : Evolution will drive out the suboptimal choice of  $\mathfrak{B}$  in the long run, since  $\mathfrak{B}$  yields a lower payoff. Second, the payoff difference ( $x - y$ ) increases in  $\delta$  for  $\alpha > \frac{1}{2}$ . Thus, if the state  $\theta_A$  is sufficiently likely, then the agent having chosen matrix 2 (since this is like increasing  $\delta$  from zero to some positive level) learns faster than the other agent.<sup>3</sup> In what follows, we want to assume  $\alpha > \frac{1}{2}$ .

The solution to the ordinary differential equation (2) equals:

$$p(t) = \frac{e^{zt}}{e^{zt} + A_0}, \quad (3)$$

where  $A_0$  reflects the initial condition,  $p(0) = 1/(A_0 + 1)$ .

### 3 Comparison of expected utility

The expected utility of an agent from a payoff stream  $\pi(t)$  for a finite period of times is given by:

$$EU(\pi(t)) = \int_0^T \left( \frac{e^{zt}}{e^{zt} + A_0} u(x) + \frac{A_0}{e^{zt} + A_0} u(y) \right) dt, \quad (4)$$

When comparing the two agents, we are interested in the question whether the expected utility can be increased by increasing  $\delta$  (from zero to some positive level). Figure 2 shows a case where this actually happens for intermediate values of  $T$ .

---

<sup>3</sup>"Faster" means: Compare two processes  $p(t)$ ,  $p(t, \delta = 0)$  and  $p(t, \delta > 0)$ , with the same initial condition. Then it holds that  $p(t, \delta = 0) \leq p(t, \delta > 0) \forall t > 0$  for  $\alpha > \frac{1}{2}$ . The proof is by contradiction and available from the author upon request.

"Faster" does not mean that  $dp/dt$  increases in  $\delta$ . This cannot be true for all values of  $t$  since  $p$  converges to 1 for all values of  $\delta$ . Thus, if the initial "speed" is larger for larger values of  $\delta$  in the beginning, it must be lower later.

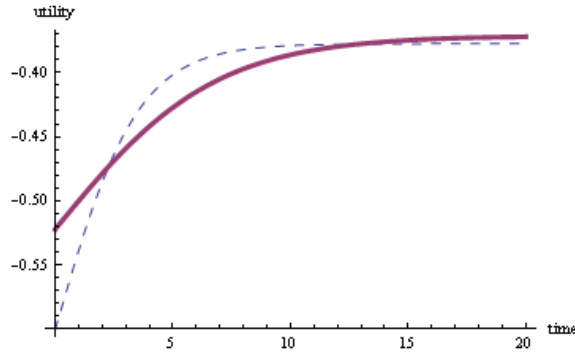


Figure 2: Per period utilities (bold = not throwing away payoff, dashed = throwing away payoff)

The parameters of the example are:  $A = 1$ ,  $C = D = 0.8$ ,  $B = -1.8$ ,  $\delta = 0.3$ , i.e. throwing away payoff yields 0.5 instead of 0.8. The probability of the state of the world  $\theta_A$  is  $\alpha = 0.96$ , the individual starts with a 50-50 probability of selecting  $\mathfrak{A}$  or  $\mathfrak{B}$ , and the level of risk aversion is  $\gamma = 1$ . The agent not throwing away payoffs must earn more in the beginning when both randomize with equal probabilities. Also in the end, when both have learned the optimal decision, the probabilities are again equal ( $p$  approaches unity for both agents), and again, the agent not throwing away payoff must earn more. However, for intermediate values, the agent who throws away payoffs earns more. Why? Because, due to the "self punishment", she has learned more quickly to avoid the wrong decision. The higher probability of choosing the "right project"  $\mathfrak{A}$  makes up for the throwing away of payoff.

However, we are interested in the cumulative payoff over time. Taking the difference in expected utilities between the two agents as a function of the "lifetime"  $T$  reveals that in this example, indeed, for intermediate values of  $T$ , the expected utility of the agent who does not throw away payoffs is lower (see Figure 3). If the lifetime is short, throwing away payoff to speed up learning has not enough time to pay off. If the lifetime is very long, both agents learnt the optimal decision, and consistently loosing  $\delta$  makes the one throwing away payoff worse off. But for intermediate values of  $T$ , agent 2 is better off.

