Regret in Dynamic Decision Problems*

Daniel Krähmer[†] and Rebecca Stone[‡]

October 17, 2008

Abstract

The paper proposes a framework to extend regret theory to dynamic contexts. The key idea is to conceive of a dynamic decision problem with regret as an intra-personal game in which the agent forms conjectures about the behaviour of the various counterfactual selves that he could have been. We derive behavioural implications in situations in which payoffs are correlated across either time or contingencies. In the first case, regret might lead to excess conservatism or a tendency to make up for missed opportunities. In the second case, behaviour is shaped by the agent's self-conception. We relate our results to empirical evidence.

Keywords: Regret, Counterfactual Reasoning, Reference Dependence, Information Aversion

JEL Classification: C72, D11, D81

*We would like to thank Helmut Bester, Paul Heidhues, Graham Loomes, Nicolas Melissas, Roland Strausz, Bob Sugden and especially Erik Eyster for helpful discussions and suggestions. We also thank seminar participants in Berlin, Birmingham, Munich, Norwich, Vienna and participants of the 2005 ESRC Game Theory Workshop at UCL, the 2005 Midlands Game Theory Workshop at Loughborough and the 2005 Toulouse Conference on Economics and Psychology for useful comments. Krähmer gratefully acknowledges support by the German Science Foundation (DFG) through SFB/TR 15.

†University Bonn, Hausdorff Center for Mathematics and Institute for Economic Theory, Adenauer Allee 24-42, 53113 Bonn, Germany, ++49-(0)228-736147, kraehmer@hcm.uni–bonn.de.

‡New York University school of law, 40 Washington Square South, New York, NY 10012, rebecca.stone@nyu.edu.

1 Introduction

A large body of evidence in psychology, economics, neuroscience, and consumer research indicates that the satisfaction people derive from their experiences is influenced by their perception of what would have occurred had they made different choices[1]: we experience regret when we perceive that we would have been better off under a different choice, and we rejoice when we perceive that we are better off than we would have been. Crucially, these emotions are responses to our *ex post* evaluation of our actions relative to unchosen alternatives. Even if our action was optimal given the information available *ex ante*, we nevertheless experience regret when we later discover that some other action would have made us better off. This irrational impulse appears, however, to be coupled with considerable foresight. The evidence suggests that people anticipate these emotions and that they are prepared to trade-off purely material considerations against their desire to avoid future regret (and maximise future rejoicing). In the light of this evidence, we study the behaviour of an agent who rationally takes such regret concerns into account when he makes his decisions.

This behavioural assumption is the centre piece of the seminal regret theories of Loomes and Sugden (1982 and 1987a) and Bell (1982 and 1983).[2] These theories are essentially static in nature: there is a single decision period, and regrets are realised at the end of the period when the agent learns the state of the world and receives his material payoff. This paper goes beyond these static approaches by extending regret theory to dynamic decision problems in which multiple decisions have to be taken at different points in time.

In the analysis of such dynamic situations, a number of conceptually novel issues arise that are absent from a static context. This is because the relative performance of an early decision cannot be fully assessed until the consequences of subsequent decisions that affect the eventual payoff have become apparent. Therefore, the agent's final regrets will only be determined at the culmination of a sequence of actions. This means that later decisions are not only driven by forward looking considerations; they are also influenced by backward looking considerations that arise from the agent's desire to minimise his ultimate regrets from prior choices. In other words, in dynamic contexts the regret associated with an action taken in a particular period is generally not "sunk" at the end of that period.

This implies that an agent's regret concerns inevitably lead him to care about his behaviour at counterfactual decision nodes, that is, decision nodes that he would have reached had he acted differently in the past. Consider, for example, a career choice problem in which an

[1] We review this evidence below.

[2] A number of studies have experimentally tested the predictions of the classic regret theories. We review this literature below.

agent first chooses whether to study economics or philosophy and then chooses an occupation. Suppose that the agent has chosen economics and is deciding on a career for himself. In doing so, he needs to assess how his current career choice will influence his future regrets that will arise from the comparison between the ultimate payoff consequences of his actual economics degree and his forgone philosophy degree. Now notice that the ultimate payoff that is associated with the decision to study philosophy depends upon the career path that the agent would have subsequently chosen. Hence, the agent's anticipated regret depends upon his conjectures about this counterfactual career path. Moreover, since he cares about regret, these conjectures will influence his optimal career choice as an economist.

We assume that the rationality of the agent imposes constraints on the conjectures he can adopt about this alternative career path. That is, we suppose that he understands that had he chosen philosophy, he would form conjectures about his behavior as an economist in a similar manner in determining his occupational choice as a philosopher. Thus, we view the agent's period 2 behaviour as the outcome of a game played between the possible selves (economist and philosopher) that arise as a result of his period 1 choice. Our behavioural prediction is that his period 2 behaviour will constitute an equilibrium of this game: a *regret equilibrium*.[3]

Beyond clarifying these conceptual issues, we aim to derive specific behavioural predictions that arise from regret equilibrium. To do so, we require a plausible specification of the agent's preferences. Therefore, before we undertake the analysis of multi-period decision making, we outline a representation of the agent's preferences and demonstrate that, in a static context, it replicates a number of empirical regularities that are suggested by the psychological evidence.

This representation of preferences has two key features. First, the agent cares about the departure of chosen outcomes from a reference point, which is given by his best estimate of what he could have gotten under another choice, given his current information. Second, he dislikes losses relative to the reference point more than he likes equivalent gains. Thus, his fear of regret is a more potent force than his love of rejoicing. This means that the agent is averse to *evaluation risk*: he dislikes uncertainty about the difference between his eventual payoff and reference point.

As a result, the agent is averse to both payoff uncertainty, holding the reference point constant and variation in the reference point, holding the payoff constant. Since new information about unchosen options puts the reference point at risk, this means that he dislikes such

[3]Formally, our game is a special case of a psychological game, and regret equilibrium is a psychological Nash equilibrium (see Geanakoplos et al., 1989). In a psychological game, players care about their beliefs about what other players do and believe. In our game, the agent cares about his beliefs about what his other possible selves do.

information.[4]

In a simple two period model, the agent's regret concerns generate a rich set of behavioural predictions. These results are driven by the fact that the agent may use his second period action to strategically minimise the overall evaluation risk to which he is exposed.

In a setting in which the agent makes two similar decisions in successive periods, we identify two opposing forces on his behaviour. On the one hand, the agent behaves *conservatively.* He exhibits a tendency to stick to his past choices, even in the face of evidence that indicates that switching would be payoff maximising. This tendency arises because by sticking to his first period choice he avoids learning what he could have received had he taken a different course of action in the first place. On the other hand, he also exhibits a *reparative* tendency, which manifests itself in a tendency to try to make up for missed opportunities. He increases the correlation between his eventual payoff and counterfactual payoff by choosing an action in the second period that is similar to the action that he turned down in the first.

In a setting in which the payoff consequences of the first and second period decisions are independent, the agent's prevailing concern in the second period is to match the actions that he believes his counterfactual selves would have chosen in period 2. Suppose, for example, that he has chosen to become a lawyer rather than a consultant and now must decide how hard he will work at his chosen career. Then the belief that he would have worked hard had he become a consultant drives him to work hard as a lawyer, since the pain of doing badly in material terms as a lawyer is then compounded by its unfavourable comparison with the counterfactual in which he became a successful consultant. Conversely, if he believes that he would have been a lazy consultant he has less incentive to work hard as a lawyer. Thus, the agent's motivation is influenced by his self-conception, specifically, his beliefs about how he would behave in certain counterfactual situations.

Related Literature An alternative means of analysing regret concerns in a dynamic context is proposed by Eyster (2002). Eyster derives behavioural implications of a "taste for consistency" whereby agents have a preference for actions which cast past decisions in a more favourable light. Specifically, agents like actions which warrant having taken a past action, which means that agents' current decisions tend to be influenced by sunk costs (or past opportunity costs). On a related note, Prendergast and Stole (1996) consider a signalling story in which agents overreact to early private information and underreact to later information in order to send a signal to their future selves (or some third party) about the quality of their decisions. Finally, Yariv (2005) considers a model in which agents can actually choose their beliefs directly in order to make themselves feel better about past decisions. Current actions

[4] This is in line with psychological evidence, which we review in Section 3.2.

are then taken in the light of the resulting beliefs and so tend to reflect past judgements too closely.

All three of these models share with ours the notion that an agent's current decisions have a backward looking component. However, these are models in which an agent, in one way or another, explicitly seeks to rationalise past decisions.[5] In our model, by contrast, the agent has a more basic concern, which is simply to *avoid information* which threatens to cast a past decision in a more unfavourable light.

In a recent paper, Hayashi (2008) provides an axiomatic foundation for an alternative model of dynamic choice with anticipated regret. In contrast to our approach, he assumes that the agent considers the regrets from past decisions as sunk at later periods and focuses on dynamic inconsistency problems.

From a methodological point of view, the work that is most relevant to the current approach is Koszegi's (2004) development of the concept of "personal equilibrium". Agents are assumed to care directly about their beliefs about the actions they will choose in the continuation games following their current action, and these beliefs are, in turn, required to correctly predict their subsequent equilibrium behaviour.[6] In contrast, in our setting, the agent cares directly about the beliefs about the actions he would choose in the continuation games following counterfactual actions.[7]

We conclude with a brief overview of the empirical evidence on regret that we alluded to at the outset. First, there is a literature in experimental economics that tested a number of predictions implied by the original regret theories by Loomes and Sugden (1982 and 1987a), such as juxtaposition effects[8] and violations of monotonicity and transitivity. Early experiments found evidence for these effects and argued that they are best explained by regret theory.[9] Subsequent studies, however, indicate that the results of these early experiments are sensitive to subtle details in the format, in which decision problems are presented to subjects (see Harless,

[5]In Eyster, current actions provide agents with a rationale for actions taken in the past; in Prendergast and Stole, actions rationalise past actions indirectly, insofar as they influence the agent's beliefs about the wisdom of his past choices; finally, in Yariv's model, agents directly manipulate their perception of their past decisions.

[6]Personal equilibrium includes the notion of loss-aversion equilibrium (see Koszegi and Rabin, 2006). This captures "disappointment", the emotional response to falling short of one's expectations.

[7]Also related to our approach are a number of papers that consider agents who, like the agents in our paper, care directly about certain beliefs that they hold. For instance, Koszegi (2006) considers implications of an agent's concern with his self image and Caplin and Leahy (2001) consider preferences which have an anticipatory component, reflecting the pleasure or pain an agent derives from his beliefs about his future prospects.

[8]A juxtaposition effect occurs when preferences over lotteries depend on the joint distribution of the lotteries in the choice set and not only on each lottery's marginal distribution.

[9]See, for example, Loomes (1988a, 1988b), Loomes and Sugden (1987a, 1987b), Loomes et al. (1991, 1992), Starmer and Sugden (1989).

1992, and Starmer and Sugden 1993, 1998). Thus, the experimental validity of regret theory, as with most choice theories, does appear to be vulnerable to framing effects[10]: when decision problems are presented in formats, in which regret considerations are made particularly salient, patterns of choice that are uniquely predicted by regret theory do occur and are best explained by regret theory but these effects seem to disappear when the presentation format makes regret considerations less salient (see Starmer and Sugden, 1998). However, Bleichrodt et al (2007) employ a procedure that controls for confounding event-splitting effects to quantify the regret function, which generates findings that are largely consistent with the assumptions of regret theory.

While this experimental economics literature focusses exclusively on testing predictions of regret theory when the outcomes of unchosen lotteries are always revealed to the agent, Bell (1983, p.1165) suggested that the "hypothesis that it may matter whether a foregone alternative is resolved or not ... is the predicted phenomenon on which experimentation should be concentrated". This proposal has inspired psychologists to conduct experiments to examine the effect that varying the amount of feedback on unchosen alternatives has on subject's choices. Their findings support a version of regret theory, like the one we use here, in which an agent's regret concerns make him averse to such feedback.[11]Relatedly, Filiz and Ozbay (forthcoming) report that bidding behavior in an experimental first price auction is influenced by the nature of feedback that bidders receive. Their theory attributes the bid differences to the differences in the regret from over- or underbidding that is evoked by the various feedback controls.

There is also a psychological literature which seeks to measure the impact of the counterfactual emotions of regret and disappointment directly. In a series of experiments Mellers et al. (1999) simultaneously measure emotions and choice and find support for their hypothesis that agents make choices to maximise their expected anticipated pleasure, where the anticipated pleasure of an outcome depends among other things on the expected (or actually realised) payoff of the other gambles in the set from which agents may choose.[12]

In addition, there is a literature in consumer research which suggests both that regret influences consumers' post-choice valuations of their purchases (Inman et al., 1997 and Taylor, 1997) and that anticipation of regret influences purchase decisions (Simonson, 1992).

