THEORETICAL PERSPECTIVES ON RECIPROCITY AND INFORMATION PROCESSING

Inaugural-Dissertation
zur Erlangung des Grades Doctor oeconomiae publicae
(Dr. oec. publ.)
an der Ludwig-Maximilians-Universität
München

2013

Vorgelegt von
Clemens Rainer König

Referent: Prof. Dr. Klaus M. Schmidt
Korreferent: Prof. Dr. Martin G. Kocher
Promotionsabschlussberatung: Mittwoch, 6. November 2013

Datum der mündlichen Prüfung: 31. Oktober 2013

Namen der Berichterstatter: Klaus M. Schmidt, Martin G. Kocher, Fabian Herweg

Danksagung

Zunächst möchte ich meinem Erstbetreuer Klaus Schmidt für seine Unterstützung danken. Seine Forschung zu sozialen Präferenzen war für mich der Anstoß zur eigenen Beschäftigung damit, wie Fairness und Reziprozität menschliches Verhalten bestimmen, und seine Aufgeschlossenheit diesen Fragen gegenüber habe ich stets geschätzt. Meinem Zweitbetreuer Martin Kocher danke ich ebenfalls für seine Hilfsbereitschaft und Freundlichkeit, die mir den Weg manches Mal erleichtert haben. Zuletzt gebührt Fabian Herweg Dank dafür, dass er sich als Drittbetreuer zur Verfügung gestellt hat und zuvor immer ein offenes Ohr für Fragen und Probleme hatte.

Meine Zeit an der Munich Graduate School of Economics habe ich nicht nur wegen der Möglichkeit zur selbständigen Forschung über Dinge, die mir am Herzen liegen, genossen, sondern auch wegen der vielfältigen Kontakte zu meinen Mitdoktoranden. Es war eine bereichernde Erfahrung, auf so viele talentierte und in ihren Persönlichkeiten überraschend verschiedene Menschen treffen zu dürfen. Ein spezieller Dank gehört meinen Lehrstuhlkollegen, mit denen zusammenzuarbeiten über die Jahre hinweg Freude bereitet hat. Im Kreis letzterer findet sich mein Ko-Autor Johannes Maier, dem ich für die gemeinsame Arbeit am dritten Kapitel und seine Ermutigung bei der Fertigstellung der beiden anderen Kapitel danken möchte.

Zu guter Letzt möchte ich meiner Familie und den Freunden außerhalb der Universität für ihre Begleitung auf diesem Weg danken. Ohne die Hilfe besonders meiner Eltern wäre vieles schwerer gefallen, und sie waren meine ersten Ratgeber und Ermutiger in schwierigen Stunden. Muhabbat danke ich für ihre Geduld und ihr Vertrauen, dass alles ein gutes Ende finden würde.

Contents

Preface 1

1 Forgone-Option Effects In Two-Player Games and Net-Loss Reciprocation 6

1.1 Introduction 6

1.2 Losses and Gains From the Other Player's Strategy 12

1.3 Reciprocating the Other Player's Strategy 16

1.4 Applications 19

1.4.1 The Hidden Cost of Control 20

1.4.2 Trust 22

1.4.3 Ultimatum Bargaining 23

1.4.4 Lost Wallets 25

1.5 Discussion 27

1.6 Conclusion 29

1.7 Appendix A: Proofs 30

1.8 Appendix B: Intention-Based Models of Reciprocity 34

2 The Power of Delegation 37

2.1 Introduction 37

2.2 The Model: Net-Loss Reciprocation 40

2.2.1 Derived Net Losses 42

2.2.2 Imposed Net Losses and Preferences 48

2.3 The Power of Delegation 50

2.3.1 Negative Reciprocity 51

2.3.2 Positive Reciprocity 61

2.3.3. Discussion 61

2.4 Conclusion 63

2.5 Appendix A: Proofs 65

3 Belief Choice Under Prior Regret 72

3.1 Introduction 72

3.2 The Model 78

3.2.1 Setup 78

3.2.2 Belief Choice Problem 79

3.2.3 No Regret 82

3.2.4 Strong Regret 84

3.2.5 Certainty 86

3.3 Applications 87

3.3.1 Conservative and Asymmetric Information Processing 89

3.3.2 Belief Dynamics 94

3.4 Conclusion 108

3.5 Appendix A: Proofs 109

3.6 Appendix B: Eliciting Individual Regret Sensitivities 123

3.7 Appendix C: Anticipatory Utility and Asymmetry 124

Bibliography 126

List of Figures

2.1 Allocations of Punishment in all Unfair Situations 53

2.2 Punishment Distributions in the Unfair Situations in D&P 55

2.3 Punishment Distributions in the Unfair Situations in Random 57

2.4 Punishment Distributions in various Unfair Situations 59

2.5 Punishment Distributions in all Fair Situations 60

3.1 Optimal and Adopted Beliefs 90

3.2 Optimal Beliefs and Over- and Underconfidence 95

3.3 Optimal Beliefs under Mild Regret 101

3.4 Optimal Beliefs under Intermediate Regret 103

3.5 Optimal Beliefs under Strong Regret 104

PREFACE

In the last thirty or so years, microeconomics has increasingly become a behavioural science. A large number of researchers now uses the tools of mathematical modelling that microeconomics offers to arrive at a deeper understanding of what drives human behaviour in economically relevant settings. Traditional microeconomics (Mas-Colell et al, 1995) often took preferences for granted. More emphasis was put on the formal analysis of decision problems with given preferences than asking how real persons would act in such situations, i.e., how they would conceive of the situation, what their motives would be, what expectations they would have about events affecting their decisions etc.

Microeconomics as a behavioural theory, which is often referred to as *behavioural economics*, addresses such questions without relinquishing the traditional strengths of microeconomics like rigorous formal modelling or theoretical parsimony (Rabin, 2002). In this quest, it has been aided by its sister field experimental economics, which by now offers a wealth of data about human behaviour generated in controlled laboratory experiments involving real monetary stakes.

This dissertation aims to situate itself within the field of behavioural microeconomics. In it, I develop (and in part test using data from experimental economics) models of human behaviour pertaining to two topics: On the one hand, *reciprocity*, which is a fundamental concept about human motivation. On the other hand, *information processing*, which addresses how people deal with new information that has become available to them.

Consider first reciprocity. The key idea here is that people's attitude towards others and in particular their willingness to help or hurt them at a material cost depends on how these others have behaved or can be expected to behave towards them. Reciprocity has been shown to be an important force in economic decision making (Fehr and Schmidt, 2006). In this dissertation, I put forward a new model of reciprocity and use it to explain various experimental phenomena.

Chapter 1 proposes an explanation for *forgone-option effects*. Forgone-option effects refer to the phenomenon that individuals ("players") choose differently from the same set of outcomes in different situations because the individuals with whom they interact have had different alternatives to their actual behaviour (different "forgone options"). For example, a player might reject an offer in a bargaining situation if a better offer could have been made, but accept the same offer if all available alternatives would have been inferior.

I explain forgone-option effects by a behavioural model entitled *net-loss reciprocation*. Net-loss reciprocation means that a player's willingness to impose a *loss* on some other player increases in the loss that the player derives from the other player's behaviour. Likewise, the willingness to impose a *gain* increases in the gain derived from the other. Taking the two together yields *net losses*, which are simply losses minus gains (where losses may loom larger). We can therefore summarise the above by saying that the player's willingness to impose a net loss on the other increases in the net loss derived from the other's behaviour.

Both imposed and derived losses and gains have a *material* and a *fairness* component. For example, the loss from receiving some offer in a bargaining situation not only depends on how much more one could have earned under the alternatives, but also on whether these alternatives are fairer than the actual offer. If some alternative offer from which one could have earned more is fairer, one derives a fairness loss on top of one's material loss.

As I argue in detail below, material and fairness considerations need not coincide. The alternative under which one could have earned most need not be the fairest alternative. As a result, overall losses being the sum of material and fairness losses, the highest overall loss need not derive from the alternative creating the highest material loss. Falk et al (2003) provide experimental evidence for the validity of this assumption in a bargaining context: They find that responders to some fixed offer do not reject it most in the situation where they derive the highest material loss. Rather, they reject it most if they derive an intermediate material loss that results from an alternative that is clearly fairer than what they are being offered. This suggests that losses are not a monotonic function of material losses, but should include a fairness component. A similar case can be made for gains.

I show that net-loss reciprocation can explain the forgone-option effects documented in a number of experimental studies, whereas existing models of intention-based reciprocity (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006) or outcome-based social preferences (Fehr and Schmidt, 1999; Charness and Rabin, 2002) fail to explain all the evidence. The evidence includes the bargaining experiments by Falk et al (2003) mentioned in the previous paragraph.

A second piece of evidence is the "hidden cost of control" (Falk and Kosfeld, 2006), which refers to the phenomenon that people dislike seeing their choice set restricted even if they would not have chosen the eliminated elements. In an organisational context, even agents who do not intend to cheat or steal may resent being too closely monitored by the organisation's principal causing them to become less cooperative than otherwise. The hidden cost of control can be thought of as a forgone-option effect: Agents whose choice set the principal has restricted, but could have left unrestricted, behave less cooperatively than agents facing the same restricted choice set in a game where the principal has had no alternative. In the experiment of Falk and Kosfeld (2006), this kind of forgone-option effect occurs despite the fact that the restriction only eliminates the most glaringly unfair options from the agents' choice set. In terms of my model, agents derive no fairness loss from the restriction. Nevertheless, net-loss reciprocation can explain the difference in agents' behaviour because of their material loss. All in all, Chapter 1 makes clear that both fairness and material factors play a crucial part in net losses.

In Chapter 2, I apply the model of net-loss reciprocation to a different economic problem, namely, the *delegation* of decision rights. I focus on a reason for delegation that has only recently received attention from economists, namely, delegation as a means to avoid responsibility and hence punishment for decisions imposing negative externalities on others. Recent experimental work confirms the validity of this particular delegation motive: As shown by Bartling and Fischbacher (2011), a principal can avoid most of the punishment for an unpopular decision by delegating the latter to an agent (whose material interests are aligned with hers). For this reason, the principal is usually better off delegating than taking the decision herself despite the loss of control that delegation entails. I refer to this phenomenon as the *power of delegation* and show that the multi-player version of net-loss reciprocation can explain the punishment patterns sustaining it well. This is also true in comparison to existing theories like intention-based reciprocity or the responsibility model of Bartling and Fischbacher (2011).

The multi-player (including Nature) version of net-loss reciprocation uses the same building blocks as the two-player version, namely, material and fairness losses and gains, where the definition of fairness has been adjusted accordingly. A critical issue that emerges is how a player's loss or gain from some other player's strategy depends on the behaviour of third parties. To address this, I posit that each player holds a belief about the likely behaviour of third parties, which affects his sense of loss and gain from the other's behaviour. Such beliefs about third parties were measured by Bartling and Fischbacher (2011). In the analysis of their punishment data, I find that the distinction between material and fairness factors, which turned out to be crucial in Chapter 1, is somewhat blurred by the fact that material and fairness considerations point into the same direction in the delegation games. Nevertheless, the preferred parameterisation in Chapter 2 for explaining the punishment data largely confirms the specification that best fits the data in Chapter 1.

The second topic that this dissertation touches upon is that of *information processing*, which refers to how people take into account new information about uncertain events affecting them. As is customary, these events are referred to as "states of the world" or simply "states". The traditional assumption in microeconomics is that people are *Bayesians*, which means that they use Bayes' rule to calculate the conditional probability of the different states given their new information and make decisions based on this updated probability called the "Bayesian posterior". In contrast, the probability assigned to the different states before receiving the information is called the "prior belief" or simply "prior". Instead of presuming that people apply Bayes' rule mechanically, Chapter 3, which is joint work with Johannes Maier, advances the idea that people respond to new information by *choosing* a new belief about the world. The proposed choice procedure is entirely instrumental because it discriminates among beliefs based solely on the actions that the decision maker would take if adopting these beliefs. In contrast, existing models of belief choice assume that beliefs have non-instrumental value, which may derive from such motives as wanting to feel good about the action implemented by one's belief *ex ante* (Brunnermeier and Parker, 2005). No such assumptions are needed in our model.