Whether such an intermediate range exists depends on the parameter constellations. The following proposition provides sufficient conditions for the occurrence of the beneficial effect from throwing away payoff, i.e., when agent 2 is better of than

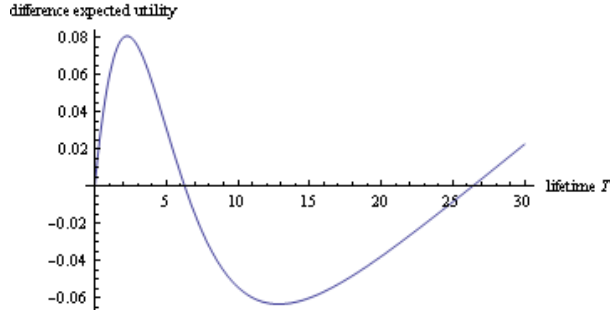


Figure 3: Difference in Expected Utility (agent 1 not throwing away – agent 2 throwing away payoff)

agent 1.

**Proposition 1** *Assume that, due to  $B$  being sufficiently small, the payoff difference  $(x - y)$  is sufficiently large and that  $\alpha > \frac{1}{2}$ . Then, for any time interval  $T$  a decision-maker with constant absolute risk aversion  $\gamma$  can increase its expected utility by throwing away payoffs of  $\delta$ , if  $\gamma$  is sufficiently large, or if  $\alpha$  is sufficiently large:*

$$\forall \gamma > \hat{\gamma} > 0 : \frac{\partial EU}{\partial \delta} > 0 \text{ for } B \rightarrow -\infty; \text{ and} \quad (5)$$

$$\exists \hat{\alpha} > \frac{1}{2} \text{ such that } \forall \alpha \in (\hat{\alpha}, 1) : \frac{\partial EU}{\partial \delta} > 0 \text{ for } B \rightarrow -\infty. \quad (6)$$

**Proof.** See Appendix. ■

The benefit from throwing away payoff is faster learning. Fast learning means that the wrong decision  $\mathfrak{B}$ , yielding the low payoff of  $B$ , is avoided more often. If  $B$  is really bad for the agents, this is an important issue, thus throwing away payoffs tends to be beneficial if  $B$  is very low. The faster learning reduces the variance of payoffs over time by avoiding  $B$ , and the value of this is higher, the more risk averse the agent is. Finally, learning will be very slow if  $\alpha$  is close to unity. The agents rarely experience the very bad outcome, and without the self punishment, have little opportunity to learn. Therefore, for high values of  $\alpha$  throwing away of payoff is also beneficial.

## 4 Discussion

In Proposition 1 we identified situations in which the competition between the two agents is won by the agent who throws away payoffs. This happens if risk averse agents have to learn to avoid a very bad outcome ( $B$ ) and where throwing away payoffs accelerates the learning process.

Instead of thinking of the model as a competition between two agents choosing different payoff matrices, one might rather think of a long run evolutionary process deciding which payoff matrix to choose, versus the short run process of learning in a very specific situation, i.e., choosing between  $\mathfrak{A}$  and  $\mathfrak{B}$ .

If we think of human decision making as "boundedly rational" in the sense of Simon (1997), then a boundedly rational agent may be well aware of his or her cognitive limitations in deciding in every situation which alternative is the best. The agent does not optimize but uses "Simple heuristics that make us smart" (Gigerenzer, Todd, and the ABC Research Group (1999)). The choice of rules or heuristics is typically long run. Within the heuristics, short term adaptation might well take place, but without questioning the heuristic as such.<sup>4</sup>

This reasoning fits well to an explanation for the sunk cost fallacy offered by psychologists, namely the *overgeneralization of rules* (Arkes and Ayton (1999)). A rule like "Don't give up too early" frequently leads to good decisions – but also supports the sunk cost fallacy. Thus, agents who follow – due to the general limitations to optimize in every situation – a rule which accelerates learning, though at the cost of sometimes destroying payoffs, might in the long run be better off than those who do not destroy payoffs.

---

<sup>4</sup>In a way, our approach is also related to the "indirect evolutionary approach" (see, e.g., Güth and Peleg (2001)) in comparing "rules" (i.e., the choice of different payoff matrices), while having evolutionary processes (learning) happening on a lower level (within a matrix).