Finally, there is a literature in neuroscience which seeks to investigate the neurological basis of emotions such as regret and disappointment. For example, using a similar experimental

[10]For example, Starmer and Sugden (1993) find that juxtaposition effects largely disappear when an experimental design is used that controls for so-called event-splitting effects. These effects arise from the tendency of subjects to put more weight on an event if it is presented as several sub-events than if it is presented as a single event.

[11]We review this evidence in more detail in Section 3.2.

[12]See also Mellers (2000) and Mellers and McGraw (2001).

setup to Mellers et al. (1999), Camille et al. (2004) find that, in contrast to normal subjects, the emotional responses of patients with orbitofrontal cortical lesions are insensitive to the nature of the feedback they receive on unchosen gambles, which suggests that they are unable to experience regret. This suggests that the experience of regret is closely associated with a particular region of the brain (the orbitofrontal cortex).[13] A recent brain imaging study by Coricelli et al. (2005) offers further confirmation of this hypothesis.

The rest of the paper is organised as follows. In Section 2, we outline the basic model of preferences that captures the agent's regret concerns. In Section 3, we derive the essential properties of these preferences that drive the main results of the paper and consider some basic implications for behaviour when there is just a single decision period. In Section 4, we extend the model to two decision periods and introduce the concept of regret equilibrium. In Section 5, we apply the equilibrium concept to the analysis of behaviour in two different scenarios: (1) a setting in which the agent makes similar decisions in each of the two periods; and (2) a setting in which payoffs from the second decision are independent of the payoffs from the first. Section 6 concludes. Throughout, we relate our findings to the psychological evidence. All proofs are relegated to an appendix.

2 The static setup

In this section, we outline the basic model that we use to analyse decision making in a single period. There is a single agent who chooses an action from a finite set Y. An action $y \in Y$ corresponds to two random variables, a real-valued random variable, X_y, that captures the *payoff* of y, and a *signal* Z_y that captures the information revealed by y. Let $\mathcal{X}_y \subset \mathbb{R}$ and $\mathcal{Z}_y \subset \mathbb{R}^m$, for some $m > 0$, be the respective supports. It is convenient to assume that $\mathcal{X}_y$ and $\mathcal{Z}_y$ are compact. Throughout the paper, we denote random variables by capital letters and their typical realisations by small letters. Let F be the joint distribution of the family $(X_y, Z_y)_{y \in Y}$, and let f be the associated probability (density) function. We indicate conditional distributions by subscripts $\cdot|\cdot$, e.g., $f_{Z|X}(z|x)$ is the conditional probability (density) that $Z = z$ conditional on $X = x$.

The agent cares about two things. First, he cares about his material payoff. Let $\phi(x_y)$ be the instantaneous material utility from payoff realisation x_y, where ϕ is a real-valued, continuous, increasing and weakly concave function. Second, he cares about his evaluation of the performance of his action relative to the performance of forgone alternatives in his choice set. When he perceives that his action has performed relatively well, he rejoices and his utility

[13] Disappointment effects were mitigated somewhat, but were nevertheless present, suggesting that regret has a different neurological basis from that of disappointment.

improves. Otherwise, he regrets his decision and his utility falls.

The agent evaluates an action's performance relative to a reference point that aggregates the agent's perception of the performance of other actions in Y. To capture this, we divide the timeline into a decision period, $t = 1$, and an evaluation period, $t = T > 1$.

Suppose the agent has chosen y in $t = 1$. Then in the evaluation period, he receives his material payoff and observes his signal, and forms his reference point, r_y, defined as the agent's best estimate of the material utility he could have obtained from the best possible unchosen alternative in his choice set given his final information:

$$r_y = \max_{y' \in Y \backslash y} E\left[\phi\left(X_{y'}\right) | x_y, z_y\right].$$

We denote by R_y the corresponding real-valued random variable.[14,15] The relative performance of action y is then given by

$$d_y = \phi\left(x_y\right) - r_y,$$

and we define the agent's overall instantaneous period T utility from action y as

$$u\left(y\right) = \phi\left(x_y\right) + \theta\rho\left(d_y\right),$$

where ρ is a real-valued, continuous, increasing, and strictly concave function with $\rho\left(0\right) = 0$, whose domain ranges over all possible values of d_y.[16] The parameter $\theta \geq 0$ measures the intensity of the agent's regret concerns relative to his material concerns. We call an agent with $\theta > 0$ a *regretful* agent and an agent with $\theta = 0$ a *standard* agent.

$\rho\left(0\right) = 0$ reflects the fact that the agent likes a positive relative performance (he rejoices) and dislikes a negative relative performance (he experiences regret). The concavity of ρ implies *loss aversion*, that is the agent's regret about a negative performance is stronger than his rejoicing about an equivalently positive performance.[17] Furthermore, concavity generates a behavioural regularity that features prominently in a number of experiments on regret: people tend to make choices that minimise their exposure to information on forgone alternatives.[18]

Two further comments are in order. First, it should be clear that the regretful agent evaluates his decision ex post in light of his final information. This is the fundamental behavioural assumption of regret theory. The agent regrets a decision with a poor outcome even though this

[14] Formally, R_y also depends on Y. Since we will not vary Y, we suppress this dependency.

[15] An alternative specification for the maximum specification is a weighted average over the expected values of the forgone options. All of our results carry over to this case.

[16] Strict concavity means that for all $\lambda \in (0,1), x, x', x \neq x'$: $\rho\left(\lambda x + (1-\lambda) x'\right) > \lambda\rho\left(x\right) + (1-\lambda)\rho\left(x'\right)$.

[17] Loss aversion was introduced by Kahneman and Tversky (1979) and is, by now, a familiar and well-established empirical regularity (Rabin, 1998). In the context of regret, e.g. Mellers et al. (1999), Camille et al. (2004), and Inman et al. (1997) find that regret looms larger than rejoicing.

[18] Section 3.2 will make these points explicit.

decision might have been optimal at the point of decision making. Indeed, if the agent evaluated his decision based upon his ex ante information, then his behaviour would be indistinguishable from that of a standard agent.

Second, our framework builds on Bell (1983) who, like us, allows for the possibility that the outcomes of unchosen alternatives are only partially resolved. He looks at independent payoff random variables and provides natural conditions under which the reference point is given by the expectation of the unchosen random variable. In contrast, Loomes and Sugden (1982) and Bell (1982) implicitly assume that the uncertainty with respect to forgone alternatives is fully resolved after the choice was made. In this case, the reference point is determined by the realisation of the unchosen random variables, and our formulation reproduces the original formulations.

3 Decision making in the static setup

This section has two purposes. First, we sketch basic properties of the static model that are helpful in understanding the dynamic decision problems studied below. Second, we argue for the plausibility of our specification by deriving behavioural predictions that are in line with empirical evidence.

Throughout we assume that the agent correctly anticipates his emotional response at the evaluation stage and maximises his expected utility taking his regret concerns into account. That is, the agent chooses $y \in Y$ so as to maximise

$$U(y) \equiv E\left[u(y)\right] = E\left[\phi\left(X_y\right)\right] + \theta E\left[\rho\left(D_y\right)\right].$$

The crucial property that we shall explore in this paper is that the regretful agent is averse to *evaluation risk*: the ex ante variability of the difference between the payoff of his chosen action and his reference payoff, D_y. This is a direct consequence of the concavity of the regret function ρ. The following Lemma is immediate (and thus stated without proof).

Lemma 1 *Let $y, y' \in Y$ with $E\left[\phi\left(X_y\right)\right] = E\left[\phi\left(X_{y'}\right)\right]$. Let D_y second order stochastically dominate $D_{y'}$. Then the agent prefers y to y'.*

Evaluation risk aversion implies that the agent is averse both to ex ante variation in his payoff holding his reference point constant, and to variation in his reference point holding his payoff constant. We refer to the former effect as *payoff risk aversion* and the latter effect as *reference point risk aversion.* We next describe these effects formally and study their implications on overall risk attitudes.

3.1 Payoff risk aversion

In the absence of feedback on foregone options, the agent behaves like an ordinary risk-averter. This is summarised formally by the following Lemma.

Lemma 2 *Suppose that $(X_y)_{y\in Y}$ are stochastically independent. Suppose there is no exogenous feedback, i.e. Z_y is constant for all $y \in Y$. Fix two actions, $y, y' \in Y$, with $E[\phi(X_y)] = E[\phi(X_{y'})]$. Let X_y second order stochastically dominate $X_{y'}$. Then the agent prefers y to y'. This is true even if ϕ is linear.*

Lemma 2 is a straightforward consequence of Lemma 1. Since $(X_y)_{y\in Y}$ are stochastically independent and since there is no additional information in the evaluation period, the reference points for y and y' are the same. Thus, the fluctuations in the difference D result entirely from fluctuations in material payoffs $\phi(X)$. But, by assumption, these are less pronounced under action y than under action y'. The result is consistent with psychological evidence which suggests that anticipated regret may be associated with more risk averse decision making.[19]

From a methodological point of view, Lemma 2 highlights the importance of distinguishing clearly between behaviour that is motivated by regret and behaviour that is motivated by standard considerations of diminishing marginal utility. Therefore, we must isolate behavioural predictions that arise in our framework that could not be generated by expected utility theory alone.

3.2 Reference point risk aversion

Evaluation risk aversion also implies reference point risk aversion. Note that the uncertainty in the reference point is generated only by the uncertainty in the agent's ex post information (X_y, Z_y). Intuitively, a more informative signal is associated with a more uncertain reference point. Thus, reference point risk aversion corresponds to *ex post information aversion.* We now provide a notion of informativeness such that this intuition is formally true. For simplicity, we restrict attention to two actions a and b.

[19] Zeelenberg (1996) reports that "implicit or explicit in most experimental work on anticipated regret is the idea that it leads to risk aversion" (p. 149). According to Kardes (1994), "concern about regret that may follow a bad decision promotes extreme risk aversion" (p. 448). In a consumer research context, Simonson (1992) finds that experimental subjects make more risk averse choices (paying a higher price for a better known brand) when asked to anticipate regret and responsibility. Finally, Josephs et al. (1992) provide experimental evidence that low self-esteem subjects who, they argue, may be more predisposed to experience regret tend to make more risk averse choices than high self-esteem agents, who may be less predisposed to such feelings.

Definition 1 *Let X_a and X_b be identically distributed with support $\mathcal{X}$. We call Z_b more informative about X_a than Z_a about X_b if there is a probability density function $\tau(z_a|z_b)$ such that for all $x \in \mathcal{X}$:*

$$f_{Z_a|X_b}(z_y|x) = \int_{\mathcal{Z}_b} \tau(z_a|z_b) f_{Z_b|X_a}(z_b|x)\, dz_b.$$

The definition is in the spirit of the notion of a Blackwell garbling, adapted to two separate state space components. For instance, Z_b is more informative about X_a than Z_a about X_b, if the partition (σ-algebra) of Z_a induced on $\mathcal{X}$ is coarser than the partition (σ-algebra) of Z_b induced on $\mathcal{X}$. It is also easy to see that in the case with binary payoffs and correlated binary signals, the definition is satisfied when Z_b has a higher precision than Z_a. The next lemma shows that a more informative signal generates a riskier estimate of the alternative option's (possibly transformed) outcome.

Lemma 3 *Let X_a and X_b be identically distributed on $\mathcal{X}$. Let $\psi : \mathcal{X} \to \mathbb{R}$ be bounded, and let Z_b be more informative about X_a than Z_a about X_b. Then $E[\psi(X_a)|Z_b]$ is a mean preserving spread of $E[\psi(X_b)|Z_a]$.*

To understand the intuition, consider the extreme case in which Z_a is not informative at all, and Z_b is fully informative. Then $E[\psi(X_b)|Z_a]$ is deterministic, and $E[\psi(X_a)|Z_b]$ is $\psi(X_a)$.

We can now derive the behavioural implication that the agent prefers the less informative action. Suppose, X_a and X_b are independent. Then the agent's observation of his actual choice does not affect his prediction of the alternative option's outcome: $R_y = E[\rho(X_{y'})|Z_y]$. Thus, Lemma 3 implies that R_b is more risky than R_a, and, in turn, since the payoff from a and b is the same, that the relative perfomance of b, D_b, is more risky than that of a, D_a:

Lemma 4 *Let X_a and X_b be i.i.d.*[20] *Let X_a and Z_a as well as X_b and Z_b be independent, and let Z_b be more informative about X_a than Z_a about X_b. Then D_a second order stochastically dominates D_b. Thus the agent prefers a to b.*

The lemma says that, all else equal, the agent chooses actions that minimise his exposure to ex-post information. Moreover, Lemma 4 suggests a trade-off between payoff risk and reference point risk: the agent's preference for a should be maintained even if the riskiness of X_a is slightly increased. While a full elaboration of this point is beyond the scope of the paper, the following example illustrates this feature of our model.