The instrumental value of beliefs has two components: One is *objective performance*, which means that the decision maker prefers beliefs implementing actions that perform as well as possible under the correct Bayesian posterior containing the new information. The second objective in belief choice involves the decision maker's prior belief. The thought here is that the decision maker regrets giving up her prior in the light of the new information if it turns out afterward that holding on to her prior would have been preferable. More specifically, the decision maker considers each state and asks herself how the action that she would choose if holding on to her prior compares to the action that she would choose if adopting some alternative belief. For every alternative belief implementing a different action than the prior, at least one state exists where the prior performs better, which creates scope for *prior regret*.

Optimal beliefs achieve the best balance between objective performance and regret avoidance, i.e., minimising expected regret relative to the reference action implemented by the prior.

We use our model of belief choice under prior regret to account for *asymmetric information processing*, which means that a person who is confident in some good (for his or her "ego") state of the world ignores news indicating an alternative bad state more than does a person who is equally confident in the bad state and receives news indicating the good state (which is at odds with Bayes' Rule). We also consider *preferences for consistency*, which refers to a propensity to replicate earlier behaviour irrespective of new information. Finally, we study the conditions under which *over*- or *underconfidence* (too optimistic or too pessimistic beliefs relative to the Bayesian benchmark) occurs in a dynamic setting.

Disparate though the topics covered in this dissertation may seem, there is a theme shared by them, which is that of *reference dependence*. Reference dependence means that the consequences of decisions are evaluated relative to a reference point. A prominent example is loss aversion (Kahnemann and Tversky, 1979; Köszegi and Rabin, 2006). Chapter 3 uses a similar idea: The reference point for the evaluation of beliefs is an action, namely, the action induced by the prior. In net-loss reciprocation, the reference point is the actual strategy chosen by the other player, relative to which losses and gains are calculated.

Both models proposed in this dissertation are consistent with the asymmetry between losses and gains that is a cornerstone of the literature on reference-dependent preferences, namely, that "losses loom larger than gains". In the belief choice model, the focus is on regret (i.e., losses) alone, which suffices to generate the desired results. In the analysis of net-loss reciprocation, it turns out that losses are more important.

Nevertheless, it would be incorrect to regard net-loss reciprocation as a model of loss aversion in the conventional sense. Loss aversion presupposes some baseline utility attached to the outcomes of decisions. Deviations of baseline utility from the reference level then translate into sensations of loss or gain. In contrast, net-loss reciprocation is a theory about baseline utility itself. It asks how people evaluate the different outcomes of the game available to them given the behaviour of others.

1 FORGONE-OPTION EFFECTS IN TWO-PLAYER GAMES AND NET-LOSS RECIPROCATION

1.1 Introduction

This chapter studies forgone-option effects in two-player games. Its point of departure is a basic decision problem faced by each player: Choosing an outcome of the game from the opportunity set created by the other player's strategy. A forgone-option effect arises if a player's choice from a fixed opportunity set of this kind is not constant across games, but varies with the other player's *forgone options*, which are simply the alternative strategies at the other's disposal. Examples are provided below. They show that forgone-option effects may occur both in the context of negative reciprocity (punishment for unkind behaviour) and positive reciprocity (reward for kind behaviour). While experimental evidence suggests an important role for forgone-option effects in negative reciprocity, experimental studies of positive reciprocity have not established strong forgone-option effects.[1] In this chapter, I propose a new behavioural theory called *net-loss reciprocation* that can explain these apparently contradictory findings.

A new theory is called for because existing theories struggle to explain forgone-option effects. According to outcome-based theories of behaviour, players are motivated by a single preference ordering on outcomes, which directly rules out any forgone-option effects.[2] Existing theories that are not outcome-based also fail to explain all the evidence.[3] From an economic point of view, forgone-option effects are important because they impact the possibility of

---

[1] See Brandts and Solá (2001) and Falk et al (2003) for evidence on negative reciprocity and Dufwenberg and Gneezy (2000), Charness and Rabin (2002), McCabe et al (2003), Cox (2004), Servatka and Vadovic (2009) and Cox et al (2010) for evidence on positive reciprocity.

[2] Outcome-based theories leave open the possibility that several outcomes are most preferred in a given opportunity set. This can rationalise isolated instances of forgone-option effects, but can hardly be regarded as a systematic explanation for their prevalence. Outcome-based theories include Fehr and Schmidt (1999) and Charness and Rabin (2002).

[3] Below, I focus on the intention-based models of reciprocity by Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006).

reaching (materially) efficient outcomes in strategic interactions. Forgone-option effects strongly influence the occurrence of punishment, which in itself harms efficiency, but may improve the overall efficiency of the interaction.[4] Positive reciprocity, which promotes efficient behaviour like trust in trading relationships, appears less affected by forgone-option effects. Yet, it is interesting to explore mechanisms behind this, which may also shed light on robustness.

Net-loss reciprocation builds on two key concepts: Firstly, the loss and gain that players derive from the other player's strategy. Secondly, the loss and gain that players themselves impose on the other through their own choice from the feasible set given the other's strategy. Net-loss reciprocation means that players' willingness to pay for increasing the net loss imposed on the other increases in the net loss derived from the other's strategy.[5] Net losses are simply losses minus gains, where the two need not count for the same. Net-loss reciprocation is consistent with an intuitive notion of reciprocity, according to which people's kindness to others increases in these others' kindness to them. More importantly, net-loss reciprocation can explain forgone-option effects because players may derive different net losses from two strategies creating the same opportunity set because of different alternatives at the other's disposal. As a result, their preferences on the same opportunity set may differ.

For an illustration of forgone-option effects in the domain of *negative reciprocity*, consider the ultimatum mini games studied by Falk et al (2003). In all of them, the proposer can make a fixed offer of dividing the surplus, namely, "8 for the proposer, 2 for the responder", and one alternative offer (his forgone option) that varies. The responder can accept or reject any offer. Consider two games: One where the alternative is "5 for both" and one where it is "2 for the proposer, 8 for the responder". Since the opportunity sets after the fixed offer are the same, we have a forgone-option effect if responders are more likely to reject the fixed offer in one game than the other. Falk et al (2003) report significantly more rejection in the first game.

For an illustration in the context of *positive reciprocity*, consider the trust games studied by Dufwenberg and Gneezy (2000). In all of them, the second mover can share 20 between himself and the first mover in the event of trust, while the games differ regarding the outcome in case of

---

[4] The efficiency-promoting role of punishment has been extensively studied in the context of public good games (Fehr and Gächter, 2000). The evidence on whether punishment opportunities promote overall material efficiency in these games is mixed.

[5] This willingness to pay may be negative or positive.

no trust. Consider two games: One where the no-trust outcome is “4 for the first mover, 0 for the second mover” and one where it is “16 for the first mover, 0 for the second mover”. Since sharing opportunities are identical, we have a forgone-option effect if second movers share the 20 differentially in the two games. Despite the fact that the no-trust outcomes differ considerably, Dufwenberg and Gneezy (2000) report no significant difference in sharing.

Below, I put forward a utility model to explain these and other experimental findings. I now sketch the basic features of the calculus of losses and gains underlying net-loss reciprocation. I focus on the loss and gain that a player derives from the other player’s strategy. Analogous procedures are used for determining the loss and gain imposed on the other.

A player’s *loss* from some strategy of the other player (called henceforth the “status quo”) is composed of his material loss and his fairness loss. The player derives a *material loss* whenever there is an alternative strategy of the other under which the player could have earned more than what he can maximally earn under the status quo. Moreover, the player derives a *fairness loss*, which he adds to his material loss, if his forgone earnings (that make up his material loss) derive from outcomes that are fairer than the fairness of the status quo. The intuition is that the player in this case feels an entitlement (“fairness claim”) to his forgone earnings, which causes him to feel an additional loss. Since his fairness loss is added to his material loss, his total loss then exceeds his material loss. Fairness is measured by a function that ranks the outcomes of the game according to their fairness and incorporates considerations of material efficiency and a concern for the less well-off player.

Likewise, a player’s *gain* from the status quo is composed of his material gain and his fairness gain. The player derives a *material gain* whenever he can earn more under the status quo than what he can maximally earn under some alternative. Moreover, the player derives a *fairness gain*, which is added to his material gain, if the earnings that make up his material gain derive from outcomes that are less fair than the fairness of the alternative. The intuition is that in the converse scenario, where the outcomes creating his material gain are fairer than the alternative, the player feels an entitlement to his material gain (“well deserved”), which causes him to feel no gain based on fairness considerations.

To illustrate the implications of this approach, consider the examples discussed above. As for the ultimatum mini games, recall that significantly more responders reject the status quo offer “8 for the proposer, 2 for the responder” if the alternative is “5 for both” than if it is “2 for the proposer, 8 for the responder". The higher willingness of responders to impose a net loss by rejecting the status quo is explained by net-loss reciprocation if they derive a higher net loss from the status quo in the first game. And indeed, if the alternative is "5 for both", responders derive both a material loss (of 3) and a fairness loss, while their loss is limited to their material loss (of 6) if the alternative is "2 for the proposer, 8 for the responder". Intuitively, "5 for both" is fairer than the status quo, whereas "2 for the proposer, 8 for the responder" is only as fair. Below, I show that the sum of material and fairness loss in the first game can exceed the higher material loss in the second game.

In the two trust games, Dufwenberg and Gneezy (2000) detect no forgone-option effect. Net-loss reciprocation can explain this if the net loss experienced by second movers is the same in each case.[6] Recall that the games differ regarding the no-trust outcome: No trust entails "4 for the first, 0 for the second mover" in the first game and "16 for the first, 0 for the second mover" in the second game. In each case, second movers suffer no loss from being trusted, and their material gain is the same (namely, 20) because they can earn up to 20 after trust, but earn zero after no trust. Their fairness gain is zero in both cases because there is no outcome after trust that is less fair than either no trust outcome, which means that second movers feel entitled to their additional earning possibilities after trust. Even the outcome "0 for the first, 20 for the second mover", which second movers can implement after trust, is not less fair than either no trust outcome because it contains a larger total payoff and the same minimal payoff. As a result, second movers derive the same (negative) net loss from trust in each case.

In contrast, existing models of social preferences struggle to account for forgone-option effects. Outcome-based models like the inequality aversion model of Fehr and Schmidt (1999) or the model of Charness and Rabin (2002), which combines a taste for material efficiency with generosity towards those who are least well-off, build on the idea that strategic behaviour derives from a single ranking of the outcomes of the interaction. For this reason, these models cannot provide a systematic explanation of forgone-option effects.[7]

---

[6] Of course, an insignificant treatment effect can have other reasons such as too few observations.

[7] Forgone-option effects can only arise from several most preferred outcomes in the fixed opportunity set.

Such an explanation can in principle be provided by intention-based models of reciprocity making use of psychological game theory.[8] Most widely used are Dufwenberg and Kirchsteiger (2004), who build on Rabin (1993), and Falk and Fischbacher (2006). The key feature of these models is their reliance on players' second-order beliefs, i.e., their beliefs about the other player's belief about their own choice of strategy. Thus, when faced with some strategy of the other player, players consult their second-order belief, which together with the strategy of the other pins down a unique outcome of the game, which serves to represent the other's strategy. In contrast, my approach does not rely on second-order beliefs, but represents the other's strategy and its alternatives by the entire sets of feasible outcomes that these strategies create. Consequently, my approach is more amenable to empirical testing using standard experimental data as it does not require measurement of higher-order beliefs.[9]

Regarding predictions, both models assert that players' willingness to be kind to the other increases in the kindness of the other's status quo strategy. Problems arise in the conceptualisation of kindness. For instance, in Dufwenberg and Kirchsteiger (2004), if the outcome representing the status quo gives the player more (less) than half of what he maximally and minimally stands to earn under the alternatives, the status quo is perceived as kind (unkind). Thus, this approach limits itself to comparing earnings without allowing for players' sense of entitlement to these earnings. The evidence from the ultimatum mini games makes clear that such a sense of entitlement may override material considerations.