## 5 Appendix

Proof of Proposition 3: Differentiating (4) with respect to  $\delta$  yields (a ' denotes partial derivatives with respect to  $\delta$ ):

$$\begin{aligned} \frac{\partial EU}{\partial \delta} &= \int_0^T \frac{z'te^{zt}(e^{zt} + A_0) - z'te^{zt}e^{zt}}{(e^{zt} + A_0)^2} + u'(x) \frac{e^{zt}}{e^{zt} + A_0} \\ &\quad - \frac{z'te^{zt}}{(e^{zt} + A_0)^2} u(y) + \frac{A_0}{e^{zt} + A_0} u'(y) dt \\ &= \int_0^T z' \frac{e^{zt}}{(e^{zt} + A_0)^2} t (A_0 u(x) - u(y)) + \frac{e^{zt} u'(x) + A_0 u'(y)}{e^{zt} + A_0} dt. \end{aligned} \quad (7)$$

The second term in the integral is negative (since  $u' = \frac{\partial u}{\partial \delta} < 0$ ), while the first can have either sign. We want to show that (7) is positive for  $B$  sufficiently low for every  $T > 0$ . A sufficient condition for this is that the integral converges pointwise to infinity. Note that  $B \rightarrow -\infty$  implies  $z = x - y \rightarrow +\infty$ . Therefore, the second term in the integral goes to a (negative) constant  $u'(x) = -(1 - \alpha) \gamma e^{-\gamma(C-\delta)}$ . The limit of the first term in the integral is not straightforward:

$$\begin{aligned} &tz' \lim_{B \rightarrow -\infty} \frac{e^{zt}}{(e^{zt} + A_0)^2} (A_0 u(x) - u(y)), \\ &= tz' \left[ \lim_{B \rightarrow -\infty} \frac{e^{zt} A_0 u(x)}{(e^{zt} + A_0)^2} + \lim_{B \rightarrow -\infty} \frac{-e^{zt} u(y)}{(e^{zt} + A_0)^2} \right]. \end{aligned}$$

The first term in the square brackets is negative (since  $u(x) < 0$ ), but goes to zero for  $B \rightarrow -\infty$ . The second term is positive (since  $u(y) < 0$ ), and both, nominator and denominator go to infinity for  $B \rightarrow -\infty$ . Applying l'Hospital's rule twice for the second term yields:

$$\begin{aligned} -tz' \lim_{B \rightarrow -\infty} \frac{e^{zt} u(y)}{(e^{zt} + A_0)^2} &= -tz' \lim_{B \rightarrow -\infty} \frac{e^{zt} (-\alpha e^{-\gamma(D-\delta)} - (1 - \alpha) e^{-\gamma B})}{(e^{zt} + A_0)^2} \\ &= -tz' \lim_{B \rightarrow -\infty} \frac{t\alpha e^{-\gamma C} + [(1 - \alpha)t + \gamma] e^{-\gamma B}}{2te^{zt} + 2tA_0} \\ &= tz' \lim_{B \rightarrow -\infty} \frac{\gamma [(1 - \alpha)t + \gamma]}{2t^2(1 - \alpha)} e^{-\gamma B - zt}. \end{aligned}$$

Note that:

$$-\gamma B - zt = B(t(1 - \alpha) - \gamma) - t[\alpha(A - (D - \delta)) + (1 - \alpha)(C - \delta)],$$

which will go to  $+\infty$  for  $B \rightarrow -\infty$  if:

$$t(1 - \alpha) < \gamma,$$

which holds for any  $t$  if  $\alpha$  is close to 1 and/or  $\gamma$  is sufficiently large, as stated in the proposition. ■

## References

- ARKES, H. R., AND P. AYTON (1999): “The Sunk Cost and Concorde Effects: Are Humans Less Rational Than Lower Animals?,” *Psychological Bulletin*, 125(5), 591–600.
- CARMICHAEL, L., AND B. W. MACLEOD (2003): “Caring About Sunk Cost: A Behavioral Solution to Holdup Problems with Small Stakes,” *Journal of Law, Economics and Organization*, 19(1), 106–118.
- DUFWENBERG, M., AND W. GÜTH (2000): “Why Do You Hate Me? On the Survival of Spite,” *Economics Letters*, 67, 147–152.
- ELLINGSEN, T., AND J. ROBLES (2002): “Does Evolution Solve the Hold-Up Problem?,” *Games and Economic Behavior*, 39, 28–53.
- GIGERENZER, G., P. M. TODD, AND THE ABC RESEARCH GROUP (1999): *Simple Heuristics That Make Us Smart*. Oxford University Press, New York.
- GÜTH, W., AND B. PELEG (2001): “When Will Payoff Maximization Survive? An Indirect Evolutionary Analysis,” *Journal of Evolutionary Economics*, 11, 479–499.
- SCHAFFER, M. E. (1989): “Are Profit Maximizers the Best Survivors? A Darwinian Model of Economic Natural Selection,” *Journal of Economic Behavior and Organization*, 12, 29–45.
- SIMON, H. A. (1997): *Models of Bounded Rationality*. MIT Press, Cambridge, MA.