Example 1 Let $Y \in \{a, b\}$. Suppose that $\phi(x) = x$ and that X_a and X_b are both symmetrically distributed with $E[X_a] = E[X_b]$. Suppose that the agent learns the realisation of

[20] independently and identically distributed.

X_a regardless of his choice while he learns the realisation of X_b only if he chooses b. Then the agent (weakly) prefers a to b.

In the example, even if X_b second order stochastically dominates X_a, the regretful agent chooses a. Moreover, it follows from Lemma 2 that the agent would choose the less risky option if no feedback was provided on unchosen options, and it is easy to see that the same is true if he always receives feedback on all options.

We close this section by arguing that these predictions are in line with empirical evidence. The relation between risk preferences and feedback has been tested experimentally in a number of psychological studies.[21] In support of our model, most experimental studies find that subjects' behaviour is influenced by the amount of feedback they expect to obtain on unchosen options in a manner that is consistent with our model.[22] For example, van de Ven and Zeelenberg (2006) conduct trading experiments in which subjects are more willing to trade a lottery ticket for another lottery ticket when they know the number of both tickets than when they know only the number of their own ticket. This is in line with the predictions of Lemma 4 andExample 1: when subjects know only the number of their own ticket they can avoid feedback about the alternative ticket by keeping their original ticket, while when they know the number of both tickets they cannot avoid this information by not trading. Similarly, subjects are more willing to trade when they know the number of neither ticket than when they only know the number of their own ticket. When subjects know the number of neither ticket, trading (or not trading) exposes them to feedback only about the ticket that ends up in their possession, while in the latter scenario, trading ensures that they are exposed to knowledge of the payoffs associated with both tickets instead of just one of them.

In the same vein, Larrick and Boles (1994), in a negotiation experiment, and Ritov (1996), in a gambling experiment, find that subjects become more willing to take risks if feedback on the risky alternative is provided. Similar findings are reported by Zeelenberg et al. (1996) in a gambling task, and by Zeelenberg and Beattie (1997) in an ultimatum game experiment.[23] For a more detailed review of the experimental evidence see Zeelenberg (1999).

Finally, in a field study, Zeelenberg and Pieters (2004) find that anticipated regret is a predictor of participation among players of the Dutch Postcode Lottery, but not among players

[21] On this point, see also Bell (1983) who considers the choice between a risky gamble and a sure thing and hypothesises that the sure thing becomes more attractive if the uncertainty about the risky gamble is not resolved (p. 1160).

[22] An exception is Kelsey and Schepanski (1991) who, in the context of a taxpayer reporting decision experiment, find no evidence that the feedback that subjects expect to obtain has an impact on behaviour.

[23] In a gambling experiment, Josephs et al (1992) induce risk averse choices of low self-esteem subjects by providing feedback on forgone options, but in their setup, the risk averse choice minimises regret.

of the State Lottery. The difference is attributed to the fact that winners of the Postcode Lottery are determined by postcode, and thus, potential participants learn whether or not they would have won regardless of whether they choose to play from their neighbours, while such forced feedback about the outcome is not a feature of the State Lottery.

In sum, this evidence strongly suggests that the fear of regret drives agents to try to minimise their exposure to information about foregone alternatives. The fact that our model captures this significant regularity through the concavity of ρ provides an important justification of our specification of the utility function.[24,25]

4 Dynamic decision making and regret equilibrium

We now turn to the analysis of dynamic decision problems in which the ultimate payoff the agent receives is the consequence of multiple decisions. In analysing such decision problems, two key issues arise. First, an agent's subsequent decision may influence the regret he experiences about early actions. Second, his conjectures about his behaviour at counterfactual decision nodes play a crucial role in determining his eventual regrets, and, hence, his actions.

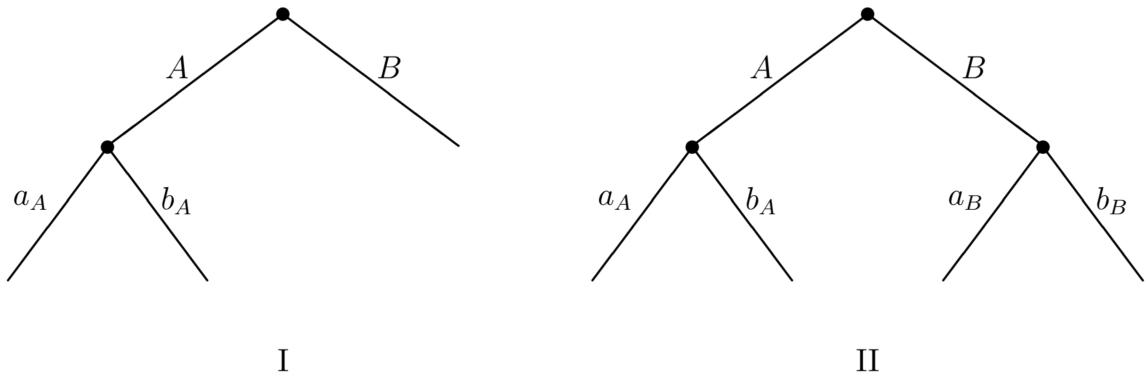

Figure 1: Two decision trees

To illustrate the first issue, consider the stylised decision problem that is represented by decision tree I in Figure 1. An agent chooses between A and B in first period. If he chooses A in the first period, he must chooses between a_A and b_A in the second period, but he faces

[24] Humphrey (2004), in a more restrictive setting, also generates this regularity. In contrast to us, he directly posits that an agent's regrets are more intense when feedback on unchosen alternatives is provided than when not.

[25] According to Mellers et al. (1999), ρ might be S-shaped, that is, concave over the rejoicing domain but convex over the loss domain. Yet most of our results can be expected to go through even for an S-shaped ρ if the convexity is not very severe.

no further decision if he chooses B. Suppose that outcomes are uncertain and all payoffs are realised in period 3. This means that the final payoff that will result from a period 1 choice of A will depend on whether the agent subsequently chooses a_A and b_A. Thus, the extent to which the agent will end up regretting a choice of A over B in period 1 will also depend on his period 2 decision. His eventual regrets are not "sunk" after period 1 and so his second period decision will have a backward looking component arising from his desire to minimise his ultimate regrets about his first period choice.

Thus, we assume that regrets about a particular decision are realised at an ultimate evaluation stage, after which no further events which may affect the apparent ex post wisdom of the decision will occur. This is in line with the central assumption of classic regret theory according to which agents evaluate their past decisions in the light of any information that they have at their disposal *once the payoffs from their choices and any available information about unchosen alternatives has been realised.*

Now consider the decision problem that is represented by tree II. This has the new feature that the final payoff of *both* initial choices depends upon a continuation action. Thus, an agent's final regrets about having chosen A for example, will depend not only on his continuation action (a_A or b_A) but also upon his conjecture about the action he would have chosen in the counterfactual scenario in which he chose B instead (a_B or b_B).

Thus, to complete the specification of dynamic decision problems with regret, we need to specify how the agent forms his conjectures about his behaviour at counterfactual decision nodes. In the next section, we develop a framework that resolves this problem in the spirit of rational expectations. For simplicity, we restrict the analysis to two-stage problems. The extension to an arbitrary finite number of stages is conceptually similar.

4.1 The dynamic setup

We extend the timeline to two decision periods, $t = 1, 2$, and one evaluation period $t = T > 2$. In period 1, the agent chooses an action y^1 from the finite set Y^1, and subsequently receives a signal $Z^1_{y^1}$ with compact support $\mathcal{Z}^1_{y^1} \subset \mathbb{R}^m, m > 0$. In period 2, contingent on y^1 and z^1, the agent chooses action y^2 in the finite set $Y^2_{y^1}$ and receives a signal $Z^2_{(y^1,y^2)}(z^1)$ with compact support $\mathcal{Z}^2_{(y^1,y^2)}(z^1) \subset \mathbb{R}^m, m > 0$. Each path (y^1, z^1, y^2) delivers the payoff random variable $X_{(y^1,y^2)}(z^1)$ with compact support $\mathcal{X}_{(y^1,y^2)}(z^1)$. This means that the second signal Z^2 is payoff irrelevant and pure ex post information. The set of feasible action paths is denoted by $Y = \{(y^1, y^2) \,|y^1 \in Y^1, y^2 \in Y^2_{y^1}(z^1), z^1 \in \mathcal{Z}^1_{y^1}\}$. Let F be the joint probability distribution of the family $\left(Z^1_{y^1}, X_y(z^1), Z^2_y(z^1)\right)_{y \in Y}$ and let f be the respective probability (density) function.

In the evaluation period, the agent receives his material payoff and evaluates his decision based on the information available in $t = T$.

This specification contains the simplifying assumption that *all* regrets are realised in the evaluation period and rules out the existence of interim regrets at stage 2. As argued above, the important point of the dynamic setup is that some regrets from the period 1 decision are realised in the evaluation period. Our assumption focuses attention on this point.

4.2 Preferences

When an agent makes decisions in two different periods, he might regret (or rejoice in) each decision. Thus, for example, an agent may regret his initial choice of degree while also rejoicing in his subsequent career choice, believing that his chosen career was the best option available to him given his (imperfect) choice of degree. To capture this, we disentangle the agent's overall regret from action (y^1, y^2) into two components, one corresponding to each decision.

Suppose that the agent has chosen $y = (y^1, y^2)$ and observed the signal realizations z^1 and z^2 in period T. When the agent evaluates the performance of his second decision, y^2, he takes his period 1 action, y^1, as given. Analogous to the single period context, he then compares the performance of y^2 with the performance of the alternative actions $\widehat{y}^2 \in Y^2_{y^1}$ that were available to him at the period 2 decision node induced by action y^1. Hence, the agent's instantaneous regret from his second decision y^2 is given by

$$v_2\left(y^1, y^2; z^1, z^2\right) = \rho\left(\phi\left(x_{(y^1,y^2)}\right) - \max_{\widehat{y}^2 \in Y^2_{y^1} \setminus y^2} E\left[\phi\left(X_{(y^1,\widehat{y}^2)}\right) \middle| x_{(y^1,y^2)}, z^1, z^2\right]\right).$$

As we noted previously, the key feature of the dynamic setup is that the outcome of a forgone action $\widehat{y}^1 \neq y^1$ depends on which period 2 action the agent would choose in the counterfactual scenario that arises in the event that he has chosen $\widehat{y}^1$ in period 1. Thus, to compare the performance of his actual choice y^1 to that of a counterfactual choice $\widehat{y}^1$, the agent needs to form a conjecture about his counterfactual continuation action at the decision node induced by $\widehat{y}^1$. We identify action $\widehat{y}^1$ with a *counterfactual self* $CS_{\widehat{y}^1}$: the incarnation of the agent that would have controlled the period 2 action in the event that $\widehat{y}^1$ had been chosen in period 1.

Given that the agent chose y^1, the agent's conjecture is expressed by a *belief,* $\mu^2_{(\widehat{y}^1,z^1)}$, that specifies for each counterfactual self, $CS_{\widehat{y}^1}, \widehat{y}^1 \neq y^1$, and all possible realisations z^1 of $Z^1_{\widehat{y}^1}$ a probability distribution over $Y^2_{\widehat{y}^1}$. That is, the agent conjectures that in the event that counterfactual self $CS_{\widehat{y}^1}$ observes signal z^1, it chooses action $\widetilde{y}^2 \in Y^2_{\widehat{y}^1}$ with probability $\mu^2_{(\widehat{y}^1,z^1)}(\widetilde{y}^2)$. We denote the profile of all (contingent) beliefs about counterfactual strategies by $\mu^2_{-y^1} = \left(\mu^2_{(\widehat{y}^1,z^1)}\right)_{\widehat{y}^1 \in Y^1 \setminus y_1}$.

The relative performance of action y^1 with respect to a forgone alternative $\widehat{y}^1$ is determined as follows. The agent first computes the conditional probability that the counterfactual self

$CS_{\widehat{y}^1}$ is of type z_1 conditional on his ex-post information. Using this he derives the probability $\mu^2_{(\widehat{y}^1,z^1)}(y^2)$ with which $CS_{\widehat{y}^1}$ has chosen action y^2 and then he uses these probabilities to calculate the expected value of $\phi\left(X_{(\widehat{y}^1,\widetilde{y}^2)}\right)$ conditional on his information.