Also, reliance on second-order beliefs may lead to unintuitive predictions because it uses players' (likely) reaction to the strategy of the other as a means to assess the strategy's kindness. Yet, if players react to some strategy that they in fact perceive as unkind in a self-serving manner and more generously to some alternative they perceive as kind, they may end up with more own payoff when faced with the former, from which Dufwenberg and Kirchsteiger (2004) would conclude that the first strategy is kinder. This problem is also shared by Falk and

---

[8] Psychological games were first defined and analysed by Geanakoplos, Pearce and Stacchetti (1989). A framework for dynamic psychological games is provided by Battigalli and Dufwenberg (2009). Another model of non-selfish preferences drawing on psychological game theory is guilt-aversion (Battigalli and Dufwenberg, 2007).

[9] Dhaene and Bouckaert (2010) investigate the performance of the model of Dufwenberg and Kirchsteiger (2004) using measured second-order beliefs in a setting unrelated to forgone-option effects.

Fischbacher (2006). When discussing applications, I explain in more detail this and other difficulties encountered by the two models, whose basic building blocks are laid out in more detail in Appendix B. All in all, net-loss reciprocation can account for larger parts of the evidence than either model of intention-based reciprocity.

At first blush, the model of net-loss reciprocation introduced in this chapter could be thought of as a model of loss aversion (Kahnemann and Tversky, 1979; Shalev 2000; Köszegi and Rabin, 2006). While I allow for the possibility that "losses loom larger than gains", which is the cornerstone of this literature, there are important differences. Loss aversion builds on the idea that the consequences of decisions are evaluated relative to some (deterministic or stochastic) reference point. If the utility of a consequence exceeds (falls short of) the reference point, individuals perceive a gain (loss). Loss aversion therefore presupposes some baseline utility attached to consequences from which losses and gains can be calculated. This is where net-loss reciprocation steps in, which is best described as a theory about how sensations of loss and gain derived from the other's behaviour act as a source of (social) preferences and hence as a source of baseline utility attached to the different outcomes of the game.[10]

That said, the utility model proposed below is qualitative in the sense that the details of the utility function up to the net-loss reciprocation property are left open. For this reason, there is also no equilibrium analysis.[11] Equilibrium analysis is refrained from because it is not needed to explain the phenomena this chapter sets out to explain. The forgone-option effects addressed below relate to the behaviour of players who have certainty about the other player's strategy because they are second movers in sequential games where each player has one move. As a result, these players' beliefs are pinned down by the game's information structure, and a notion of best response is enough to explain their choices.[12] Regarding best responses, I take no stance on which specification of utility consistent with net-loss reciprocation is most appropriate.

[10] In this sense, loss aversion is orthogonal to net-loss reciprocation. Shalev (2000) studies loss aversion in games.

[11] Cox et al (2008) propose a non-equilibrium model of reciprocity in sequential games. However, it is not suited to studying forgone-option effects.

[12] Of course, net-loss reciprocation could also be used to explain the behaviour of first movers with the added complication that their beliefs about the other's strategy are unobservable. These beliefs could be measured experimentally. However, there are to the best of my knowledge no economic experiments documenting forgone-option effects in the behaviour of players who must form beliefs about others.

There are several plausible ways of incorporating net-loss reciprocation into a full-fledged utility model.[13] Comparing the relative performance of these modelling options is left for future work. The substantive question addressed in this chapter is: Can net-loss reciprocation together with the method for calculating losses and gains introduced below explain the forgone-option effects we observe in experiments? The answer is largely affirmative.

The remainder of this chapter is structured as follows: I first show how to calculate the net loss that a player derives from the other player's strategy, which is followed by the method for calculating the net loss that the player himself imposes on the other as well as the utility model incorporating net-loss reciprocation. I then show how this qualitative model can account for forgone-option effects in a number of well-known experimental studies and compare its performance to intention-based models of reciprocity. All proofs are in Appendix A.

1.2 Losses and Gains From the Other Player's Strategy

I limit attention to finite-horizon two-player (*i*, *j*) multi-stage games with observable past actions.[14] A player's inactivity at a stage is modelled by the respective action set being singleton. Let H be the set of non-terminal histories of the game. Player *i*'s *pure strategy* $s_i \in S_i$ assigns to each history $h \in H$ an action available to *i* at *h*.[15] I restrict attention to pure strategies. The set of pure strategy profiles is $S = S_i \times S_j$. Outcomes $\pi = (\pi_i, \pi_j)$ of the game are two-dimensional vectors of material payoffs. The function $\pi : S \rightarrow \mathbb{R}^2$ is the *outcome function* and $\Pi = \{\pi(s) : s \in S\}$ the set of attainable outcomes in the game. Moreover, the set of attainable outcomes or *opportunity set* for player *i* given that player *j* plays strategy $s_j \in S_j$ is given by $\Pi^{s_j} = \{\pi(s_i, s_j) : s_i \in S_i\}$ with $\Pi^{s_j} \subseteq \Pi$.

---

[13] E.g., players could be willing to sacrifice own payoff to match the net loss they impose on the other to the net loss imposed on them. Players could also be endowed with some baseline preferences on outcomes whose degree of altruism decreases in the net loss they derive from the other (although this specification does not perfectly fit the definition of net-loss reciprocation given below).

[14] I refer to *i* as "he" and *j* as "she".

[15] Action sets are assumed to be finite.

I first define player *i*'s *loss* from $s_j \in S_j$. Player *i*'s overall loss is derived from a more basic notion, namely, his loss from s_j relative to a particular alternative $\tilde{s}_j \in S_j$. I also refer to s_j as the *status quo* and to $\tilde{s}_j$ as the *alternative*. The basic idea is the following: Player *i* suffers a loss from s_j relative to $\tilde{s}_j$ only if he can earn more given $\tilde{s}_j$ than what he can maximally earn given s_j. Or, more formally and with a slight abuse of notation, let $\Pi^{\tilde{s}_j > s_j}$ be the set of outcomes in $\Pi^{\tilde{s}_j}$ that yield *i* a higher payoff than his highest attainable payoff in Π^{s_j}. A necessary condition for *i* suffering a loss is then that $\Pi^{\tilde{s}_j > s_j}$ is non-empty. The magnitude of his loss is determined by considering the different outcomes in $\Pi^{\tilde{s}_j > s_j}$. For each $\pi \in \Pi^{\tilde{s}_j > s_j}$, *i* calculates both his *material loss*, which is the amount by which his payoff from π exceeds his maximal payoff in Π^{s_j}, and his *fairness loss*, which tracks the extent to which π is fairer than the fairest outcomes in Π^{s_j}.

While *i*'s material loss is guaranteed to be positive by the definition of $\Pi^{\tilde{s}_j > s_j}$, π may or may not be fairer than the fairest outcomes in Π^{s_j}. If π is not fairer, *i* perceives no fairness loss because he then has no fairness claim to π even though he could have earned more from it than what he can maximally earn given s_j. His loss is then limited to his material loss. In contrast, if π is fairer than the fairest outcomes in Π^{s_j}, *i* suffers a fairness loss because he now has a fairness claim to his additional earnings from π. All in all, *i*'s loss from s_j relative to a particular $\pi \in \Pi^{\tilde{s}_j > s_j}$ is the weighted sum of his material and fairness loss, while his loss from s_j relative to $\tilde{s}_j$ at large is his maximal loss from s_j relative to the outcomes in $\Pi^{\tilde{s}_j > s_j}$.

I now put more formal structure on these ideas. As stated above, Π^{s_j} is the set of attainable outcomes given s_j. The set $\Pi_i^{s_j}$ is the set of payoffs to player *i* contained in Π^{s_j}. Its maximal element is $\overline{\pi}_i^{s_j} = max\,\Pi_i^{s_j}$. Furthermore, $\Pi^{\tilde{s}_j > s_j} = \left\{\pi \in \Pi^{\tilde{s}_j} : \pi_i > \overline{\pi}_i^{s_j}\right\}$ is the set of attainable outcomes given $\tilde{s}_j$ that yield *i* more payoff than what he can maximally earn given s_j. Fairness is measured by a fairness function, isoquants of which are called *fairness curves*:

DEFINITION 1 The *fairness function* $f : \mathbb{R}^2 \to \mathbb{R}$ is given by

$$f(\pi) = \alpha\left(\pi_i + \pi_j\right)/2 + (1-\alpha)\underline{\pi}$$

where $\underline{\pi} = min\left\{\pi_i, \pi_j\right\}$ and $\alpha \in [0,1]$.

For an interpretation of the parameter α, consider the polar cases $\alpha = 1$ and $\alpha = 0$: If $\alpha = 1$, fairness boils down to material efficiency meaning that for any two outcomes π and $\tilde{\pi}$ we have $f(\pi) > f(\tilde{\pi})$ if and only if $\pi_i + \pi_j > \tilde{\pi}_i + \tilde{\pi}_j$. In contrast, if $\alpha = 0$, we have $f(\pi) > f(\tilde{\pi})$ if and only if $\underline{\pi} > \underline{\tilde{\pi}}$. For fairness to increase in this case, the player with less payoff must receive more.[16] In general, the lower α, the smaller (larger) the relative weight attached to efficiency (equality) considerations in fairness assessments. Yet, it is not equality *per se* that enters the fairness function, but the payoff of the less well-off player. It is this payoff that must increase for fairness to increase. If equality as such mattered, we could also reduce the payoff of the better-off player for fairness to increase.

Player *i* derives a fairness loss from s_j relative to $\pi \in \Pi^{\tilde{s}_j > s_j}$ if and only if π lies on a higher fairness curve than the highest fairness curve reached in Π^{s_j}. Moreover, *i*'s fairness loss increases in the extent to which the fairness of π exceeds the maximal fairness in Π^{s_j}. To formalise this idea, let $\overline{f}^{s_j} = max_{\pi \in \Pi^{s_j}} f(\pi)$ be the highest fairness level attained in Π^{s_j}. The fairness gap between π and Π^{s_j} can then be expressed as $f(\pi) - \overline{f}^{s_j}$.[17] This lead to

DEFINITION 2 Player *i*'s *loss* from strategy $s_j \in S_j$ relative to strategy $\tilde{s}_j \in S_j$ is given by

$$l_i(s_j, \tilde{s}_j) = \begin{cases} max_{\pi \in \Pi^{\tilde{s}_j > s_j}} \left[\beta\left(\pi_i - \overline{\pi}_i^{s_j}\right) + (1-\beta) max\left\{ f(\pi) - \overline{f}^{s_j}, 0 \right\} \right] & \text{if } \Pi^{\tilde{s}_j > s_j} \neq \varnothing \\ 0 & \text{otherwise} \end{cases}$$

with $\beta \in [0,1]$. Moreover, *i*'s loss from s_j is given by $l_i(s_j) = max_{\tilde{s}_j \in S_j} l_i(s_j, \tilde{s}_j)$.

Thus, *i* assesses his loss from s_j relative to $\tilde{s}_j$ by considering the set $\Pi^{\tilde{s}_j > s_j}$. For each outcome $\pi \in \Pi^{\tilde{s}_j > s_j}$, he determines the weighted sum of his material loss $\pi_i - \overline{\pi}_i^{s_j} > 0$ and his fairness loss $max\left\{ f(\pi) - \overline{f}^{s_j}, 0 \right\}$. The weighting is provided by the parameter β: If $\beta = 1$, *i*'s loss coincides

[16] The second case is reminiscent of a Rawlsian (or max-min) social welfare function. The first case has a utilitarian flavour.