TRÖGER, T. (2002): “Why Sunk Cost Matter for Bargaining Outcomes: An Evolutionary Approach,” *Journal of Economic Theory*, 102, 375–402.

## Supplementary Material (Not intended for publication)

Derivation of replicator dynamics: Consider a bounded rational player who repeatedly plays the game. She employs a simple learning rule which is made conditional on her own past payoff. She uses a mixed strategy, where  $p(t)$  denotes the probability to select  $\mathfrak{A}$  at time  $t$ . If she played  $\mathfrak{A}$  and received a payoff  $\pi$ , then she would reduce the deviation from selecting  $\mathfrak{A}$  with certainty, which is given by  $1 - p(t)$ , by the factor  $\pi/A$ . For  $\mathfrak{B}$ , vice versa. This payoff monotonic learning rules yields the following as an expected probability for playing  $\mathfrak{A}$  at time  $t + 1$  :

$$p(t + \tau) = p(t) + p(t)(1 - p(t)) \frac{\alpha A + (1 - \alpha) C}{A} \tau - (1 - p(t)) p(t) \frac{\alpha D + (1 - \alpha) B}{A} \tau.$$

The rate of change is then given by  $\lim_{\tau \rightarrow 0} \frac{p(t+\tau) - p(t)}{\tau} = \frac{1}{A} p(t)(1 - p(t))(x - y)$ , which is (2) for  $A = 1$ .

Proof of the claim in Footnote 3: Compare two different situations with the same initial condition  $p_0$ : In situation 1, in the underlying decision problem, the agent has  $\delta = 0$ ; in situation 2 the agent has  $\delta > 0$ . Both situations define evolutionary processes,  $p(t, \delta = 0)$  and  $p(t, \delta > 0)$ , which cannot intersect: At  $p(0)$ ,  $p(t, \delta = 0)$  will have the larger growth rate if and only if  $\alpha > \frac{1}{2}$ . Now assume that at some  $\hat{t} > 0$  the two functions intersect, which would require that  $p(\hat{t}, \delta = 0) = p(\hat{t}, \delta > 0)$  and  $\frac{\partial p(\hat{t}, \delta = 0)}{\partial t} > \frac{\partial p(\hat{t}, \delta > 0)}{\partial t}$ , which cannot be the case for  $\alpha > \frac{1}{2}$ . Thus,  $p(t, \delta = 0) \leq p(t, \delta > 0) \forall t > 0$  for  $\alpha > \frac{1}{2}$ . ■

Alternative payoff structures: In the main body of the paper we assumed  $A > D$  and  $C > B$ . Alternatively, one might assume that  $A > D$  but  $C < B$ , while still  $(x - y) > 0$  and  $u(x) - u(y) > 0$ , i.e.  $\mathfrak{A}$  is still the better choice in terms of payoffs and of utility from payoffs, but only in expected terms. Proposition 1 requires some qualifications, since we discussed only the case where  $(x - y)$  goes to infinity due to  $B \rightarrow -\infty$ . We also want to investigate the interesting cases where  $(x - y)$  becomes large due to either  $D \rightarrow -\infty$  or  $A \rightarrow +\infty$ .

**Proposition 2** *Assume that, due to  $D$  being sufficiently small, the payoff difference  $(x - y)$  is sufficiently large. Then, for any time interval  $T$ , a decision-maker with*

constant absolute risk aversion  $\gamma$ , can increase its expected utility by throwing away payoffs,  $\delta > 0$ , if  $\gamma$  is sufficiently large:

$$\forall \gamma > \hat{\gamma} > 0 : \frac{\partial EU}{\partial \delta} > 0 \text{ for } D \rightarrow -\infty.$$

**Proof.** The proof is along the lines of the proof of Proposition 1. We want to show the (7) is positive for  $D$  sufficiently small. We do so by showing that the integral converges pointwise to infinity for  $D \rightarrow -\infty$  ( $z', x', y'$  denote partial derivatives with respect to  $\delta$ ):

$$\lim_{D \rightarrow -\infty} z' \frac{e^{zt}}{(e^{zt} + A_0)^2} t (A_0 u(x) - u(y)) + \frac{e^{zt} u'(x) + A_0 u'(y)}{e^{zt} + A_0}. \quad (8)$$