The agent's first period reference point, r_1, is then given by the performance of the best possible unchosen alternative in his first period choice set. Formally, given a conjecture $\mu^2_{-y^1}$ and given actions (y^1, y^2), the instantaneous regret that the agent experiences about his first decision y^2 in the evaluation period is given by

$$v_1\left(y^1, y^2; \mu^2_{-y^1}, z^1, z^2\right) = \rho\left(\phi\left(x_{(y^1,y^2)}\right) - \max_{\widehat{y}^1 \in Y^1 \setminus y^1} E_{Z^1(\widehat{y}^1)}\left[E_{\widetilde{y}^2}\left[\phi\left(X_{(\widehat{y}^1,\widetilde{y}^2)}\right) \middle| x_{(y^1,y^2)}, z^1, z^2\right] \middle| x_{(y^1,y^2)}, z^1, z^2\right]\right),$$

where the first expectation is taken with respect to the conditional distribution of the signal $Z^1(\widehat{y}^1)$, and the second expectation is taken with respect to the agent's conjecture $\mu^2_{(\widehat{y}^1,z^1(\widehat{y}^1))}$ and the conditional distribution of $X_{(\widehat{y}^1,\widetilde{y}^2)}$ conditional on $x_{(y^1,y^2)}, z^1, z^2$.

Finally, we assume that the total instantaneous regret that the agent experiences in the evaluation period as a result of actions (y^1, y^2) is a weighted sum of his first and second period regrets:

$$v\left(y^1, y^2; \mu^2_{-y^1}, z^1, z^2\right) = \theta_1 v_1\left(y^1, y^2; \mu^2_{-y^1}, z^1, z^2\right) + \theta_2 v_2\left(y^1, y^2; , z^1, z^2\right).$$

The parameters $\theta_t \geq 0$ measure the extent to which the agent cares about his first and second period decision in regret terms.

Notice that our formulation embodies the psychological assumption that if the agent believes that his counterfactual self chooses his second period action randomly, either because it plays a mixed strategy, or because it makes its choices contingent on the realisation of a private signal, he incoporates this uncertainty directly into the reference point. This is in line with our assumptions about the agent's informational preferences in the static setup.

4.3 Behaviour

We now argue that in this dynamic setup the agent's behaviour is naturally described as the outcome of a non-cooperative game played between the possible incarnations of the agent (actual and counterfactual) that arise at each of his period 2 decision nodes.[26]

Consider first the agent's period 2 behaviour and suppose that his period 1 choice was y^1 and he has observed the signal z^1. We make two assumptions. First, we assume that,

[26] Throughout we assume that the agent cannot commit to a period 2 action contingent on his first period choice. If this was the case, we would be back in the static case: the agent would simply select a contingent strategy in period 1 and compare it to other contingent strategies that he could have chosen in period 1. In the absence of a commitment device, we think this is a natural assumption.

given a conjecture $\mu^2_{-y^1}$, the agent maximises his expected utility taking into account his regret concerns. That is, he chooses $y^2 \in Y^2_{y^1}$ so as to maximise

$$U_{y^1}\left(y^2;\mu^2_{-y^1},z^1\right) \equiv E\left[\phi\left(X_{(y^1,y^2)}\right) + v\left(y^1,y^2;\mu^2_{-y^1},z^1,Z^2_{(y^1,y^2)}\left(z^1\right)\right)|z^1\right].$$

Implicit in this formulation is the assumption that the agent is bound to his conjecture over time. That is, in the evaluation period, the agent's conjecture about his counterfactual self's strategy is the same as his period 2 conjecture.

Second, we assume that in forming his conjectures he rationally understands that had he chosen $\widehat{y}^1$, he would, in order to determine his behaviour in this counterfactual situation, reason in a similar way about his current behaviour. This suggests that the agent's period 2 behaviour is the outcome of an intra-personal game G played between the agent's counterfactual selves where counterfactual self CS_{y^1}'s objective function is given by U_{y^1}.

Because each self cares directly about his conjectures about his counterfactual selves' actions, yet not about their physical consequences, the game G is a psychological game (see Geanakoplos et al., 1989) and we are naturally led to assume that the agent's period 2 behaviour constitutes a psychological Nash equilibrium. More precisely, let $s^2_{y^1} \in Y^2_{y^1}$ be a (possibly mixed) strategy for self CS_{y^1}, and denote by $s^2 = \left(s^2_{y^1}\right)_{y^1\in Y^1}$ a strategy profile and by $\mu^2 = \left(\mu^2_{y^1}\right)_{y^1\in Y^1}$ a profile of conjectures. We define a regret equilibrium as a psychological Nash equilibrium of G.

Definition 2 *A pair (μ^2, s^2) is a (period 2) regret equilibrium if*
(i) $\mu^2 = s^2$,
(ii) for each $y^1 \in Y^1, z^1 \in \mathcal{Z}^1_{y^1}$ and $\widehat{s}^2 \in Y^2_{y^1}$: $\quad U_{y^1}\left(s^2;\mu^2_{-y^1},z^1\right) \geq U_{y^1}\left(\widehat{s}^2;\mu^2_{-y^1},z^1\right).$
The set of all strategies s^{2^*} such that there is a (period 2) regret equilibrium $\left(s^{2^*}, s^{2^*}\right)$ is denoted by S^{2^*}.

Our key behavioural assumption is that the agent's period 2 behaviour constitutes a (period 2) regret equilibrium: each self plays a best response to his conjecture about the other self's behaviour and these conjectures correctly predict the other selves' behaviour.[27] Note that (i) implies that an equilibrium is fully specified by s^2 alone. Thus, henceforth we ommit any explicit mention of μ^2.

Next consider period 1 behaviour. We assume that in period 1 the agent anticipates that he will play a regret-equilibrium in period 2 and chooses action y^1 so as to maximise

$$\begin{aligned} &U\left(y^1\right) \equiv E\left[\phi\left(X_{\left(y^1,s^2_{y^1}\right)}\right) + v\left(y^1,s^2_{y^1};s^2_{-y^1},Z^1_{y^1},Z^2_{(y^1,y^2)}\right)\right] \\ &\text{s.t.} \quad s^2 \in S^{2^*}. \end{aligned}$$

[27] Since the two selves operate in different possible worlds, the actions of one do not affect the other's physical payoffs, but each self cares directly about their beliefs about the other self's behaviour.

We call the agent's overall behaviour (y^1, s^2) with an optimal period 1 action y^1 and $s^2 \in S^{2^*}$ a *regret equilibrium*.

The assumption that the agent's behaviour constitutes a regret equilibrium means, in effect, that the agent is not free to choose his counterfactual conjectures and must instead form them rationally based on his understanding of the nature of the choices that his alternative possible selves must make. This captures the spirit of the original regret theory: the agent's propensity toward ex post evaluation is combined with rational foresight.

Indeed, the agent thinks of his counterfactual behaviour just as he would think about the behaviour of another agent. Consider a simple two agent entry game in which a regretful manager of a firm must decide whether or not to enter a market in which the incumbent will respond by either accomodating or fighting. Suppose the incumbent's payoffs are such that he will certainly fight any entering firm. Then, in line with backwards induction, it seems natural to suppose that if the manager chooses to stay out he compares his actual payoff to the payoff he would have gotten had he entered and been fought (rather than accomodated). Analogously, in our two-period single agent decision problem, the agent bases his ex post comparisons on rational beliefs about what he would have done had he chosen otherwise.[28]

5 Implications of regret equilibrium

We now derive behavioural implications of our regret equilibrium concept. Notice that the strategic interaction between the agent's possible selves arises through the agent's regret concerns with respect to his period 1 choices. We therefore set $\theta_2 = 0$ from now on in order to isolate these strategic effects. For $\theta_2 > 0$, all our results go through so long as period 1 regret concerns are sufficiently large relative to period 2 regret concerns. Also, we are mainly interested in the agent's behaviour at stage 2. We do not explicitly look at his (rather straightforward) first period behaviour.

We consider two kinds of situation. First, we consider repeated actions whose payoffs are correlated across time. Second, we look at situations in which the payoffs of period 2 actions are correlated across all possible decision nodes.

[28] It seems interesting, however, to speculate about alternative assumptions. For instance, the agent might naively suppose that his counterfactual selves maximise their expected material utility. More radically, the agent might choose his (unverifiable) beliefs in a self-serving manner so as to minimise his regrets. Such an agent's beliefs would be irrational in a truth-seeking sense, but not necessarily from the point of view of maximising overall utility.

5.1 Correlation across time

Suppose the agent can choose repeatedly from actions with correlated payoffs. We identify two opposing forces that drive the agent's incentives in period 2. On the one hand, the agent's aversion to information on forgone payoffs will drive him to stick to past actions, that is, to behave in an *excessively conservative* manner. On the other hand, his desire to correlate his actual outcome with his counterfactual reference outcome will drive him to try to make up for opportunities that he turned down in early periods by switching his course of action, a tendency that might psychologically be interpreted as a *reparative tendency*.

In the next two subsections, we identify conditions under which each effect is likely to occur and summarise the relevant evidence. We consider the following setup.[29] In each of two periods, $t = 1, 2$, the agent has to make a binary choice $y^t \in Y^t = \{a, b\}$. The payoffs of an action are perfectly correlated across time. That is, action a yields payoff X_a in period 1 and payoff λX_a in period 2 for $\lambda \geq 0$. Likewise, action b yields payoff X_b in period 1 and payoff λX_b in period 2. X_a and X_b are stochastically independent. After the first period, the agent receives a signal that is informative about X_a and X_b.

The parameter λ reflects the weight on the second period payoff relative to the first period payoff. When analysing the agent's behaviour, we will be interested in the effects of varying λ.

5.1.1 Excess conservatism

We now consider how a regretful agent responds to news that questions the wisdom of his first period choice. Suppose that after having made his first choice, the agent gets a binary signal Z_y^1 with values in $\left\{z_y', z_y''\right\}$. The realization z_y' indicates that action y is worse than action $\hat{y} \neq y$ in payoff terms, and the realization z_y'' indicates that action y is better than action $\hat{y} \neq y$ in payoff terms.

We assume that Z_y^1 is not fully revealing about X_y, that is, $F_{X_y|Z_y^1}$ is non–degenerate. Moreover, for $\hat{y} \neq y$, the signal Z_y^1 is uninformative about $X_{\hat{y}}$, that is, $F_{X_{\hat{y}}|Z_y^1} = F_{X_{\hat{y}}}$. To capture the idea that z_y' is bad and that z_y'' good news, we assume that for $\hat{y} \neq y$, $F_{X_y|Z_y^1}\left(\cdot|z_y'\right)$ is first order stochastically dominated by $F_{X_{\hat{y}}}(\cdot)$, and that $F_{X_y|Z_y^1}\left(\cdot|z_y''\right)$ first order stochastically dominates $F_{X_{\hat{y}}}(\cdot)$.

After his second period choice, the agent only learns the payoff(s) of the action(s) he has chosen. This means that the agent can affect the amount of information he ultimately receives through his second period choice. We are interested in the regretful agent's behavior in period 2 after having received "bad" news z'. We say that the agent's behavior is *conservative* if he does not change his action in period 2 in response to bad news.

[29] The setup can be easily generalised without changing our main results.

Before we analyze the regretful agent's behavior, it is useful to discuss the benchmark behavior of a standard agent. Of course, a rational risk-neutral agent switches action in response to bad news and sticks with his choice in response to good news. In contrast to a risk-neutral agent, a risk-averse agent seeks a compromise between his expected payoff and the aggregate payoff risk to which he is exposed. Switching to a different action in period 2 reduces the aggregate risk, because it diversifies the agent's portfolio. Hence, the agent sticks to his first period choice only when it has an advantage in payoff terms. But this is never the case, when he receives bad news. Thus, a standard agent always switches in response to bad news. This stands in stark contrast to the behavior of the regretful agent as summarized in the next proposition.

Proposition 1 *There exist strictly positive $\overline{\lambda}$ and $\overline{\theta}_1$ such that for all $\lambda \leq \overline{\lambda}$ and $\theta_1 \geq \overline{\theta}_1$ there is a regret-equilibrium in which the agent's period 2 behaviour is conservative.*

The key point of Proposition 1 is that in period 2 the regretful agent may stick to his first period choice even though he receives unfavourable news and switching his action would both raise his material payoff and reduce the aggregate payoff risk to which he is exposed. In this sense, the agent acts as if he ignores evidence against his previous choice and exhibits *excess conservatism*.

To understand what drives this result, notice that for a very small period 2 payoff weight, λ, the period 2 decision has negligible payoff consequences while its information content is undiminished. Switching course of action reveals what the agent would have obtained had he acted differently in period 1 and so exposes the agent to evaluation risk whereas sticking with his original choice adds nothing to his final information set and so eliminates all evaluation risk.

This conservative tendency provides an explanation for brand loyalty: by continuing to consume a particular brand, consumers avoid the risk of learning that they would have been better off had they consumed a different brand. In effect, the fear of regret generates a psychological switching cost.