[17] The two terms can be thought of as the unique payoffs to *i* yielding the fairness levels $f(\pi)$ and $\overline{f}^{s_j}$, respectively, assuming that all other players earn the same. This reading of the fairness gap is invariant to rescalings of the fairness function.

with his material loss. Conversely, if $\beta = 0$, *i* only pays heed to his fairness loss. Player *i*'s loss from s_j relative to $\tilde{s}_j$ is the maximal sum of this kind with respect to all outcomes in $\Pi^{\tilde{s}_j > s_j}$, while his loss from s_j at large is his maximal loss relative to all alternatives in S_j.

I next address player *i*'s *gain* from s_j. Relative to a particular $\tilde{s}_j$, *i* derives a gain only if $\Pi^{s_j > \tilde{s}_j}$, the set of feasible outcomes under s_j that give him more payoff than what he can maximally earn under $\tilde{s}_j$, is non-empty. Regarding the magnitude of his gain, fairness curves again play a central role. Player *i* derives a *fairness gain* from a given $\pi \in \Pi^{s_j > \tilde{s}_j}$ if and only if π lies on a lower fairness curve than the highest curve reached in $\Pi^{\tilde{s}_j}$. The intuition is that if π lay on the same or a higher curve, *i* would consider his material gain from π to be well-deserved for contributing to no decrease in fairness. This would lead him to feel no fairness gain, and his gain would be limited to his material gain. These considerations motivate

DEFINITION 3 Player *i*'s gain from strategy $s_j \in S_j$ relative to strategy $\tilde{s}_j \in S_j$ is given by

$$g_i\left(s_j, \tilde{s}_j\right) = \begin{cases} max_{\pi \in \Pi^{s_j > \tilde{s}_j}} \left[\beta\left(\pi_i - \overline{\pi}_i^{\tilde{s}_j}\right) + \left(1-\beta\right) max\left\{ \overline{f}^{\tilde{s}_j} - f\left(\pi\right), 0 \right\} \right] & \text{if } \Pi^{s_j > \tilde{s}_j} \neq \varnothing \\ 0 & \text{otherwise} \end{cases}$$

with $\beta \in [0,1]$. Moreover, *i*'s gain from s_j is given by $g_i\left(s_j\right) = max_{\tilde{s}_j \in S_j}\, g_i\left(s_j, \tilde{s}_j\right)$.

Thus, to assess his gain from s_j relative to $\tilde{s}_j$, *i* determines for each $\pi \in \Pi^{s_j > \tilde{s}_j}$ the weighted sum of his material gain $\pi_i - \overline{\pi}_i^{\tilde{s}_j} > 0$ and his fairness gain $max\left\{ \overline{f}^{\tilde{s}_j} - f\left(\pi\right), 0 \right\}$. Crucially, for the fairness gain to be positive, π must be *less* fair than what is maximally achievable given $\tilde{s}_j$. If π were more fair, *i* would feel entitled to his material gain and perceive no fairness gain.[18]

I assume that the same parameters α and β are used in the calculation of gains and losses. I allow for asymmetries between the two in defining *net losses*:

DEFINITION 4 Player *i*'s *net loss* from strategy $s_j \in S_j$ is given by

[18] Hence, the expression "fairness gain" does not refer to an increase in fairness, but to a sensation of gain based on fairness considerations.

$$nl_i(s_j) = l_i(s_j) - \gamma g_i(s_j)$$

with $\gamma \in [0,1]$.

The case $\gamma < 1$ allows for the possibility that "losses loom larger than gains", which is a key assumption in the literature on loss aversion (Kahnemann and Tversky, 1979; Köszegi and Rabin, 2006). Yet, as discussed in the Introduction, this chapter is not about loss aversion as understood by that literature. The following is immediate:

LEMMA 1 If $S_j = \{s_j\}$, we have $nl_i(s_j) = 0$.

The lemma addresses the case where *j* is passive. If *j* has only one strategy, $\Pi^{\tilde{s}_j > s_j}$ and $\Pi^{s_j > \tilde{s}_j}$ are empty for all $\tilde{s}_j \in S_j$, which implies $l_i(s_j) = 0$ and $g_i(s_j) = 0$ and therefore $nl_i(s_j) = 0$.

1.3 Reciprocating the Other Player's Strategy

Given the strategy s_j of player *j*, player *i* must choose an outcome from the opportunity set Π^{s_j} created by s_j. In this section, I define a preference for *net-loss reciprocation* to explain this choice. Net-loss reciprocation means that *i*'s willingness to pay for increasing the net loss that he imposes on *j* increases in his own net loss from *j*'s strategy.

I first define the net loss imposed on *j*. Let $\pi^c \in \Pi^{s_j}$ be the outcome chosen by *i* and let $\Pi^{s_j > c} = \{\pi \in \Pi^{s_j} : \pi_j > \pi_j^c\}$ and $\Pi^{s_j < c} = \{\pi \in \Pi^{s_j} : \pi_j < \pi_j^c\}$ be the outcomes in Π^{s_j} yielding *j* more and less payoff than π^c, respectively. This leads to

DEFINITION 5 Player *j*'s *loss* from $\pi^c \in \Pi^{s_j}$ is

$$l_j(\pi^c, \Pi^{s_j}) = \begin{cases} max_{\pi \in \Pi^{s_j > c}} \left[\beta(\pi_j - \pi_j^c) + (1-\beta) max\{f(\pi) - f(\pi^c), 0\} \right] & \text{if } \Pi^{s_j > c} \neq \varnothing \\ 0 & \text{otherwise} \end{cases}$$

Moreover, *j*'s *gain* from π^c is

$$g_j\left(\pi^c,\Pi^{s_j}\right)=\begin{cases} max_{\pi\in\Pi^{s_j<c}}\left[\beta\left(\pi_j^c-\pi_j\right)+\left(1-\beta\right)max\left\{f\left(\pi\right)-f\left(\pi^c\right),0\right\}\right] & \text{if } \Pi^{s_j<c}\neq\varnothing \\ 0 & \text{otherwise} \end{cases}$$

Finally, j's *net loss* from π^c is

$$nl_j\left(\pi^c,\Pi^{s_j}\right)=l_j\left(\pi^c,\Pi^{s_j}\right)-\gamma g_j\left(\pi^c,\Pi^{s_j}\right).$$

The procedure for calculating j's loss and gain from π^c is analogous to calculating i's loss and gain from s_j. In particular, π^c takes the role of Π^{s_j} and the alternative outcomes $\pi\in\Pi^{s_j}$ the roles of the different $\Pi^{\tilde{s}_j}$. I also assume that the same parameters α, β and γ are used.

I now turn to player i's preferences governing his choice from Π^{s_j}:

ASSUMPTION 1 Player i's preferences on the outcomes in Π^{s_j} are represented by

$$u_i\left(\pi,s_j\right)=v\left(\pi_i\right)+r\left(nl_j\left(\pi,\Pi^{s_j}\right),nl_i\left(s_j\right)\right)$$

where $v:\mathbb{R}\to\mathbb{R}$ and $r:\mathbb{R}^2\to\mathbb{R}$ are continuous and $dv/d\pi_i>0$ as well as $\partial^2 r/\partial nl_j\partial nl_i>0$.

Thus, given s_j, i's utility from $\pi\in\Pi^{s_j}$ is additively separable into the utility from his own payoff and a reciprocation term that depends on the net loss that π imposes on j as well as the net loss that i himself derives from s_j. The marginal utility of i's own payoff is always positive. Moreover, $WTP=\partial r/\partial nl_j/dv/d\pi_i$, which is i's willingness to pay for increasing the net loss imposed on j, increases in i's net loss from s_j, where WTP itself may be negative or positive. Whenever $nl_j'>nl_j$, we therefore have $\partial\left[r\left(nl_j',nl_i\right)-r\left(nl_j,nl_i\right)\right]/\partial nl_i>0$, which means that the impact of an increase in i's net loss from s_j is such that for any two outcomes that differ in terms of the net loss that they impose on j the utility advantage (disadvantage) of the outcome imposing the larger net loss becomes larger (smaller).[19]

[19] An example is $u_i=\pi_i-\left(nl_j-nl_i\right)^2$. The partial derivative of $-\left(nl_j'-nl_i\right)^2+\left(nl_j-nl_i\right)^2$ with respect to nl_i is $2\left(nl_j'-nl_i\right)-2\left(nl_j-nl_i\right)$, which is positive if and only if $nl_j'>nl_j$.

Furthermore, I follow McFadden (1974) and McKelvey and Palfrey (1995, 1996) in making

ASSUMPTION 2 The probability of player *i* choosing outcome $\pi \in \Pi^{s_j}$ is given by

$$Pr\left(\pi, s_j\right) = exp\left[u_i\left(\pi, s_j\right)\right] \Big/ \sum_{\breve{\pi} \in \Pi^{s_j}} exp\left[u_i\left(\breve{\pi}, s_j\right)\right].$$

As explained by Goeree et al (2008), this quantal response structure can be justified by disturbances on individual decision making reflecting the effects of unobservables such as mood or perceptual variations. According to this interpretation, $u_i\left(\pi, s_j\right)$ corresponds to the average utility attached to $\pi \in \Pi^{s_j}$ with each player realising a mean-zero perturbation of $u_i\left(\pi, s_j\right)$.[20] Perturbations are assumed i.i.d. across outcomes and players.[21] Assumption 2 implies that every player chooses every available outcome with positive probability. This helps in interpreting experimental data, where typically not all subjects facing a given opportunity set $\hat{\Pi}$ choose the same outcome. In this context, a forgone-option effect refers to a statistically significant shift in the empirical choice distribution on $\hat{\Pi}$ for two strategies s_j and s_j' satisfying $\Pi^{s_j} = \Pi^{s_j'} = \hat{\Pi}$. Using Assumption 2, we can explain this shift if we can show a corresponding shift in the theoretical choice distributions. I draw extensively on this method in the following section.

We conclude this section by the following lemma, which is useful in what follows:

LEMMA 2 Consider any pair of strategies $s_j, s_j' \in S_j$ with $\Pi^{s_j} = \Pi^{s_j'} = \hat{\Pi}$ where $\hat{\Pi}$ is a fixed set of outcomes and let $Pr\left(\pi\right)$ and $Pr'\left(\pi\right)$ denote the probabilities of player *i* choosing a given $\pi \in \hat{\Pi}$ when faced with s_j and s_j', respectively. Suppose that $nl_i\left(s_j\right) = nl_i\left(s_j'\right)$. We then have

$$Pr\left(\pi\right) = Pr'\left(\pi\right)$$

---

[20] In this setup, the choice of scale for utilities is not without loss of generality. In particular, multiplication of all utilities by some constant $c > 1$ makes players likelier to choose the options yielding them the highest utility. For this reason, quantal response models contain a scaling parameter λ intended to capture the degree of players' "rationality", i.e., their likelihood of choosing their most preferred options. Since the level of λ does not affect the conclusions drawn below, I set λ to 1.

[21] Moreover, in order to generate the assumed logit structure, the perturbations must take a particular stochastic form. See Goeree et al (2008) for details.

for all $\pi \in \hat{\Pi}$. Suppose next that $nl_i\left(s'_j\right) > nl_i\left(s_j\right)$. We then have

$$Pr'\left(\pi'\right)/Pr\left(\pi'\right) > Pr'\left(\pi\right)/Pr\left(\pi\right)$$

for all $\pi, \pi' \in \hat{\Pi}$ such that $nl_j\left(\pi', \hat{\Pi}\right) > nl_j\left(\pi, \hat{\Pi}\right)$.

Thus, if *i* is faced with two strategies creating the same opportunity set $\hat{\Pi}$ for him and from which he derives the same net loss, his probability of choosing any given outcome in $\hat{\Pi}$ is the same across the two situations. If he does not derive the same net loss, his choice probability possesses the *monotone likelihood ratio property* for outcomes that can be ordered according to the net loss imposed on *j*. For any such pair, there is a relative shift in probability mass towards the outcome imposing the higher net loss in the situation where *i* derives the higher net loss.