The second term in (8) goes to a (negative) constant  $u'(x)$ . The first term can be written as:

$$\lim_{D \rightarrow -\infty} \frac{e^{zt} (\alpha e^{-\gamma(D-\delta)} + (1-\alpha) e^{-\gamma B})}{e^{2zt} + 2A_0 e^{zt} + A_0^2},$$

applying l'Hospital's rule twice yields:

$$\begin{aligned} & \lim_{D \rightarrow -\infty} \frac{e^{zt} (\alpha e^{-\gamma(D-\delta)} + (1-\alpha) e^{-\gamma B})}{e^{2zt} + 2A_0 e^{zt} + A_0^2} \\ &= \lim_{D \rightarrow -\infty} \frac{at (\alpha e^{-\gamma(D-\delta)} + (1-\alpha) e^{-\gamma B}) + \alpha \gamma e^{-\gamma(D-\delta)}}{2\alpha t e^{zt} + 2\alpha t A_0} \\ &= \lim_{D \rightarrow -\infty} \frac{at\gamma + \gamma^2}{2\alpha t^2} e^{-\gamma(D-d)-zt}. \end{aligned} \quad (9)$$

Note that

$$-\gamma(D-d) - zt = -D(\gamma - t\alpha) + \gamma\delta - t(\alpha A + \alpha\delta + (1-\alpha)((C-\delta) - B)),$$

implying that (9) goes to  $+\infty$  for  $D \rightarrow -\infty$  for  $\gamma > t\alpha$ . ■

A similar logic as for Proposition 1 applies. If an individual is sufficiently risk averse it will benefit from throwing away payoff if there is a very low outcome which the individual will learn to avoid faster due to the self punishment. The condition is, however, more restrictive. It is no longer true that for any level of risk aversion throwing away payoff will be beneficial if only the occurrence of the bad outcome is sufficiently unlikely. The reason is that if the bad outcome  $D$  is realized, this will

already have a strong learning impact. Reducing the payoff to  $D - \delta$  has only a relatively small additional learning effect. Only very risk averse individuals will find this learning gain worthwhile compared to the loss of expected payoff.

Finally, the payoff difference might be large not due to the danger of very low payoffs but also due to the chance of very high payoffs. Proposition 3 shows that the effect that causes destroying of payoffs to be beneficial is not symmetric in this sense.

**Proposition 3** *If the payoff difference becomes infinitely large due to  $A \rightarrow +\infty$ , the expected utility  $EU$  can not be increased by throwing away payoffs,*

$$\frac{\partial EU}{\partial \delta} < 0 \text{ for } (x - y) \rightarrow \infty \text{ due to } A \rightarrow \infty.$$

**Proof.** The proof is again along the lines of the proof of Proposition 1. For  $A \rightarrow \infty$ , the integral (7) converges pointwise to some negative constant. The second term in the integral of (7) goes to a negative constant  $u'(x) = \partial u(x) / \partial \delta = -\gamma(1 - \alpha) e^{-\gamma(C - \delta)} < 0$ . Therefore, investigate:

$$\lim_{A \rightarrow \infty} tz' \left[ \frac{e^{zt} A_0 u(x)}{(e^{zt} + A_0)^2} - \frac{e^{zt} u(y)}{(e^{zt} + A_0)^2} \right].$$

The second term in the square brackets goes to zero ( $u(y)$  does not depend on  $A$ ). The first term goes to zero as well; it can be written as:

$$\frac{A_0 u(x)}{e^{zt} + 2A_0 + \frac{A_0^2}{e^{zt}}},$$

where the numerator goes to zero while the denominator tends to infinity for  $A \rightarrow \infty$ .

■

The reason for this asymmetry stems from the two opposing effects from high payoff differences. High payoff differences imply a very fast learning. In the replicator dynamics the effect is exponential. Any further increase in learning speed must therefore have a very high marginal valuation to be worthwhile giving up some payoff  $\delta$ . If the payoff structure is such that there is a very bad payoff to be avoided, any risk averse individual attributes a very high utility to marginally decreasing the risk of this very bad outcome. These are the cases of Proposition 1 and 2. However, risk aversion is not "symmetric" in that it attributes the same marginal utility to in-

creasing the probability of getting very high payoffs. The marginal utility from very high payoffs is bounded above for any concave utility function. Therefore, throwing away payoffs cannot increase expected utility in this case.