At a general level, as Camerer and Weber (1999) demonstrate, there is evidence that people and organisations sometimes exhibit a tendency to maintain or increase a commitment to a project even when the cost-benefit calculus suggests that they would be wise to abandon it. Our theory suggests that the fear of regret may contribute to such escalation phenomena. Indeed, Tykocinski and Pittman (1998) use fear of regret to explain the results of experiments that suggest that people who have forgone attractive opportunities tend to decline similar (but substantially less attractive) current opportunities.[30]

[30] "Declining the subsequent action opportunity may appear to be illogical because this opportunity is still attractive in an absolute sense, but it may also be psychologically rewarding because it mitigates the unpleasant experience of regret produced by dwelling on a perceived loss" (p. 608).

These experiments do not exactly replicate the features of our setup since subjects perfectly learn the value of the option initially forgone before making their subsequent choice. However, in explaining their results, Tycocinski and Pittman implicitly assume that agents are able to forget such information about unchosen options that they may have once known unless they are explicitly reminded of it. Thus, agents stick to an original course of action in order to avoid *reminding* themselves of the value of the forgone alternative. By contrast, in our theory agents stick to an original course of action in order to avoid *learning* about a forgone option.

Interestingly, Tycocinski and Pittman also find that this tendency to avoid the foregone course of action is attenuated when the agent will be reminded of it even if he does not subsequently choose it. Similarly, in our setup, if we make the alternative assumption that the regretful agent will receive full feedback on his available actions, irrespective of the choices that he makes, then he no longer displays conservatism.

5.1.2 Reparative action

The previous result rests critically on the assumption that the second decision's payoff consequences are rather unimportant. In this case, the regretful agent's desire to avoid information about foregone alternatives drives him to stick to his first period action in the second period. In this subsection, we identify a countervailing force on his behaviour, which drives him to try to make up for missed opportunities by choosing a second period action that is similar to the action he turned down in the first period. This force predominates when the two decisions are of equal importance in payoff terms.

In contrast to the previous section we now assume, for simplicity, that the agent does *not* get any feedback after period 1. The behavioural tendencies we identify will, however, also be generated by a more complex model in which the agent does receive such interim feedback.

It turns out that the relative riskiness of the options a and b is a critical determinant of the agent's incentives. Thus, in order to characterise the forces at work, we shall want to vary the degree of risk of a relative to b. A convenient way to do so, is to assume that X_a is symmetrically distributed with mean $E[X_a]$, and that X_b is given by the random variable $\gamma \widehat{X}_a$, where $\widehat{X}_a$ and X_a are i.i.d. The parameter γ reflects the relative riskiness of a and b. If $\gamma \geq 1$, then X_b second order stochastically dominates X_a, and the reverse if $\gamma < 1$.

In addition, our point becomes most transparent if $\lambda = 1$. So, for the rest of this section, we focus exclusively on this case. We also assume that ϕ is strictly concave.[31]

Our aim is to identify conditions under which the behaviour of the regretful agent departs

[31] The case with linear ϕ is a boundary case, which is qualitatively similar, but requires slight changes in the proofs. To save space, we ommit the argument.

from that of the standard agent. We thus first characterise the standard agent's behaviour in the current setup.

Lemma 5 *Consider the standard agent. There are $\overline{\gamma}, \overline{\overline{\gamma}}$ with $\overline{\gamma} < 1 < \overline{\overline{\gamma}}$ such that the agent's choice is given by*

$$(y^1, y^2) = \begin{cases} (b,b) & \text{if} \quad \gamma < \overline{\gamma} \\ (a,b) \ \text{ or } \ (b,a) & \text{if} \quad \gamma \in [\overline{\gamma}, \overline{\overline{\gamma}}) \\ (a,a) & \text{if} \quad \gamma \geq \overline{\overline{\gamma}}. \end{cases}$$

The lemma says that the standard agent switches course of action if and only if the random variables are similarly risky, that is, if $\gamma \in [\overline{\gamma}, \overline{\overline{\gamma}})$. (Due to the symmetry of our setup, he is indifferent between (a,b) and (b,a)). In this case, by switching course of action the agent diversifies the aggregate payoff risk to which he is exposed. As the relative riskiness of one alternative becomes extreme ($\gamma < \overline{\gamma}$ or $\gamma \geq \overline{\overline{\gamma}}$), then choosing the less risky alternative in both periods minimises the agent's overall payoff risk.

We now turn to the regretful agent. The following lemma describes the best responses given the counterfactual self switches action in period 2.

Lemma 6 *Consider the regretful agent's actual period 2 decision.*
(i) Suppose the agent chose a in period 1 and the counterfactual self switches action in period 2. Then there is $\widehat{\widehat{\gamma}}$ with $\widehat{\widehat{\gamma}} > \overline{\overline{\gamma}}$ such that the agent switches course if and only if $\gamma < \widehat{\widehat{\gamma}}$.
(ii) Suppose the agent chose b in period 1 and the counterfactual self switches action in period 2. Then there is $\widehat{\gamma}$ with $\widehat{\gamma} < \overline{\gamma}$ such that the agent switches course if and only if $\gamma > \widehat{\gamma}$.

The following proposition describes the equilibria that result from these best responses. (It is an immediate consequence of the previous lemma and thus stated without proof.)

Proposition 2 *For all $\gamma \in \left(\widehat{\gamma}, \widehat{\widehat{\gamma}}\right)$ there is a regret equilibrium in which the regretful agent switches action in period 2.*

The proposition implies that the regretful agent switches for a greater range of parameters than the standard agent. To grasp the intuition for this result, notice first that the regretful agent's payoff risk aversion means that, like the standard agent, he has an incentive to diversify his portfolio. In addition, the agent's behaviour is driven by evaluation risk aversion. If the weight on the action's payoff is the same across periods and the counterfactual strategy prescribes switching in period 2, then if the agent switches in period 2, he obtains in period 1 what he counterfactually would have obtained in period 1. Hence, by switching he eliminates all evaluation risk.

One interpretation of the regretful agent's behaviour is suggested by psychological "dissonance" theories that view the mere act of choice as inherently painful due to the regret that is caused by the necessity of having to give up certain opportunities (see Festinger, 1964). The agent in our model makes up for the opportunity missed in the first period, and thereby eliminates the regrets that necessarily arise from being forced to choose an action in period 1. In this sense he appears to exhibit a "reparative" tendency as he seeks to relieve the dissonance associated with his period 1 choice.

We close this section with an informal discussion of what happens if the weight on an action's payoff is different across periods ($\lambda \neq 1$). In this case, the agent's opportunities to obtain this type of insurance against regret are limited. On the one hand, if λ is small, then the second period action cannot contribute much to reduce the fluctuations between overall actual and counterfactual payoffs. As in the previous section, the period 2 incentive is then determined by the agent's feedback concerns and there is a regret equilibrium in which the agent sticks to his first period action.

On the other hand, if λ is large, the main fluctuations in overall relative performance come from period 2 payoffs. Thus, the highest correlation between actual and counterfactual payoff can be achieved by matching one's counterfactual period 2 action. Thus, while a standard agent optimally chooses the second period action whose associated payoff is less risky, a regretful agent may optimally choose an action whose associated payoff is stochastically dominated if he believes that his counterfactual self also chooses this action. It follows that when λ is large, asymmetric equilibria can arise in which the agent chooses the same action in period 2 irrespective of this first period action. This matching incentive is at the core of the results of the next section, to which we now turn.

5.2 Correlation across counterfactuals

In this section, we consider a scenario in which the payoffs from first and second period actions are independent and the payoffs from second period actions are correlated across all possible decision nodes. We first derive a result that identifies a tendency of the regretful agent to match the actions that he believes his counterfactual self would choose in period 2 and then consider a career choice example as an application.

The setup is as follows. In period 1, the agent's choice set is $Y^1 = \{a, b\}$ with associated payoffs X_a and X_b. In period 2, his choice set is $Y^2_{y^1} = \{c, d\}$.[32] We denote the payoff associated with a period 2 choice y^2 by $X^{y^1}_{y^2}$, where the superscript y^1 indicates that the random variable depends on which action y^1 was chosen in period 1. We assume that $X^a_{y^2}$ and $X^b_{y^2}$ are stochas-

[32] Strictly speaking, c and d depend on y^1. We suppress this dependency for notational simplicity.

tically independent of X_a and X_b. The key point is that $X^a_{y^2}$ and $X^b_{y^2}$ might be correlated. To provide a parsimonious means of varying the degree of correlation, let X_{y^2} and $\widehat{X}_{y^2}$ be two i.i.d. random variables. Let $X^a_{y^2} = X_{y^2}$, and assume that $X^a_{y^2}$ is equal to X_{y^2} with probability α, and equal to $\widehat{X}_{y^2}$ with probability $1-\alpha$. That is, if $\alpha = 1$, the payoffs from actions are perfectly correlated across decision nodes, and if $\alpha = 0$, they are stochastically independent.[33] The agent's overall material payoff from path (y^1, y^2) is given by $X_{(y^1,y^2)} = X_{y^1} + X^{y^1}_{y^2}$. Finally, we assume for simplicity that the agent does not receive interim information and that the outcome of unchosen random variables is always revealed to the agent.[34]

5.2.1 Matching the counterfactual self

We now describe the agent's period 2 incentives. To isolate the effects that arise as result of the correlation in period 2, we focus on the case, in which the first stage is inconsequential in terms of material payoff. That is, we assume that X_a and X_b are deterministic and equal to zero. Notice that this does *not* mean that the first decision is inconsequential in terms of regret. Rather, it is precisely the distinctive feature of our approach that the regrets about the first decision also include the regrets about the second period action opportunities that arise as a consequence of this first decision.

We analyse the interplay between the relative riskiness of the two options c and d and the degree of correlation across decision nodes. To do so, we assume that X_d second order stochastically dominates X_c and vary the correlation parameter α.

Notice that the standard agent chooses d at the second stage for all α. However, for the regretful agent this does not need to be true. Our first result says that if payoffs are sufficiently correlated across decision nodes, the agent has an incentive to match his counterfactual behaviour.

Lemma 7 *Consider the agent's actual period 2 decision and suppose his counterfactual self chooses action y^2 in period 2. Then there is a $\widehat{\theta}_1 > 0$ such that for all $\theta_1 \geq \widehat{\theta}_1$, there is an $\widehat{\alpha} \in [0,1]$ such that for all $\alpha \geq \widehat{\alpha}$, it is a best response to choose y^2, too.*

To illustrate the intuition, suppose the counterfactual self chooses the riskier option c. The actual self then faces a trade-off between risk-minimisation and regret-minimisation. In terms of risk, option d is superior to option c by assumption. In terms of regret, option c is superior to option d. This is once more due to the agent's desire to minimise the fluctuations in his overall

[33] In terms of the underlying joint distribution F, this means that $F_{X^b_{y^2}|X^a_{y^2}}\left(x^b|x^a\right) = \alpha 1_{[x^a,\infty)}\left(x^b\right) + (1-\alpha) F_{\widehat{X}_{y^2}}(x)$ where $F_{\widehat{X}_{y^2}}$ denotes the marginal distribution of $\widehat{X}_{y^2}$.

[34] In this section, the nature of the feedback the agent receives is unsubstantial.

relative performance. Because payoffs are correlated across decision nodes, he can reduce these fluctuations by chooseing c. When the counterfactual self chooses d, then the same two forces both favour option d for the actual self. Proposition 3 is a direct consequence of the agent's best response (and thus stated without proof).

Proposition 3 *Let θ_1 and α be sufficiently large. Then there are two period 2 regret equilibria, one in which the agent always chooses c in period 2, and one in which he always chooses d in period 2 irrespective of his period 1 choices.*[35]

5.2.2 Application: belief determined motivation

To further illustrate this point, we now consider a career example. Suppose that in period 1, the agent chooses between two occupations, a and b, where, for concreteness, a corresponds to "consultancy" and b corresponds to "the law". In the second stage, the agent chooses an effort level $e \in \{0, 1\}$ at cost ke, $k > 0$, that influences the likelihood of his career success. The key point is that we think of success to be determined by some general unknown ability of the agent that is the same in both occupations. That is, we suppose that the likelihood of success is correlated across occupations. Formally, the agent's ability is given by a state $\omega \in \{0, 1\}$ where each state's prior probability is 1/2. In state ω, the agent's material payoff (gross of effort costs) is ωe in both occupations. For simplicity, we assume that the agent learns his true ability after his effort choice.