1.4 Applications

In this section, I show that net loss reciprocation can explain the forgone-option effects that have been documented in a number of experimental studies. Each piece of evidence considered lends additional structure to the model: Firstly, the evidence from *The Hidden Cost of Control* suggests that material factors are not irrelevant in the calculation of losses and gains ($\beta > 0$). Secondly, the evidence from *Trust* implies that gains are not fully discounted when calculating net losses ($\gamma > 0$). Thirdly, the evidence from *Ultimatum Bargaining* rules out a purely efficiency-oriented notion of fairness and provides an upper bound for the importance of material factors ($\alpha < \left(1-2\beta\right)/\left(1-\beta\right)$ and $\beta < 0.5$). Finally, the evidence from *Lost Wallets* pinpoints the fairness parameter α to equal $\left(1-2\beta\right)/\left(2-2\beta\right)$.

Discussing each piece of evidence in turn, I also address the problems faced by the reciprocity models of Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006) in accounting for these experimental results. Recall that both models make use of players' second-order beliefs, i.e., their beliefs about the other player's belief about their own choice of strategy. As the studies considered in this section do not measure these beliefs, I use the actual behaviour of experimental subjects as a "stand-in" for second-order beliefs. This is in keeping with the equilibrium spirit of these models, which requires beliefs to coincide with actual behaviour.

Although both models rely on second-order beliefs, they differ in how they define the kindness of the other player's strategy. Suppose that we want to evaluate the kindness of s_j to player *i*. To do so, Dufwenberg and Kirchsteiger (2004) compare *i*'s expected payoff from s_j given his second-order belief to what he could have minimally and maximally earned under the alternatives given again his second-order belief. The strategy is viewed as kind (unkind) if it gives *i* more (less) than half of what he could have minimally and maximally earned. In contrast, Falk and Fischbacher (2006) focus on the expected outcome implemented by s_j given *i*'s second-order belief. If *i* earns more (less) than *j* under this outcome, s_j is viewed as kind (unkind). The model also takes into account the alternatives available to *j*. For instance, if the outcome implemented by s_j puts *i* ahead, this is only viewed as fully kind if there are alternatives under which *i* would have earned less. The intuition is that the kindness embodied by s_j cannot be regarded as fully intentional otherwise. Appendix B contains a more detailed exposition of the two models.

1.4.1 The Hidden Cost of Control

Falk and Kosfeld (2006) study the reaction of agents (*A*) to a principal's (*P*) decision to control their choice of how much productive effort to exert. Two of their treatments are directly relevant for us: In the control game (*CG*), the principal first decides whether to control the agent or not. If not controlled by the principal, the agent can exert any effort level $e \in \{0,1,2,...,120\}$. Payoffs are given by $2e$ for the principal and $120-e$ for the agent. In contrast, if the principal has controlled the agent, effort is restricted to be at least ten, i.e., $e \in \{10,11,...,120\}$, the mapping from effort to payoffs being the same as after no control. The second treatment is a dictator game (*DG*) that is identical to the subgame of *CG* after control. As a result, the agent (who takes the role of player *i*) chooses from $\{(120-e,2e) : e \in \{10,11,...,120\}\}$ both in *DG* and after control in *CG*.[22] We have

[22] Recall that outcomes have the format $\pi = (\pi_i, \pi_j)$.

PROPOSITION 1 Suppose that $\beta > 0$. The net loss that the agent derives from the principal is given by $nl_A^{CG} = 10\beta$ after control in *CG* and by $nl_A^{DG} = 0 < nl_A^{CG}$ in *DG*. The net loss that the agent imposes on the principal decreases in effort. Hence, average effort is higher in *DG*.

The agent derives a net loss of zero in *DG* because the principal is passive in this treatment. After control in *CG*, the agent only derives a material loss (of 10) because control only rules out outcomes that are less fair than the outcomes attainable under control for being less efficient and containing a smaller minimal payoff. Hence, for the agent to derive a loss from control, the weight on his material loss must not be zero ($\beta > 0$). Regarding the net loss that the agent imposes on the principal, a one-unit increase in effort decreases the principal's material loss and increases her material gain. Fairness losses and gains do not counteract this: The principal's fairness loss is non-increasing in effort. In fact, it decreases except for high effort, where the principal may not feel entitled to additional effort if her fairness notion leans toward a concern for the less well-off. Likewise, the principal's fairness gain is non-decreasing in effort. It is zero for low effort levels, where the principal feels entitled to effort, but may increase for higher effort. As a result, as long as $\beta > 0$, the principal's net loss unambiguously decreases in effort.

Lemma 2 then implies that the effort distributions in the two situations possess the monotone likelihood ratio property, with probability mass shifting towards higher effort in the situation where agents derive the lower net loss. Consequently, average effort is predicted to be higher in *DG*. This matches the results of Falk and Kosfeld (2006), who report that average effort is significantly lower after control in *CG*. The agents in their experiment provide a mean effort of 17.5 after control in *CG*, but of 28.7 in *DG*.[23]

The reciprocity model of Dufwenberg and Kirchsteiger (2004), referred to as DK in what follows, struggles to account for these findings. Given that agents in *CG* exert more effort after no control than after control and exerting more effort means less payoff for agents themselves,[24] DK view control (somewhat counter-intuitively) as kind to agents. Consequently, they predict more effort, which is kinder to principals, after control in *CG* than in *DG*, where the principal is

---

[23] Falk and Kosfeld (2006) also study other control levels. However, they implement no corresponding dictator games, which means that no foregone-option effects can be studied.

[24] Falk and Kosfeld (2006) report a mean effort after no control of 23.0.

passive and hence neither kind nor unkind. Reliance on second-order beliefs has a perverse consequence here. From an intuitive viewpoint, it is clear agents are disgruntled at being controlled if the principal has had the choice of not controlling them. The effect of this disgruntlement, namely, that agents keep more money for themselves, is used by DK as "evidence" for the conclusion that they have no reason for being disgruntled. This confuses cause and effect of agents' emotional response to the principal's behaviour. My approach avoids this problem because it relies on agents' opportunity sets given the principal's strategy without taking into account their reaction to the latter.

The reciprocity model of Falk and Fischbacher (2006), denoted FF in what follows, runs into similar problems. Given that responders exert less effort after control in *CG* than in *DG*, agents view themselves as being put further ahead of principals in *CG*. Moreover, principals could have put agents less ahead in *CG* by not controlling them. As a result, control in *CG* is kinder, which is at odds with the fact that agents are kinder to principals in *DG* by exerting more effort.

1.4.2 Trust

McCabe et al (2003) study a simple trust game (*TG*) in which the first mover (*FM*) can either implement the no-trust outcome $(20,20)$ or trust the second mover (*SM*). If trusted by the first mover, the second mover can choose between $(30,15)$ and $(25,25)$. The first entry in each payoff vector denotes the payoff to the second mover, who takes the role of player i in what follows. The authors compare second mover behaviour in this game to behaviour in a dictator game (*DG*) in which the first mover is passive and the second mover has a choice between the same two outcomes as after trust in *TG*. We have

PROPOSITION 2 Suppose that $\gamma > 0$ and either $\alpha < 2/3$ or $\beta > 0$. The net loss that the second mover derives from trust in *TG* is $nl_{SM}^{TG} = -\gamma\left(\beta 10 + (1-\beta) max\{5-7.5\alpha, 0\}\right)$, while his net loss in *DG* is $nl_{SM}^{DG} = 0 > nl_{SM}^{TG}$. The net loss that the second mover imposes on the first mover is $nl_{FM} = -10\gamma\beta$ if the second mover chooses $(25,25)$ and $nl'_{FM} = 10 - 7.5\alpha + 7.5\alpha\beta > nl_{FM}$ if he chooses $(30,15)$. Thus, second movers are more likely to choose $(25,25)$ after trust in *TG*.

We have $nl_{SM}^{DG} = 0$ because the first mover is passive in *DG*. As for *TG*, for the second mover's net loss from trust to be negative, we must have $\gamma > 0$ and either $\alpha < 2/3$ or $\beta > 0$. On the one hand, the gain from trust must not be fully discounted, which is guaranteed by $\gamma > 0$. Further, there are two ways for the gain from trust to be positive: Either we have $\beta > 0$, which ensures that the second mover puts positive weight on his material gain of 10. Or we have $\alpha < 2/3$, which means that fairness does not lean too much towards efficiency. In this case, the second mover perceives $(30,15)$, which he can implement after trust, as less fair than the no-trust outcome $(20,20)$ because of the smaller minimal payoff. He then derives a fairness gain from trust, which makes his overall gain positive irrespective of β.

Furthermore, the second mover clearly imposes a lower net loss on the first mover by choosing $(25,25)$. Net-loss reciprocation implies that the second mover is more likely to choose $(25,25)$ instead of $(30,15)$ in *TG* because of his lower derived net loss This forgone-option effect is confirmed by McCabe et al (2003), who report that second movers choose $(25,25)$ with a frequency of 0.65 after trust in *TG* and of only 0.33 in *DG*, this difference being significant.[25]

1.4.3 Ultimatum Bargaining

Falk et al (2003) study four binary ultimatum games in each of which the offer “2 for the responder, 8 for the proposer” is available to the proposer. The games differ regarding the second offer. In three, there exists a true alternative, namely, “5 for both”, “8 for the responder, 2 for the proposer” and “0 for the responder, 10 for the proposer”, respectively. In the fourth, the proposer is effectively passive because the alternative offer is also “2 for the responder, 8 for the proposer”. Thus, letting the responder take the role of player *i*, acceptance of the alternative implements $(5,5)$, $(8,2)$, $(0,10)$ and $(2,8)$, respectively. In what follows, I refer to the four treatments by these outcomes. We have

---

[25] The approaches of Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006) can also account for this. Under the former, trust gives the second mover more in expected terms than no trust, which leads the second mover to view trust as kind, giving him an incentive to be kind to the first mover by repaying trust. Under the latter, given that trust is not always reciprocated, trust puts the second mover ahead of the first mover in expected terms. Moreover, the second mover would have earned less had he not been trusted. As a result, the kindness embodied by trust is fully intentional.

PROPOSITION 3 Suppose that $\alpha < (1-2\beta)/(1-\beta)$, $0<\beta<0.5$ and $\gamma>0$. The net losses that the responder derives from "2 for the responder, 8 for the proposer" in the four treatments are given by $nl_R^{(5,5)} = 3-3\alpha+3\alpha\beta$, $nl_R^{(8,2)} = 6\beta$, $nl_R^{(2,8)} = 0$ and $nl_R^{(0,10)} = -2\gamma\beta$, respectively, where we have $nl_R^{(5,5)} > nl_R^{(8,2)} > nl_R^{(2,8)} > nl_R^{(0,10)}$. The net loss that the responder imposes on the proposer by accepting "2 for the proposer, 8 for the responder" is $nl_P = -8\gamma\beta$, while rejection imposes $nl'_p = 2+3\alpha+6\beta-3\alpha\beta > nl_p$. We therefore have $Pr^{(5,5)} > Pr^{(8,2)} > Pr^{(2,8)} > Pr^{(0,10)}$, where Pr^x denotes the rejection probability in treatment $x \in \{(5,5),(8,2),(2,8),(0,10)\}$.

Crucially, Proposition 3 asserts that losses do not track material losses. The responder derives the highest overall loss in treatment $(5,5)$, whereas his material loss is largest in $(8,2)$. To see this, notice that "2 for the responder, 8 for the proposer" means a material loss of 6 for the responder if the alternative has been "8 for the responder, 2 for the proposer", but of only 3 if it has been "5 for both". The responder's fairness loss is $3-3\alpha$ in $(5,5)$, but is necessarily zero in $(8,2)$ because the outcome $(8,2)$ is always as fair as the outcome $(2,8)$ irrespective of α. The dissociation of losses from material losses is achieved by the fairness loss in $(5,5)$ being sufficiently large and fairness losses playing a sufficiently large role. Formally, $nl_R^{(5,5)} > nl_R^{(8,2)}$ is equivalent to $3-3\alpha+3\alpha\beta > 6\beta \Leftrightarrow \alpha < (1-2\beta)/(1-\beta)$. This can only be satisfied by $\alpha \geq 0$ if we have $1-2\beta>0 \Leftrightarrow \beta<0.5$.

Further, we have $nl_R^{(2,8)} = 0$ because the proposer is passive in $(2,8)$ and $nl_R^{(0,10)} = -2\gamma\beta$ because the responder derives no fairness gain in $(0,10)$. The reason is that he feels entitled to "2 for the responder, 8 for the proposer" if the alternative is "0 for the responder, 10 for the proposer". We have $6\beta > 0 > -2\gamma\beta$ because of $\beta>0$ and $\gamma>0$.

All in all, since rejection of the fixed offer "2 for the responder, 8 for the proposer" imposes a higher net loss on the proposer than acceptance, the responder is most likely to reject in $(5,5)$, second most likely in $(8,2)$ etc. This is largely consistent with the results of Falk et al (2003), who report the following rejection frequencies:

$$Pr^{(5,5)} = 0.44 > Pr^{(8,2)} = 0.27 > Pr^{(2,8)} = 0.18 > Pr^{(0,10)} = 0.09\,,$$

these differences being statistically significant except for the last one. Although the last difference has the right sign, my approach faces a difficulty here. The absence of a significant difference would be explained by net-loss reciprocation if $nl_R^{(2,8)} = nl_R^{(0,10)}$. Since we have $nl_R^{(2,8)} = 0$ because the proposer is passive in this treatment, whereas $nl_R^{(0,10)} = -2\gamma\beta$, we would have to assume either $\beta = 0$ or $\gamma = 0$ or both. Alternatively, if both γ and β are positive but small, their product can be viewed as approximately zero. I return to this issue below.

In contrast, DK face difficulties in rationalising the difference between the treatments $(5,5)$ and $(8,2)$. According to DK, the status quo offer "2 for the responder, 8 for the proposer" is *less* kind in $(8,2)$ than $(5,5)$ for the following reason: In $(5,5)$, the status quo is accepted with a probability of 0.66 and the alternative for sure, which makes for a kindness of the status quo of $k = 0.66 \cdot 2 - 0.5[5 + 0.66 \cdot 2] = -1.84$.[26] In $(8,2)$, the acceptance probabilities are 0.73 for the status quo and 0.98 for the alternative. The fact that the status quo is accepted with a higher probability tends to make the status quo kinder. However, what works in the opposite direction is that responders could have earned more under the alternative. This second effect dominates since we have $k = 0.73 \cdot 2 - 0.5[0.98 \cdot 8 + 0.73 \cdot 2] = -3.72 < -1.84$. What DK do not take into account is responders' sense of entitlement, in particular, that they feel less entitled to their forgone earnings in $(8,2)$ because these forgone earnings derive from an outcome that is only as fair as status quo. Such considerations are at the heart of my approach.

The basic model in FF also fails to explain the difference in responder behaviour between $(5,5)$ and $(8,2)$. The reason is the binary nature of the intention factor (see Appendix B), which simply asks whether or not responders could have earned more in expected terms than under the status quo, which is the case in both treatments. Hence, the intention factor does not capture the fairness difference between the two alternatives and the differential sense of entitlement that this difference creates. The appendix in FF contains a richer version of their model designed to address this problem.

1.4.4 Lost Wallets

Dufwenberg and Gneezy (2000) fail to establish forgone-option effects in a series of trust games termed "Lost Wallet Games". Their common feature is that the second mover can split 20 units of payoff between himself and the first mover in the event of trust. The games differ regarding

[26] See Appendix B for a detailed exposition of the kindness function *k*.

the no-trust outcome, which, letting the second mover take the role of player *i*, is given by $(0, f)$ with $f \in \{4,7,10,13,16\}$. That is, the no-trust payoff for the second mover is always zero, while the games differ with respect to the no-trust payoff for the first mover. For simplicity, I focus on the polar cases $f = 4$ and $f = 16$ because the absence of a forgone-option effect is most puzzling between them. I refer to the two treatments by the respective no-trust outcome, namely $(0,4)$ and $(0,16)$. Servatka and Vadovic (2009) draw on the basic setup of Dufwenberg and Gneezy (2000), while varying the inequality of the no-trust outcome. In their two treatments, the no-trust outcomes are given by $(0,10)$ and $(5,5)$. Like Dufwenberg and Gneezy (2000), they fail to establish a significant difference in return transfers, i.e., there is again no forgone-option effect. Consistent with these empirical findings, we have

PROPOSITION 4 Suppose that we have either $\gamma > 0$, $0 < \beta \leq 0.5$ and $\alpha = (1-2\beta)/(2-2\beta)$ or $\gamma > 0$, $\beta = 0$ and $\alpha \geq 0.5$ or $\gamma = 0$ The net loss that the second mover derives from the first mover's trust is $nl_{SM}^{(0,4)} = -20\gamma\beta$ if the no-trust outcome is $(0,4)$ and $nl_{SM}^{(0,16)} = -20\gamma\beta = nl_{SM}^{(0,4)}$ if it is $(0,16)$. As a result, the repayment distributions do not differ between the two situations. If the no-trust outcome is $(0,10)$, the net loss derived from trust is $nl_{SM}^{(0,10)} = -20\gamma\beta$, while it is $nl_{SM}^{(5,5)} = -\gamma\left(15\beta + (1-\beta)\left(max\{5-10\alpha, 0\}\right)\right) = nl_{SM}^{(0,10)}$ if the no-trust outcome is $(5,5)$. Again, the repayment distributions do not differ between the two situations.

In the first two cases, the second mover's gain from trust is limited to his material gain of 20 because the no-trust outcomes $(0,4)$ and $(0,16)$ are not fairer than any outcome in the second mover's opportunity set after trust. Representing a payoff sum of less than 20, both no-trust outcomes are less efficient than the outcomes available after trust. Moreover, the minimal payoff is zero in each case, which is also the minimal payoff available after trust (if the second mover shares nothing). As a result, the responder feels entitled to his material gain causing his gain to be limited to the latter. As there is no loss from trust, the net loss is $-20\gamma\beta$ in each case.

For the same reasons, the second mover's net loss is $-20\gamma\beta$ if the no-trust outcome is $(0,10)$. If it is $(5,5)$, the second mover's material gain from trust is 15. For $nl_{SM}^{(5,5)} = -20\gamma\beta$ to hold, we can impose $\gamma = 0$ meaning that gains are fully discounted. Alternatively, if $\gamma > 0$, we can let $\beta = 0$ and $\alpha \geq 0.5$ meaning that the weight on material gains is zero, but the second mover derives no fairness gain from trust. Indeed, if $\alpha \geq 0.5$, the least fair outcome after trust, namely,

$(20,0)$, is at least as fair as $(5,5)$ because efficiency receives sufficient weight in the fairness function. Finally, if $\gamma > 0$ and $\beta > 0$, the second mover must derive a positive fairness gain from trust in $(5,5)$ to offset his larger material gain in $(0,10)$. This is the case if $\alpha < 0.5$ because fairness then leans towards a concern for the less well-off. In these conditions, we have $nl_{SM}^{(5,5)} = -\gamma(5 - 10\alpha + 10\beta + 10\alpha\beta)$, which equals $-20\gamma\beta$ if and only if $\alpha = (1-2\beta)/(2-2\beta)$. This equality can only be satisfied by $\alpha \geq 0$ if $\beta \leq 0.5$. Also, given $\beta > 0$, $\alpha < 0.5$ as assumed.

Dufwenberg and Gneezy (2000) also implement a dictator treatment (*DG*) in which dictators face the same opportunity set as second movers after trust. The authors report no significant difference in transfers between *DG* on the one hand and the treatments $(0,4)$ and $(0,16)$ on the other. For net-loss reciprocation to explain this, we must have $-20\gamma\beta = nl_{SM}^{DG} = 0$. We can make the equality hold by imposing either $\beta = 0$ or $\gamma = 0$. I return to this issue below.

Regarding the intention-based models, DK cannot account for there being no difference between $(0,10)$ and $(5,5)$. The problem is that DK pick up on second movers' differential earnings from the no-trust outcome. Given that second movers behave in the same way after trust, this yields the conclusion that trust is kinder in $(0,10)$ because second movers gain more from it in expected terms. My approach can navigate around this problem because second movers' higher material gain in $(0,10)$ can be offset by a higher fairness gain in $(5,5)$.

In contrast, FF can explain most absences of treatment differences. Given that average second-mover behaviour is the same, second movers view first movers as intending the same expected outcome (putting them ahead of first movers) in all treatments. Moreover, first movers could have treated second movers worse by not trusting them (except in *DG*). Hence, first movers are equally kind to second movers in all treatments rationalising the absence of a forgone-option effect. FF (like my approach) only struggle to explain behaviour in *DG* relative to the other treatments because first movers are passive there, which should cause second movers to share less.

1.5 Discussion

The examples considered in the preceding section are instructive with regard to the calibration of the model. The preferred specification is

$0 < \gamma \leq 1$, $0 < \beta < 0.5$ and $\alpha = (1-2\beta)/(2-2\beta)$

which is well-supported by the experimental data considered in this chapter. The interpretation is that gains are not fully discounted in the calculation of net losses ($0 < \gamma \leq 1$) and that the weight on material losses and gains is neither zero nor too large ($0 < \beta < 0.5$). Given the restrictions on β, the condition on α implies $0 < \alpha < 0.5$ meaning that fairness leans towards a concern for the less well-off.

Imposing $\beta > 0$ and $\gamma > 0$ fails to explain two pieces of evidence, namely, the treatment $(0,10)$ from *Ultimatum Bargaining* and *DG* from *Lost Wallets*. To account for them, the above specification could be modified by setting $\gamma = 0$. This parameterisation, which implies that gains are fully discounted, can account for all the evidence except that from *Trust*. In particular, the evidence from *Lost Wallets* is explained almost trivially by reducing net losses to zero in all treatments. Effectively, this specification negates the importance of positive reciprocity by asserting that people do not react to gains that they derive from others. Charness and Rabin (2002) provide further evidence that positive reciprocity is a less important motivational force than negative reciprocity.[27]

All in all, this section has demonstrated that net-loss reciprocation in conjunction with the method for calculating net losses developed in this chapter can by and large account for the existence or absence of forgone-option effects in a number of experimental studies. I have also shown that existing models of intention-based reciprocity face problems in explaining this evidence comprehensively. This is particularly true for the model of Dufwenberg and Kirchsteiger (2004), while at least the extended version of Falk and Fischbacher (2006) performs relatively well. Yet, even in its extended form, the latter only captures players' sense of entitlement in an approximate, qualitative fashion. My approach allows to precisely quantify this sense via fairness losses and gains.

[27] In applications of loss aversion, it is often assumed for the sake of simplicity that only losses count. In Chapter 3, I propose a model of belief choice under regret that also draws on losses only.

1.6 Conclusion

This chapter presents a qualitative preference model for two-player interactions building on the idea of net-loss reciprocation. Net-loss reciprocation asserts that a player's willingness to impose net losses on the other increases in the net loss that he derives from the other player's strategy. The chapter shows that net-loss reciprocation can account for various forgone-option effects (or absences thereof) that arise in a number of experimental studies.