The matching incentive gives rise to a complementarity between the agent's conjecture about his counterfactual effort choice and his own effort choice. To see this formally, define Δ_e as the agent's period 2 incentive to choose high effort (the difference between the utility he obtains from working hard and the utility he obtains from shirking) given that his counterfactual self chooses effort e. Then we have:

Proposition 4 *The agent's incentive to choose high effort is higher the higher the counterfactual self's effort, i.e. $\Delta_1 \geq \Delta_0$.*

The result is a straightforward consequence of the matching incentive. If as a consultant the agent believes that he would have worked hard as a lawyer, he better work hard as a consultant so as not to regret not having become a successful lawyer. In other words, the agent's motivation is determined by his general views about himself: his self-image. An agent who is generally optimistic about himself with respect to other possible worlds, will tend to be more motivated

[35] It is easy to see that, generically, these two equilibria are the only pure-strategy equilibria.

than a more pessimistic but otherwise identical agent.[36]

The idea that motivation is driven by self-perceptions is familiar to psychologists. In their influential review of the evidence, Taylor and Brown (1988) conclude that "self-enhancing perceptions ... appear to foster motivation, persistence at tasks, and ultimately, more effective performance." (p. 199) and argue that differences in self-images across individuals can be explained as the result of a sort of self-fulfilling prophecy: someone who is optimistic (pessimistic) about his talents will tend to be more (less) motivated, resulting in better (worse) outcomes. Better (worse) outcomes, in turn, will confirm his optimistic (pessimistic) expectations. For an elaboration of this view, see Aspinwall et al. (2001).

Our approach shares with these accounts the theme of self-fulfilling expectations. However, in our model, it is not expectations about actual future outcomes but about counterfactual comparisons that determine current incentives.[37]

6 Conclusion

In this paper, we have developed an extension of regret theory, which can be used to analyse dynamic decision problems. We have argued that in dynamic decision problems, the need arises to model counterfactual thought processes which extend beyond the mere enumeration of unchosen options to the consideration of an agent's own behaviour in alternative possible worlds. Accordingly, we have gone beyond the original regret theories by developing a conceptual framework within which an agent's counterfactual beliefs and actual behaviour are simultaneously determined as part of an equilibrium of an intrapersonal game. Furthermore, we have shown that our framework generates behavioural predictions that do not arise from risk aversion alone and are therefore testable in principle.

Our approach raises a series of issues that go beyond the analysis that we have undertaken in this paper. For example, it is easy to see that in a static setup, a regretful agent might be better off with a smaller choice set because it reduces the number of alternatives that he has to forgo. In our dynamic setup we can ask whether the agent would actually choose to pre-commit

[36] Of course, in equilibrium, effort choice and self-image have to be consistent. It is straightforward to see that multiple equilibria might arise in our setup. Thus, even though the agent must form his beliefs about his counterfactual self in a rational way, our model is compatible with different, self-sustaining self-images and behaviours.

[37] Markus and Nurius (1986) introduce the notion of possible selves and argue that the desire or fear to become a specific future self provides incentives for current action. While they do emphasise the importance of counterfactual comparisons for construing the *current* self-image (p. 963), they maintain that the attractiveness of a future self is determined by its actual features only but not by the counterfactual comparisons it will engage in.

to less choice given that, after the fact, he might regret his first period commitment.[38]

At a more conceptual level, our theory highlights the need to investigate further the close interaction between regret and counterfactual reasoning. In classic regret theory, the agent engages in comparisons between what he actually did and what he *possibly could* have done counterfactually. We have argued, however, that once we move into the realm of dynamic decision problems, the agent's reference point is inevitably determined by his conjectures about what he *reasonably would or should* have done at a second stage had he made a different first period choice. We elaborate on these considerations in another paper (Krähmer and Stone, 2008).

Finally, evidence suggests that, in reality, the experience of regret interacts in important ways with the experience of disappointment that is generated by the discovery that reality has fallen short of expectations (Roese, 1997). Combining our model with the previously mentioned models of personal equilibrium is likely to be a fruitful way to understand this interaction in more detail.

Appendix

Proof of Lemma 2 By Lemma 1, it is sufficient to show that D_y second order stochastically dominates (SOSD) $D_{y'}$. We begin by showing that the reference points R_y and $R_{y'}$ are constants with $R_y = R_{y'}$. Indeed since Z_y is a constant, we have that $R_y = \max_{\widehat{y}\in Y\setminus y} E\left[\phi\left(X_{\widehat{y}}\right)|X_y\right]$. Moreover, since $\left(X_y\right)_{y\in Y}$ are stochastically independent, $R_y = \max_{\widehat{y}\in Y\setminus y} E\left[\phi\left(X_{\widehat{y}}\right)\right]$. Similarly it follows that $R_{y'} = \max_{\widehat{y}\in Y\setminus y'} E\left[\phi\left(X_{\widehat{y}}\right)\right]$. Now note that $E\left[\phi\left(X_{y'}\right)\right] = E\left[\phi\left(X_y\right)\right]$ by assumption. Thus, $\max_{\widehat{y}\in Y\setminus y} E\left[\phi\left(X_{\widehat{y}}\right)\right] = \max_{\widehat{y}\in Y\setminus y'} E\left[\phi\left(X_{\widehat{y}}\right)\right]$, and hence $R_y = R_{y'}$.

Now recall that $D_y = \phi\left(X_y\right) - R_y$. Since R_y and $R_{y'}$ are constants with $R_y = R_{y'}$, it thus follows that D_y SOSD $D_{y'}$ if $\phi\left(X_y\right)$ SOSD $\phi\left(X_{y'}\right)$. To see the latter, note that X_y SOSD $X_{y'}$ by assumption and that the SOSD property is preserved under a concave transformation. This completes the proof. □

Proof of Lemma 3: We have to show that there is a density function $k\left(z_b|z_a\right)$ such that:

$$E\left[\psi\left(X_b\right)|Z_a = z_a\right] = \int_{\mathcal{Z}_b} E\left[\psi\left(X_a\right)|Z_b = z_b\right] k\left(z_b|z_a\right) dz_b, \tag{1}$$

$$f_{Z_b}\left(z_b\right) = \int_{\mathcal{Z}_a} k\left(z_b|z_a\right) f_{Z_a}\left(z_a\right) dz_a. \tag{2}$$

[38] In a recent paper, Sarver (2008) provides axiomatic foundations for menu preferences such that the agent may want to commit to smaller menus so as to reduce his regrets. His agent does not, however, regret his commitment choice.

Define

$$k\left(z_b|z_a\right) = \tau\left(z_a|z_b\right)\frac{f_{Z_b}\left(z_b\right)}{f_{Z_a}\left(z_a\right)}.$$

As for (1). Note first that for all $x \in \mathcal{X}$, we have $f_{X_b|Z_a}\left(x|z_a\right) = f_{Z_a|X_b}\left(z_a|x\right) f_{X_b}\left(x\right)/f_{Z_a}\left(z_a\right)$. Thus, with the "garbling" equation from Definition 1,

$$f_{X_b|Z_a}\left(x|z_a\right) = \int_{\mathcal{Z}_b} \tau\left(z_a|z_b\right) f_{Z_b|X_a}\left(z_b|x\right) dz_b \frac{f_{X_b}\left(x\right)}{f_{Z_a}\left(z_a\right)}.$$

Using the definition of k, we can write this as

$$f_{X_b|Z_a}\left(x|z_a\right) = \int_{\mathcal{Z}_b} k\left(z_b|z_a\right) f_{Z_b|X_a}\left(z_b|x\right) \frac{1}{f_{Z_b}\left(z_b\right)} dz_b f_{X_b}\left(x\right).$$

Now we can write the conditional density under the integral as

$$f_{Z_b|X_a}\left(z_b|x\right) = \frac{f_{X_a|Z_b}\left(x|z_b\right)}{f_{X_a}\left(x\right)} f_{Z_b}\left(z_b\right)$$

and obtain

$$f_{X_b|Z_a}\left(x|z_a\right) = \int_{\mathcal{Z}_b} k\left(z_b|z_a\right) f_{X_a|Z_b}\left(x|z_b\right) dz_b \frac{f_{X_b}\left(x\right)}{f_{X_a}\left(x\right)}.$$

Since X_a and X_b are identically distributed, the last fraction is 1. Thus, the LHS of (1) is

$$\begin{aligned} E\left[\psi\left(X_b\right)|Z_a = z_a\right] &= \int_{\mathcal{X}} \psi\left(x\right) f_{X_b|Z_a}\left(x_b|z_a\right) dx \\ &= \int_{\mathcal{X}} \psi\left(x_b\right) \int_{\mathcal{Z}_b} k\left(z_b|z_a\right) f_{X_a|Z_b}\left(x|z_b\right) dz_b dx. \end{aligned}$$

Swapping the order of integration delivers the claim. (These operations are well-defined since ψ is integrable, and supports are compact by assumption.)

As for (2). Note that

$$k\left(z_b|z_a\right) f_{Z_a}\left(z_a\right) = \tau\left(z_a|z_b\right) f_{Z_b}\left(z_b\right).$$

Integrating with respect to z_a over both sides and noting that $\tau\left(z_a|z_b\right)$ is a density, i.e. $\int \tau\left(z_a|z_b\right) dz_a = 1$, yields the claim. □

Proof of Lemma 4 Since X_a and X_b are identically distributed, we have to show that $E\left[\chi\left(D_a\right)\right] \geq E\left[\chi\left(D_b\right)\right]$ for any concave $\chi : \mathcal{X} \rightarrow \mathbb{R}$ where $\mathcal{X} = \mathcal{X}_a = \mathcal{X}_b$. Since X_a and Z_a as well as X_a and X_b are independent, we have

$$\begin{aligned} E\left[\rho\left(D_a\right)\right] &= E_{X_a}\left[E\left[\chi\left(\phi\left(X_a\right) - E\left[\phi\left(X_b\right)|Z_a, X_a\right]\right)|Z_a\right]\right] \\ &= E_{X_a}\left[E_{Z_a}\left[\chi\left(\phi\left(X_a\right) - E\left[\phi\left(X_b\right)|Z_a\right]\right)\right]\right]. \end{aligned} \tag{3}$$

By the previous lemma $E[\phi(X_a)|Z_b]$ is a mean preserving spread of $E[\phi(X_b)|Z_a]$. Hence, for all $x_a \in \mathcal{X}$,

$$E_{Z_a}[\chi(\phi(x_a) - E[\phi(X_b)|Z_a])] \geq E_{Z_b}[\chi(\phi(x_a) - E[\phi(X_a)|Z_b])].$$

Thus, (3) is larger than

$$E_{X_a}[E_{Z_b}[\chi(\phi(X_a) - E[\phi(X_a)|Z_b])]].$$

Now note that X_a and X_b are identically distributed and that X_b and Z_b are independent. Thus, instead of integrating over X_a we can integrate over X_b, and the previous expression becomes

$$E_{X_b}[E_{Z_b}[\chi(\phi(X_b) - E[\phi(X_a)|Z_b])]].$$

But using again independence of X_b and Z_b and of X_b and X_a this is equal to $E[\chi(D_b)]$, and this proves the claim. □

Proof of Example 1 We have to show that the agent's incentive to choose a, defined as $U(a) - U(b)$, is non-negative. Since $E[X_a] = E[X_b]$ and $\phi(x) = x$, this incentive is given by

$$\theta E[\rho(X_a - E[X_b|X_a])] - \theta E[\rho(X_b - X_a)].$$

To see that this is non-negative, notice first that since X_a and X_b are symmetrically distributed, the second term in this expression equals $\theta E[\rho(X_a - X_b)]$. Moreover, Jensen's inequality for conditional expectations implies that

$$\rho(X_a - E[X_b|X_a]) \geq E[\rho(X_a - X_b)|X_a].$$

Hence, the first term is not larger than $\theta E[\rho(X_a - X_b)]$. These two observations imply the claim. □

Proof of Proposition 1 We shall look for conditions such that in period 2, the agent sticks to his first period action for all types of news, that is, $s_a(z_a) = a$ for all $z_a \in \{z'_a, z''_a\}$ is a best response against the strategy $s_b(z_b) = b$ for all $z_b \in \{z'_b, z''_b\}$, and thus (s_a, s_b) is a period 2 regret-equilibrium. Suppose that a was chosen in period 1 and suppose that $s_b(z_b) = b$ for all $z_b \in \{z'_b, z''_b\}$. (The case in which b was chosen in period 1 is symmetric.)

If the agent chooses a in the second period, then because X_a and X_b are independent, he learns nothing more about X_b after making his second period decision. Therefore, given his conjecture about his counterfactual behaviour, the reference point R_1 against which he evaluates his first decision is deterministic and given by the expected payoff from the path (b, b)

conditional on the realisation z_a. Hence, $R_1 = E[\phi((1+\lambda)X_b)|z_a]$. Using this reference point and the payoff risk associated with a, the expected utility from choosing a in period 2 is given by

$$u(a,a;z_a) = E[\phi((1+\lambda)X_a)|z_a] + \theta_1 E[\rho(\phi((1+\lambda)X_a) - E[\phi((1+\lambda)X_b)|z_a])|z_a].$$

If the agent chooses b in the second period, he will eventually learn both X_a and X_b. Therefore, his reference point is $R_1 = \phi((1+\lambda)X_b)$, and his expected utility from choosing b is given by

$$u(a,b;z_a) = E[\phi(X_a + \lambda X_b)|z_a] + \theta_1 E[\rho(\phi(X_a + \lambda X_b) - \phi((1+\lambda)X_b))|z_a].$$

Thus, the agent's incentive to choose a, $u(a,a;z_a) - u(a,b;z_a)$, can be written as the sum $\xi_0 + \xi_1$, where the first part is the difference in the material payoff:

$$\xi_0 = E[\phi((1+\lambda)X_a)|z_a] - E[\phi(X_a + \lambda X_b)|z_a],$$

and the second part is the difference in regret with respect to the first decision:

$$\begin{aligned}\xi_1 = \theta_1 (&E[\rho(\phi((1+\lambda)X_a) - E[\phi((1+\lambda)X_b)|z_a])|z_a] \\ &-E[\rho(\phi(X_a + \lambda X_b) - \phi((1+\lambda)X_b))|z_a])\end{aligned}$$

We want to show that for any realisation z_a, the sum $\xi_0 + \xi_1$ is positive if λ is small and θ_1 is large. To do so, notice that ξ_0 is bounded from below. (This is, because the support of X_a and X_b is compact and ϕ is continous, and thus bounded on the support of X_a and X_b.) Hence, it suffices to show that for any realisation z_a, ξ_1 is strictly positive for small λ. We can then choose θ_1 large enough such that ξ_1 outweighs the possibly negative impact of ξ_0.

To see that ξ_1 is strictly positive for small λ, we first subtract and add $E[\rho(\phi((1+\lambda)X_a) - \phi((1+\lambda)(X_b)))|z_a]$ to ξ_1/θ_1 to obtain

$$\begin{aligned}\xi_1/\theta_1 = &E[\rho(\phi((1+\lambda)X_a) - E[\phi((1+\lambda)X_b)|z_a]) - \rho(\phi((1+\lambda)X_a) - \phi((1+\lambda)X_b))|z_a] \\ &+ E[\rho(\phi((1+\lambda)X_a) - \phi((1+\lambda)(X_b))) - \rho(\phi(X_a + \lambda X_b) - \phi((1+\lambda)X_b))|z_a].\end{aligned}$$

Define the difference in the first line of the right hand side as $\delta(\lambda)$. Jensen's inequality yields that $\delta(\lambda) > 0$. The inequality holds strict, because ρ is strictly concave and the conditional distribution of X_a conditional on Z is non-degenerate. Define $\delta_{\min} = \min_{\lambda \in [0,1]} \delta(\lambda)$. Since the minimum is taken over a compact set, we have that $\delta_{\min}$ is strictly positive. We thus obtain:

$$\xi_1/\theta_1 > \delta_{\min} + E[\rho(\phi((1+\lambda)X_a) - \phi((1+\lambda)X_b)) - \rho(\phi(X_a + \lambda X_b) - \phi((1+\lambda)X_b))|z_a].$$

To conclude the proof, it therefore remains to show that the second term in this expression is not smaller than $-\delta_{\min}$ for small λ. To see this, define the arguments in ρ respectively as

$$\begin{aligned}\Delta_1(\lambda, x_a, x_b) &= \phi((1+\lambda)x_a) - \phi((1+\lambda)x_b), \\ \Delta_2(\lambda, x_a, x_b) &= \phi(x_a + \lambda x_b) - \phi((1+\lambda)x_b).\end{aligned}$$

Notice first that $\Delta_1(\lambda, x_a, x_b) - \Delta_2(\lambda, x_a, x_b) \to 0$ for $\lambda \to 0$. Moreover, because the support of X_a and X_b is compact, $\Delta_1(\lambda, x_a, x_b) - \Delta_2(\lambda, x_a, x_b)$ is uniformly continuous for all $\lambda \leq 1$. It follows from these two observations that for all $\tau > 0$ there is a $\lambda_\tau > 0$ such that for all realisations x_a, x_b:

$$|\Delta_1(\lambda, x_a, x_b) - \Delta_2(\lambda, x_a, x_b)| < \tau \text{ for all } \lambda \leq \lambda_\tau .$$

In addition, ρ is uniformly continuous on the compact support of X_a and X_b. Hence, there is a $\tau > 0$ such that for all realisations of Δ_1 and Δ_2 with $|\Delta_1 - \Delta_2| < \tau$: $\rho(\Delta_1) - \rho(\Delta_2) > -\delta_{\min}/2$. This implies that if we now choose $\lambda \leq \lambda_\tau$, then

$$E\left[\rho\left(\phi\left((1+\lambda) X_a\right) - \phi\left((1+\lambda) X_b\right)\right) - \rho\left(\phi\left(X_a + \lambda X_b\right) - \phi\left((1+\lambda) X_b\right)\right) | z_a\right] > -\delta_{\min}/2,$$

which is what we sought to prove. □

Proof of Lemma 5 Suppose first that $\gamma \geq 1$. Notice first that the choice $(y^1, y^2) = (b, b)$ is dominated by the choice (a, a). This is so, because (b, b) is associated with the random variable $2X_b = 2\gamma \widehat{X}_a$, and the choice (a, a) is associated with $2X_a$. But since X_a and $\widehat{X}_a$ are i.i.d., $2X_a$ second order stochastically dominates $2\gamma \widehat{X}_a$ for $\gamma \geq 1$.

Moreover, the incentive to choose (a, b) or (b, a) rather then (a, a) is given by

$$\xi_0(\gamma) = E\left[\phi\left(X_a + \gamma \widehat{X}_a\right)\right] - E\left[\phi\left(2X_a\right)\right]. \tag{4}$$

We shall now show that (a) $\xi_0(\gamma)$ is declining in γ with $\lim_{\gamma\to\infty} \xi_0(\gamma) = -\infty$, and (b) $\xi_0(1) > 0$. This then implies that there is a $\overline{\overline{\gamma}} > 1$ such that the agent chooses (a, b) if and only if $\gamma < \overline{\overline{\gamma}}$, which is the claim that we need to show.

We begin with (b). This follows from the fact that $X_a + \widehat{X}_a$ second order stochastically dominates $2X_a$ and because ϕ is strictly concave.

As for (a), we use the symmetry of X_a and $\widehat{X}_a$ to write

$$\begin{aligned} E\left[\phi\left(X_a + \gamma \widehat{X}_a\right)\right] = & \int_0^\infty \int_0^\infty \left[\phi\left(-\left(x_a + \gamma \widehat{x}_a\right)\right) + \phi\left(x_a + \gamma \widehat{x}_a\right)\right] \\ & + \left[\phi\left(-\left(x_a - \gamma \widehat{x}_a\right)\right) + \phi\left(x_a - \gamma \widehat{x}_a\right)\right] dF_a(x_a)\, dF_a(\widehat{x}_a). \end{aligned}$$

The strict concavity of ϕ implies that for all x_a, $\widehat{x}_a$ the terms in the squared brackets are negative and converge monotonically to $-\infty$ as $\gamma \to \infty$. Thus, the Monotone Convergence Theorem implies that the integral converges monotonically to $-\infty$ as $\gamma \to \infty$. This establishes the claim for $\gamma \geq 1$. The case $\gamma < 1$ follows with identical arguments. □

Proof of Lemma 6 To establish part (i), suppose that $s_b = a$, and that the agent has chosen a in period 1. Then, if he chooses a in period 2, he does not learn X_b and his reference point is $R_1 = E\left[\phi\left(X_b + X_a\right)\right]$. Hence, with $X_b = \gamma \widehat{X}_a$, the expected utility from choosing a in period 2 is

$$u(a) = E\left[\phi\left(2X_a\right)\right] + \theta_1 E\left[\rho\left(\phi\left(2X_a\right) - E\left[\phi\left(X_a + \gamma\widehat{X}_a\right)\right]\right)\right].$$

If he chooses b in period 2, he learns X_b and his reference point is $R_1 = \phi\left(X_b + X_a\right)$. Note that his reference point is equal to his actual overall payoff. Thus, the agent will have no regrets and his expected utility is

$$u(b) = E\left[\phi\left(2X_a\right)\right].$$

Let

$$\xi_1(\gamma) = E\left[\rho\left(\phi\left(2X_a\right) - E\left[\phi\left(X_a + \gamma\widehat{X}_a\right)\right]\right)\right],$$

then the incentive to switch to b in period 2 is given by

$$\Delta_b(\gamma) = u(b) - u(a) = \xi_0(\gamma) - \theta_1 \xi_1(\gamma),$$

where $\xi_0(\gamma)$ is defined by (4). We now show that (a) $\Delta_b(\gamma)$ is declining in γ with $\lim_{\gamma\to\infty} \Delta_b(\gamma) = -\infty$, and (b) $\Delta_b(0) > 0$, and (c) $\Delta_b\left(\overline{\overline{\gamma}}\right) > 0$, where $\overline{\overline{\gamma}}$ is the cutoff at which the standard agent is indifferent between (a, b) and (a, a). This then implies that there is a $\widehat{\widehat{\gamma}} > \overline{\overline{\gamma}}$ such that the regretful agent switches to b if and only if $\gamma < \overline{\overline{\gamma}}$, which is the claim that we want to show.

We begin with (b) and (c) and the observation that Jensen's inequality implies that

$$\xi_1(\gamma) < \rho\left(E\left[\phi\left(2X_a\right)\right] - E\left[\phi\left(X_a + \gamma\widehat{X}_a\right)\right]\right) = \rho\left(-\xi_0(\gamma)\right),$$

where the inequality is strict, since ρ is strictly convave. Thus, $\Delta_b(\gamma) > \xi_0(\gamma) - \theta_1\rho\left(-\xi_0(\gamma)\right)$. As for (b), notice now that $\xi_0(0) > 0$ since X_a second order stochastically dominates $2X_a$ and because ϕ is strictly concave. Hence, all terms on the right hand side of the previous inequality are strictly positive and so $\Delta_b(0) > 0$, establishing (b). As for (c), note that by definition, $\xi_0\left(\overline{\overline{\gamma}}\right) = 0$. Hence, $\Delta_b\left(\overline{\overline{\gamma}}\right) > 0$.

As for (a), we have already established in the proof of Lemma 5 that $\lim_{\gamma\to\infty} \xi_0(\gamma) = -\infty$. Hence, it is sufficient to show that $\lim_{\gamma\to\infty} \xi_1(\gamma) = -\infty$. To see this notice that we have shown in part (a) of the proof of Lemma 5 that $E\left[\phi\left(X_a + \gamma\widehat{X}_a\right)\right]$ converges monotonically to $-\infty$ as $\gamma \to \infty$. Thus, since ρ is increasing, $\rho\left(\phi\left(2x_a\right) - E\left[\phi\left(X_a + \gamma\widehat{X}_a\right)\right]\right)$ converges monotonically to $-\infty$ for all x_a as $\gamma \to \infty$. Thus the integral over this expression converges monotonically to $-\infty$ by the Monotone Convergence Theorem. This establishes part (i).

Part (ii), which deals with the case, in which $s_a = b$, and the agent has chosen b in period 1, follows with identical arguments. □

Proof of Lemma 7 Let $s_{y^2} \in Y^2$ be the agent's period 2 strategy after having chosen action y_1 in period 1. To show the claim, we have to show that for $\alpha \geq \widehat{\alpha}$, $s_a = y^2$ is a best response against $s_b = y^2$ and vice versa. We only show that $s_a = c$ is a best response against $s_b = c$. All the other cases follow with identical arguments.

So let $s_b = c$ and suppose the agent has chosen a in period 1. Then, the counterfactual self obtains $\phi(X_c)$ with probability α and $\phi\left(\widehat{X}_c\right)$ with probability $1-\alpha$. Thus, the agent's reference point is $\phi(X_c)$ with probability α and $\phi\left(\widehat{X}_c\right)$ with probability $1-\alpha$. Hence, his expected utility from choosing c is

$$\begin{aligned} u(c) &= E[\phi(X_c)] + \theta_1 E\left[\alpha\rho(\phi(X_c) - \phi(X_c)) + (1-\alpha)\rho\left(\phi(X_c) - \phi\left(\widehat{X}_c\right)\right)\right] \\ &= E[\phi(X_c)] + \theta_1(1-\alpha)E\left[\rho\left(\phi(X_c) - \phi\left(\widehat{X}_c\right)\right)\right]. \end{aligned}$$

Likewise, his expected utility from choosing d is

$$\begin{aligned} u(d) &= E[\phi(X_d)] + \theta_1 E\left[\alpha\rho(\phi(X_d) - \phi(X_c)) + (1-\alpha)\rho\left(\phi(X_d) - \phi\left(\widehat{X}_c\right)\right)\right] \\ &= E[\phi(X_d)] + \theta_1 E[\rho(\phi(X_d) - \phi(X_c))], \end{aligned}$$

where the second equality holds because X_c and $\widehat{X}_c$ are identically distributed.