The main difficulty faced by net-loss reciprocation relates to the status of positive reciprocity. In the light of the evidence considered in this chapter, it is not clear whether players fully discount any gain they derive from the other player's strategy or whether they take this gains into account leading them to become less willing to impose a net loss on the other. Apart from this, I find conclusive evidence that both material and fairness considerations matter to the determination of net losses, with fairness being somewhat more important. I also establish that a regard for the less well-off as opposed to a pure concern for material efficiency plays an important role in fairness assessments.

Given the relative success of my approach in explaining forgone-option effects when compared to existing models of intention-based reciprocity, the development of full-fledged utility models incorporating net-loss reciprocation seems worthwhile. These models could be used to analyse more general classes of games.[28] An advantage of such models compared to intention-based models is their direct testability using standard experimental data as they do not rely on higher-order beliefs.

[28] In Chapter 2 of this dissertation, an extension of the two-player model developed in this chapter to more players (including Nature) is proposed.

1.7 Appendix A: Proofs

PROOF OF LEMMA 2

Let $nl_i\left(s'_j\right)=nl'_i$ and $nl_i\left(s_j\right)=nl_i$. Also, let $nl_j\left(\pi',\hat{\Pi}\right)=nl'_j$ and $nl_j\left(\pi,\hat{\Pi}\right)=nl_j$. The first part, where $nl_i=nl'_i$, is immediate since we have

$$Pr\left(\pi\right)=exp\left[v\left(\pi_i\right)+r\left(nl_j,nl_i\right)\right]\Big/\sum_{\breve{\pi}\in\hat{\Pi}}exp\left[v\left(\breve{\pi}_i\right)+r\left(n\breve{l}_j,nl_i\right)\right]$$

and

$$Pr'\left(\pi\right)=exp\left[v\left(\pi_i\right)+r\left(nl_j,nl'_i\right)\right]\Big/\sum_{\breve{\pi}\in\hat{\Pi}}exp\left[v\left(\breve{\pi}_i\right)+r\left(n\breve{l}_j,nl'_i\right)\right]$$

for all $\pi\in\hat{\Pi}$ by Assumptions 1 and 2.

Next, I show that $Pr'\left(\pi'\right)/Pr\left(\pi'\right)>Pr'\left(\pi\right)/Pr\left(\pi\right)\Leftrightarrow Pr'\left(\pi'\right)/Pr'\left(\pi\right)>Pr\left(\pi'\right)/Pr\left(\pi\right)$ for all $\pi,\pi'\in\hat{\Pi}$ if $nl'_i>nl_i$ and $nl'_j>nl_j$. By Assumption 2, we can express the second inequality as

$$\frac{exp\left[v\left(\pi'\right)\right]\cdot exp\left[r\left(nl'_j,nl'_i\right)\right]}{exp\left[v\left(\pi\right)\right]\cdot exp\left[r\left(nl_j,nl'_i\right)\right]}>\frac{exp\left[v\left(\pi'\right)\right]\cdot exp\left[r\left(nl'_j,nl_i\right)\right]}{exp\left[v\left(\pi\right)\right]\cdot exp\left[r\left(nl_j,nl_i\right)\right]}\Leftrightarrow$$

$$\frac{exp\left[r\left(nl'_j,nl'_i\right)\right]}{exp\left[r\left(nl_j,nl'_i\right)\right]}>\frac{exp\left[r\left(nl'_j,nl_i\right)\right]}{exp\left[r\left(nl_j,nl_i\right)\right]}.$$

Logarithmation of both sides yields

$$r\left(nl'_j,nl'_i\right)-r\left(nl_j,nl'_i\right)>r\left(nl'_j,nl_i\right)-r\left(nl_j,nl_i\right),$$

which holds by our assumptions on $r\left(\cdot\right)$.■

PROOF OF PROPOSITION 1

The principal is passive in *DG*. By Lemma 1, we have $nl_A^{DG}=0$. In *CG*, letting *NC* denote no control, we have $\Pi^{NC>C} = \{(120-e,2e): e \in \{0,1,...,9\}\}$. Thus, if not controlled, the agent can earn more than what he can maximally earn if controlled. As a result, $g_A^{CG} = 0$ and

$$l_A^{CG} = max_{\pi \in \Pi^{NC>C}} \left[\beta(\pi_A - 110) + (1-\beta) max\left\{ f(\pi) - \bar{f}^C, 0 \right\} \right],$$

where $\bar{f}^C$ is the highest fairness level attained in Π^C. However, the fairness of the outcomes in $\Pi^{NC>C}$ is below $\bar{f}^C$ because both efficiency and a concern for the less well off mandate increasing effort to 40. Consequently, we have $l_A^{CG} = \beta 10 = nl_A^{CG} > 0$.

Next, I show that the net loss that the agent imposes on the principal decreases in *e*. Suppose that the agent chooses $e = x \in \{10,...,119\}$. We then have $\pi^c = (120-x,2x)$,

$$l_P = max_{\pi \in \Pi^{C>c}} \left[\beta(\pi_P - 2x) + (1-\beta) max\left\{ f(\pi) - f(\pi^c), 0 \right\} \right] \text{ and}$$

$$g_P = max_{\pi \in \Pi^{C<c}} \left[\beta(2x - \pi_P) + (1-\beta) max\left\{ f(\pi) - f(\pi^c), 0 \right\} \right]$$

where $\Pi^{C>c} = \{(120-e,2e): e \in \{x+1,...,120\}\}$ and $\Pi^{C<c} = \{(120-e,2e): e \in \{10,...,x-1\}\}$. If effort increases by one unit, i.e., if $e = y = x+1$, we have $\pi^{c'} = (120-x-1,2x+2)$ implying

$$l'_P = max_{\pi \in \Pi^{C>c'}} \left[\beta(\pi_P - 2x - 2) + (1-\beta) max\left\{ f(\pi) - f(\pi^{c'}), 0 \right\} \right] \text{ and}$$

$$g'_P = max_{\pi \in \Pi^{C<c'}} \left[\beta(2x + 2 - \pi_P) + (1-\beta) max\left\{ f(\pi) - f(\pi^{c'}), 0 \right\} \right]$$

where $\Pi^{C>c'} = \{(120-e,2e): e \in \{x+2,...,120\}\}$ and $\Pi^{C<c'} = \{(120-e,2e): e \in \{10,...,x\}\}$.

Suppose first that $x < 40$. We have $f(\pi^{c'}) > f(\pi^c)$ because both efficiency and a concern for the less well off point towards increasing effort. As a result, $l_P > l'_P$ because the maximisation for determining l'_P takes place on the set $\Pi^{C>c'}$, which is a subset of the set $\Pi^{C>c}$ used for establishing l_P and $\beta > 0$. As for gains, the outcomes in $\Pi^{C<c}$ are less fair than π^c because they represent effort further away from 40. The same holds for $\Pi^{C<c'}$ and $\pi^{c'}$. As a result, gains are limited to material gains, and we have $g'_P > g_P$ because $\Pi^{C<c'}$ is a super-set of $\Pi^{C<c}$ and $\beta > 0$. All in all, $e = y$ imposes a smaller net loss on the principal than $e = x$.

Suppose next that $x \geq 40$. Regarding losses, if $f\left(\pi^{c'}\right) \geq f\left(\pi^{c}\right)$ because the fairness function leans towards efficiency, we have $l_P > l'_P$ for the same reasons as above. If $f\left(\pi^{c'}\right) < f\left(\pi^{c}\right)$ because the fairness function leans towards a concern for the less well off, the linear nature of the fairness function implies that π^c is fairer than all elements in $\Pi^{C>c}$ and likewise for $\pi^{c'}$ and $\Pi^{C>c'}$. As a result, losses are limited to material losses and we have $l_P > l'_P$ because $\Pi^{C>c}$ is a super-set of $\Pi^{C>c'}$ and $\beta > 0$. As for gains, if $f\left(\pi^{c'}\right) \geq f\left(\pi^{c}\right)$, all outcomes in $\Pi^{C<c}$ are not fairer than π^c and likewise for $\Pi^{C<c'}$ and $\pi^{c'}$, which implies that there are only material gains. We have $g'_P > g_P$ because $\Pi^{C<c'}$ is a super-set of $\Pi^{C<c}$ and $\beta > 0$. For the same reason, we have $g'_P > g_P$ if $f\left(\pi^{c'}\right) < f\left(\pi^{c}\right)$. Again, $e = y$ imposes a smaller net loss on the principal.

Since the principal's net loss decreases in effort, Lemma 2 together with $nl_A^{CG} > nl_A^{DG}$ implies that the effort distribution in *DG* first-order stochastically dominates the distribution after control in *CG*, which implies that average effort is higher in *DG*.■

PROOF OF PROPOSITION 2

Since first movers are passive in *DG*, we have $nl_{SM}^{DG} = 0$. In *TG*, second movers gain from trust in material terms, which implies their loss is zero. Their gain is given by

$$g_{SM}^{TG} = \beta 10 + (1-\beta)\left(max\{20 - \alpha 22.5 - (1-\alpha)15, 0\}\right) = \beta 10 + (1-\beta)max\{5 - 7.5\alpha, 0\}$$

because $(30,15)$ corresponds to a higher material and fairness gain than $(25,25)$. We have $nl_{SM}^{TG} = -\gamma\left(\beta 10 + (1-\beta)max\{5 - 7.5\alpha, 0\}\right) < 0$ because of our parameter assumptions.

I next show that the net loss imposed on first movers through $(30,15)$ exceeds that imposed through $(25,25)$. First movers derive no loss from $(25,25)$ and a material gain of 10. They derive no fairness gain because $(25,25)$ is superior from the viewpoint of both efficiency and a concern for the less well off. As a result, $nl_{FM} = -10\gamma\beta$. From $(30,15)$, first movers derive a loss of

$$\beta 10 + (1-\beta)(25 - \alpha 22.5 - (1-\alpha)15) = \beta 10 + (1-\beta)(10 - 7.5\alpha) = 10 - 7.5\alpha + 7.5\alpha\beta$$

and no gain, which implies that $nl'_{FM} = 10 - 7.5\alpha + 7.5\alpha\beta > nl_{FM}$. Lemma 2 together with

$nl_{SM}^{DG} > nl_{SM}^{TG}$ then implies that second movers are more likely to choose $(25,25)$ in *TG*.■

PROOF OF PROPOSITION 3

I refer to the offer "2 for the responder, 8 for the proposer" as *X* and the alternative offer in a given treatment as *Y*. We have $\Pi^X = \{(2,8),(0,0)\}$. In $(5,5)$, $\Pi^{Y>X} = \{(5,5)\}$. As a result, responders derive no gain from *X* in this case and a material loss of 3. The highest fairness level reached in Π^X is $f = \alpha 5 + (1-\alpha)2$, whereas $f((5,5)) = 5$. As a result,

$$nl_R^{(5,5)} = \beta 3 + (1-\beta)(5 - \alpha 5 - (1-\alpha)2) = 3 - 3\alpha + 3\alpha\beta .$$

In $(8,2)$, we have $\Pi^{Y>X} = \{(8,2)\}$, which implies a material loss from *X* of 6 and no gain. Responders derive no fairness loss because $(8,2)$ and $(2,8)$ lie on the same fairness curve irrespective of α. Consequently, $nl_R^{(8,2)} = 6\beta$. Since proposers are passive in $(2,8)$, $nl_R^{(2,8)} = 0$. Finally, we have $\Pi^{Y>X} = \varnothing$, but $\Pi^{X>Y} = \{(2,8)\}$ in $(0,10)$ meaning that responders derive no loss from *X* and a material gain of 2. The highest fairness level reached in Π^Y is $f = \alpha 5$. Since $\alpha 5 + (1-\alpha)2 \geq \alpha 5$, responders feel entitled to their material gain. We thus have $nl_R^{(0,10)} = -2\gamma\beta$. From our parameter assumptions, it follows that $nl_R^{(5,5)} > nl_R^{(8,2)} > nl_R^{(2,8)} > nl_R^{(0,10)}$.