We want to show that $u(c)$ is larger than $u(d)$ for sufficiently large θ_1 and α. To see that is true, notice first that $E[\phi(X_d)] \geq E[\phi(X_c)]$, because X_d second order stochastically dominates X_c by assumption. Thus, we need to show that the regret from c is smaller than that from d for sufficiently large α, that is

$$(1-\alpha)E\left[\rho\left(\phi(X_c) - \phi\left(\widehat{X}_c\right)\right)\right] > E[\rho(\phi(X_d) - \phi(X_c))] \tag{5}$$

for sufficiently large α. (We can then choose θ_1 sufficiently large to override material incentives.) To (5), notice first that $\rho(E[\phi(X_d) - \phi(X_c)])$ is a strict upper bound for the right hand side of (5) by Jensen's inequality and because ρ is strictly concave. Moreover, because X_d second order stochastically dominates X_c, this upper bound is negative, and hence the right hand side of (5) is strictly negative. Observe now that the left hand side (5) converges to 0 as α converges to 1. Thus, (5) holds for sufficiently large α, and this completes the proof. □

Proof of Proposition 4 Suppose first the counterfactual self chooses $e = 0$. This gives rise to the reference point $R_1 = \phi(0)$. Thus, the agent's expected utility from choosing $e = 0$ and $e = 1$, respectively, is

$$\begin{aligned} u(0,0) &= \phi(0) + \theta_1\rho(0) = \phi(0), \\ u(1,0) &= \phi\left(\frac{1}{2} - k\right) + \theta_1\left(\frac{1}{2}\rho(\phi(1-k) - \phi(0)) + \frac{1}{2}\rho(\phi(-k) - \phi(0))\right). \end{aligned}$$

Hence,

$$\Delta_0 = u(1,0) - u(0,0) = \phi\left(\frac{1}{2} - k\right) - \phi(0) + \theta_1\left(\frac{1}{2}\rho(\phi(1-k) - \phi(0)) + \frac{1}{2}\rho(\phi(-k) - \phi(0))\right).$$

Suppose next the counterfactual self chooses $e = 1$. If the agent chooses $e = 0$, then his reference point depends on the true state that is revealed to the agent. If he learns that $\omega = 1$, his reference point is $R_1 = \phi(1-k)$, while if he learns that $\omega = 0$, his reference point is $R_1 = \phi(-k)$. Hence, his expected utility is

$$u(0,1) = \phi(0) + \theta_1\left(\frac{1}{2}\rho(\phi(0) - \phi(1-k)) + \frac{1}{2}\rho(\phi(0) - \phi(-k))\right).$$

If the agent chooses $e = 1$, then he matches exactly his counterfactual self's payoff and thus receives expected utility

$$u(1,1) = \phi\left(\frac{1}{2} - k\right) + \theta_1\rho(0) = \phi\left(\frac{1}{2} - k\right)$$

Hence,

$$\Delta_1 = u(1,1) - u(0,1) = \phi\left(\frac{1}{2} - k\right) - \phi(0) - \theta_1\left(\frac{1}{2}\rho(\phi(0) - \phi(1-k)) + \frac{1}{2}\rho(\phi(0) - \phi(-k))\right).$$

We want to show that $\Delta_1 - \Delta_0 \geq 0$. From the above expressions, this difference can be written as

$$\begin{aligned}\Delta_1 - \Delta_0 = -\frac{1}{2}\theta_1 \{ & [\rho(\phi(0) - \phi(1-k)) + \rho(\phi(1-k) - \phi(0))] + \\ & + [\rho(\phi(0) - \phi(-k)) + \rho(\phi(-k) - \phi(0))].\end{aligned}$$

But because ρ is concave, the terms in the square brackets are non-positive, and this implies the claim. □

References

Aspinwall, L. G., L. Richter, and R. R. Hoffman (2001). "Understanding how optimism "works": An examination of optimists' adaptive moderation of belief and behavior." In *Optimism and Pessimism: Theory, Research, and Practice*, edited by E. C. Chang, Washington, American Psychological Association.

Bell, D. E. (1982). "Regret in decision making under uncertainty." *Operations Research*, 30: 961-981.

Bell, D. E. (1983). "Risk premiums for decision regret." *Management Science*, 29: 1156-1166.

Bleichrodt H., A. Cillo, and E. Diecidue (2007), "A Quantitative Measurement of Regret Theory." *mimeo.*

Camerer, C. F. and R. A. Weber (1999). "The econometrics and behavioral economics of escalation of commitment: a reexamination of Staw and Hoang's NBA data." *Journal of Economic Behavior and Organisation*, 39: 59-82.

Camille, N., G. Coricelli, J. Sallet, P. Pradat-Diehl, J.-R. Duhamel, and A. Sirigu (2004): "The involvement of the orbitofrontal cortex in the experience of regret." *Science*, 304: 1167-1170.

Caplin, A. and J. Leahy (2001). "Psychological expected utility theory and anticipatory feelings." *Quarterly Journal of Economics*, 116: 55-79.

Coricelli, G., H. D., Critchley, M. Joffily, J. P. O'Doherty, A. Sirigu, and R. J. Dolan (2005). "Regret and its avoidance: a neuro imaging study of choice behavior." Nature Neuroscience, 8 (9): 1255-1262.

Eyster (2002). "A taste for consistency." *mimeo.*

Festinger, L. and E. Walster (1964). "Post-decision regret and decision reversals." In Festinger, L. (ed). *Conflict, Decision Making, and Dissonance.* Stanford University Press, Stanford, California.

Filiz, E. and E. Y. Ozbay (forthcoming): "Auctions with anticipated regret: theory and evidence." *American Economic Review.*

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989). "Psychological games and sequential rationality." *Games and Economic Behavior*, 1: 60-79.

Harless, D. (1992). "Actions versus prospects: the effect of problem representation on regret." *American Economic Review*, 82: 634-649.

Hayashi, T. (2008): "Context dependence and consistency in dynamic choice under uncertainty: the case of anticipated regret." *mimeo.*

Humphrey, S. J. (2004). "Feedback-conditional regret theory and testing regret-aversion in risky choice." *Journal of Economic Psychology*, 25(6): 839-857.

Inman, J. J., J. S. Dyer, and J. Jia (1997). "A generalized utility model of disappointment and regret effects on post-choice valuation." *Marketing Science*, 16: 97-111.

Irons, B. and C. Hepburn (2003). "Regret Theory and the tyranny of choice." *mimeo.*

Josephs, R. A., R. P. Larrick, C. M. Steele, and R. E. Nisbett (1993): "Protecting the self from the negative consequences of risky decisions." *Journal of Personality and Social Psychology*, 1992, 62 (1): 26-37.

Kahneman, D. and A. Tversky (1979). "Prospect theory: An analysis of decision under risk." *Econometrica*, 47: 263-291.

Kahneman, A. and D. T. Miller (1986). "Norm theory: comparing reality to its alternatives." *Psychological Review*, 93 (2): 136-153.

Kardes, F. R. (1994). "Consumer judgement and decision processes." In Wyer, R. S. and T. C. Srull (eds). *Handbook of Social Cognition*, Vol 2 (2), Hillsdale, NJ, Lawrence Erlbaum Associates.

Kelsey, D. and A. Schepanski (1991). "Regret and disappointment in taxpayer reporting decisions: an experimental study," *Journal of Behavioral Decision Making*: 4, 33-53.

Koszegi, B. (2006). "Ego utility, overconfidence and task choice." *Journal of the European Economic Association*, forthcoming.

Koszegi, B. (2004). "Utility from anticipation and personal equilibrium." *mimeo.*

Koszegi, B, and M. Rabin (2006). "A model of reference dependent preferences." *Quarterly Journal of Economics*, forthcoming.

Krähmer, D. and R. Stone (2008). "Regret and uncertainty aversion." *mimeo.*

Larrick, R. P. and T. J. Boles (1995). "Avoiding regret in decisions with feedback: A negotiation example." *Organizational Behavior and Human Decision Processes*, 63(1): 87-97.

Loomes, G. (1988a). "Further evidence of the impact of regret and disappointment in choice under uncertainty." *Economica*, 55: 47-62.

Loomes G. (1988b). "When actions speak louder than prospects." *American Economic Review*, 78: 463-470.

Loomes, G. and R. Sugden (1982). "Regret theory: An alternative approach to rational choice under uncertainty." *The Economic Journal*, 92: 805-824.

Loomes, G. and R. Sugden (1987a). "Some implications of a more general form of regret theory." *Journal of Economic Theory*, 41: 270-87.

Loomes, G. and R. Sugden (1987b). "Testing for regret and disappointment in choice under uncertainty." *The Economic Journal*, 97 (Supplement): 118-129.

Loomes, G., C. Starmer and R. Sugden (1991). "Observing violations of transitivity by experimental methods." *Econometrica*, 59(2): 425-439.

Loomes, G., C. Starmer and R. Sugden (1992). "Are preferences monotonic? Testing some predictions of regret theory." *Economica*, 59: 17-33.

Markus, H. and P. Nurius (1986). "Possible selves." *American Psychologist*, 41(9): 954-969.

Mellers, B. (2000). "Choice and the relative pleasure of consequences." *Psychological Bulletin*, 126: 910-924.

Mellers, B., A. Schwartz and I. Ritov (1999). "Emotion-based choice." *Journal of Experimental Psychology: General*, 128: 332-345.

Mellers, B. and A. P. McGraw (2001). "Anticipated emotions as guides to choice." *Current Directions in Psychological Science*, 10(6): 210-214.

Prendergast, C. and L. Stole (1996). "Impetuous youngsters and jaded old timers: acquiring a reputation for learning." *Journal of Political Economy*, 104(6): 1105-1134.

Rabin, M. (1998). "Psychology and Economics." *Journal of Economic Literature*, 36: 11-46.

Ritov, I. (1996). "Probability of regret: anticipation of uncertainty resolution in choice." *Organizational Behavior and Human Decision Processes*, 66(2): 228-236.

Roese, N. J. (1997). "Counterfactual thinking." *Psychological Bulletin*, 121(1): 133-148.

Sarver, T. (2008). "Anticipating regret: why fewer options are better." *Econometrica*, 76(2), 263-305..

Simonson, I. (1992). "The influence of anticipating regret and responsibility on purchase decision." *Journal of Consumer Research*, 19:105-118.

Starmer, C., and R. Sugden (1993). "Testing for juxtaposition and event-splitting effects." *Journal or Risk and Uncertainty*, 6: 235-254.

Starmer, C., and R. Sugden (1998). "Testing alternative explanations for cyclical choice." *Economica*, 65: 347-361.

Taylor, S. E. and J. D. Brown (1988). "Illusion and well-being: a social psychological perspective on mental health." *Psychological Bulletin*, 103(2): 193-210.

Taylor, K. A. (1997). "A regret theory approach to assessing consumer satisfaction." *Marketing Letters*, 8(2): 229-238.

Tsiros, M. and V. Mittal: (2000). "Regret: A model of its antecedents and consequences in consumer decision making." *Journal of Consumer Research*, 26: 401 - 417.

Tykocinski, O. E. and T. S. Pittman (1998). "The consequences of doing nothing: Inaction inertia as avoidance of anticipated counterfactual regret." *Journal of Personality and Social Psychology*, 75(3): 607-616.

Van de Ven, N. and M. Zeelenberg (2006). "Consequences of regret aversion: Two lottery experiments." *mimeo*

Yariv, L. (2005). "I'll see it when I believe it: A simple model of cognitive consistency." *mimeo.*

Zeelenberg, M. (1999). "Anticipated regret, expected feedback, and behavioral decision making." *Journal of Behavioral Decision Making*, 12: 93-106.

Zeelenberg, M. (2004). "Consequences or regret aversion in real life: the case of the Dutch postcode lottery." *Organizationsl Behavior and Human Decision Processes*, 93: 155-168.

Zeelenberg, M., J. Beattie, J. Van Der Pligt, and N. K. De Vries (1996). "Consequences of regret aversion: Effects of expected feedback on risky decision making." *Organizational Behavior and Human Decision Processes*, 65(2): 148-158.

Zeelenberg, M., and J. Beattie (1997). "Consequences of regret aversion 2: Additional evidence of the effects of feedback on decision making." *Organizational Behavior and Human Decision Processes*, 72(1): 63-78.