I now turn to net losses imposed on proposers. Since $\Pi^X = \{(2,8),(0,0)\}$, proposers derive no loss from acceptance. Their gain is limited to $\beta 8$ because $(2,8)$ is fairer than $(0,0)$. All in all, $nl_P = -8\gamma\beta$. Conversely, next to a material loss of 8 from rejection, proposers suffer a fairness loss of $f((2,8)) - f((0,0)) = \alpha 5 + (1-\alpha)2 = 2 + 3\alpha$. As a result, their net loss is given by $nl'_p = \beta 8 + (1-\beta)(2+3\alpha) = 2 + 3\alpha + 6\beta - 3\alpha\beta > nl_p$. From $nl_R^{(5,5)} > nl_R^{(8,2)} > nl_R^{(2,8)} > nl_R^{(0,10)}$ and Lemma 2, it then follows that $Pr^{(5,5)} > Pr^{(8,2)} > Pr^{(2,8)} > Pr^{(0,10)}$.■

PROOF OF PROPOSITION 4

The opportunity set given trust is $\Pi^T = \{(20-r,r) : r \in \{0,1,...,20\}\}$ where r is the amount shared. Denoting no trust by *NT*, we have $\Pi^{T>NT} = \{(20-r,r) : r \in \{0,1,...,19\}\}$ in both treatments because all outcomes in Π^T except $(0,20)$ give the second mover more than the no-trust outcome. The fairness associated with no trust is $f = \alpha 2$ in $(0,4)$ and $f = \alpha 8$ in $(0,16)$, while the fairness reached in $\Pi^{T>NT}$ as a function of $r \in \{0,1,...,19\}$ is given by

$$f = \alpha 10 + (1-\alpha) min\{r, 20-r\} > \alpha 8 > \alpha 2,$$

which implies that gains are limited to material gains. Hence, we have $g_{SM}^{(0,4)} = g_{SM}^{(0,16)} = \beta 20$ and $nl_{SM}^{(0,4)} = nl_{SM}^{(0,16)} = -\gamma\beta 20$. By Lemma 2, the repayment distributions are then the same.

Next, we show that $nl_{SM}^{(0,10)} = nl_{SM}^{(5,5)}$. In $(0,10)$, we have $nl_{SM}^{(0,10)} = -\gamma\beta 20$ for the same reasons as before. Since the second mover derives no loss from trust, $nl_{SM}^{(0,10)} = nl_{SM}^{(5,5)}$ is trivially satisfied if $\gamma = 0$ because all gains are then fully discounted.

Suppose instead that $\gamma > 0$, $0 < \beta \leq 0.5$ and $\alpha = (1-2\beta)/(2-2\beta)$ and notice that the last two conditions imply $0 \leq \alpha < 0.5$. In $(5,5)$, we have $\Pi^{T>NT} = \{(20-r, r) : r \in \{0,1,...,14\}\}$. Given that $\alpha < 0.5$, the least fair outcome in $\Pi^{T>NT}$, namely, $(20,0)$, is less fair than $(5,5)$, which implies a positive fairness gain from trust. As the maximisation of material and fairness gains points into the same direction, with $(20,0)$ maximising both, we obtain

$$g_{SM}^{(5,5)} = \beta 15 + (1-\beta)(5-\alpha 10) = 5 - 10\alpha + 10\beta + 10\alpha\beta.$$

We then have

$$nl_{SM}^{(5,5)} = nl_{SM}^{(0,10)} \Leftrightarrow -\gamma(5 - 10\alpha + 10\beta + 10\alpha\beta) = -\gamma\beta 20 \Leftrightarrow \alpha = (1-2\beta)/(2-2\beta),$$

as assumed.

Finally, if $\gamma > 0$, $\beta = 0$ and $\alpha \geq 0.5$, the second mover disregards his material gains, but his fairness gain from trust is zero in $(5,5)$ because fairness leans towards efficiency. This implies $nl_{SM}^{(5,5)} = 0 = nl_{SM}^{(0,10)}$. ■

1.8 Appendix B: Intention-Based Models of Reciprocity

In this section, I sketch the main features of the reciprocity models of Dufwenberg and Kirchsteiger (2004) and Falk and Fischbacher (2006). I focus on how player *i* evaluates the *kindness* of player *j*'s pure strategy s_j. Conceptually, the kindness of s_j plays the same role as *i*'s net loss from s_j in my model. Since *i* observes s_j in the applications considered in this chapter, *i* need not form a belief about s_j. At the same time, both approaches draw on what is called *i*'s *second-order belief*, i.e., *i*'s belief about *j*'s belief about *i*'s own strategy. Both allow this belief to refer to a behaviour strategy σ_i. I denote *i*'s second-order belief about σ_i by σ_{iji}. In this chapter, I use for σ_{iji} the empirically observed choices of players *i*. The justification is that both models are equilibrium models and hence require beliefs to coincide with actual behaviour.

Dufwenberg and Kirchsteiger (2004) define

$$k\left(s_j,\sigma_{iji}\right)=\pi_i\left(s_j,\sigma_{iji}\right)-\pi_i^e\left(\sigma_{iji}\right)$$

where $\pi_i\left(s_j,\sigma_{iji}\right)$ is *i*'s expected payoff from *j*'s strategy s_j given his second-order belief σ_{iji}, i.e., the payoff to himself *i* thinks *j intends* him to receive, and $\pi_i^e\left(\sigma_{iji}\right)$ the payoff to himself *i* views as "equitable" given σ_{iji}. It is defined by

$$\pi_i^e\left(\sigma_{iji}\right)=0.5\cdot\left(\max_{s_j\in S_j}\pi_i\left(s_j,\sigma_{iji}\right)+\min_{s_j\in S_j}\pi_i\left(s_j,\sigma_{iji}\right)\right).$$

This formulation slightly simplifies the original model of Dufwenberg and Kirchsteiger (2004), which is inconsequential in the examples considered here. The interpretation is that *i* feels neutral about s_j $(k=0)$ if he believes *j* intends him to receive half of what he maximally and minimally stands to earn given *j*'s strategy set S_j and his second-order belief σ_{iji} and feels s_j is (un)kind whenever he receives more (less), to which correspond $k>0$ $(k<0)$. A direct implication is that the kindness of s_j equals zero if *j* is passive. Player *i* responds to the kindness of s_j as follows: If s_j is (un)kind, he is willing to increase the (un)kindness of his own behaviour to *j* at some material payoff cost to himself.

The reciprocity model of Falk and Fischbacher (2006) differs from Dufwenberg and Kirchsteiger (2004) in that distributional concerns directly influence kindness perceptions. The kindness of strategy $s_j\in S_j$ as perceived by player *i* is given by

$$k\left(s_j,\sigma_{iji}\right)=\left[\pi_i\left(s_j,\sigma_{iji}\right)-\pi_j\left(s_j,\sigma_{iji}\right)\right]\cdot\Delta\left(s_j,\sigma_{iji}\right).$$

The first term is called the "outcome term". It consists of the inequality associated with the outcome $\pi\left(s_j,\sigma_{iji}\right)$ implemented by s_j given *i*'s second-order belief. If $\pi\left(s_j,\sigma_{iji}\right)$ puts player *i* ahead of (behind) *j*, *i* tends to view s_j as (un)kind. At the same time, the outcome term does not reflect the alternatives to s_j that *j* has at her disposal. This is where the second term (the "intention factor") comes into play. It takes on either the value 1 or $\varepsilon \in \left[0,1\right]$. For example, if $\pi\left(s_j,\sigma_{iji}\right)$ puts *i* ahead of *j*, the intention factor equals 1 if the feasible set of outcomes given σ_{iji} contains a payoff to *i* smaller than $\pi_i\left(s_j,\sigma_{iji}\right)$ and ε otherwise. The idea is that in the first case *j* could have treated *i* worse than giving him $\pi_i\left(s_j,\sigma_{iji}\right)$, whereas no such option was available in the second case. As a result, *i* discounts his advantage $\pi_i\left(s_j,\sigma_{iji}\right)-\pi_j\left(s_j,\sigma_{iji}\right)>0$ in the second case, but not in the first. The procedure for the case where $\pi\left(s_j,\sigma_{iji}\right)$ puts *i* behind is analogous.

2 THE POWER OF DELEGATION

2.1 Introduction

In economic settings, decision rights may be delegated for various reasons.[1] This chapter focuses on delegation as a means to avoid responsibility and hence punishment for hurtful or unpopular decisions. For example, external consultants are often used in organisations to devise restructuring measures that are painful for at least some members of the organisation. While such consultants may possess superior expertise and provide an outside perspective, an additional explanation for their use is that they attract most of the blame for the unpopular decisions they propose, allowing the organisation's own management to avoid "punishment" (e.g., in the form of low effort or sabotage). The example points to what is called here the *power of delegation*. The power of delegation holds if a principal's expected payoff from delegating her choice from a fixed set of options to an agent exceeds her expected payoff from choosing any of these options directly. Since delegation entails a loss of control for the principal, the power of delegation can only be satisfied if there are third parties who can punish (or reward) the principal.[2]

This chapter sets out to explain the power of delegation in the presence of punishment opportunities. The main setting studied below is the following: A principal can either choose between a fair ("popular") and unfair ("unpopular") option herself or delegate this decision to an agent, where the unfair option yields a higher payoff to both the principal and agent and a lower payoff to two "recipients", who can subsequently punish the principal or agent or both. In the experimental study by Bartling and Fischbacher (2011), the power of delegation is reported to hold in this setting. Next to the behaviour of agents, the key contributing factor to this result is the punishment behaviour of recipients. In particular, the principal is punished much less if the

[1] Bartling and Fischbacher (2011) review the different strands of literature.

[2] Otherwise, the principal's payoff from delegation cannot exceed her highest payoff from choosing directly. Note also that the power of delegation can hold both empirically and in theory.

agent has chosen the unfair option on her behalf than if she has done so herself. Below, I develop a theoretical model that can account for recipients' punishment behaviour in this and related settings.

The model builds on the following idea: When faced with some behaviour of the principal and agent (e.g., if the principal has delegated and the agent chosen the unfair option), recipients determine the *net losses* they derive from these behaviours. Net losses are simply losses minus gains, where the two need not count for the same.[3] Recipients assess their net losses because of their preference for *net-loss reciprocation*: The higher their net loss from a player's behaviour, the higher their willingness to impose a net loss on that player by punishing him or her. As a result, recipients punish more the player from whose behaviour they derive the higher net loss.

This approach accounts for the punishment patterns constitutive of the power of delegation as follows: If the principal delegates and the agent chooses the unfair option, recipients derive a low net loss from the principal and a high net loss from the agent. The reason is that delegation leaves open the possibility that the fair option is chosen according to recipients' *ex-ante beliefs* about the agent, while the agent rules out the fair option directly. On the other hand, if the principal chooses the unfair option, she rules out the fair option directly and hence imposes the same high net loss on recipients as the agent in the first situation, while the agent's choice remains unobserved. Given their beliefs about the agent, recipients derive an *expected net loss* from the agent that equals their low net loss from the principal in the first situation. As a result, the two situations are mirror images in terms of net losses derived by recipients. This explains why the latter punish the agent severely and the principal lightly in the first situation, while this pattern is exactly inverted in the second situation (see Bartling and Fischbacher, 2011).

The model of net-loss reciprocation developed in this chapter is a generalisation of the model of net-loss reciprocation developed in Chapter 1 to explain the context-dependency of social preferences in two-player games without moves of Nature ("forgone-option effects"). In contrast, the model proposed in this chapter involves $n \geq 3$ players, one of whom is Nature. Yet, this setting nests the setting considered in Chapter 1 because it effectively collapses to a two-

[3] Losses and gains have a material as well as a fairness component. Broadly, recipients derive a *material loss* from some behaviour x of some player if there exists an alternative behaviour under which they could have earned more than what they can maximally earn under x. Moreover, recipients derive a *fairness loss*, which they add to their material loss, if their material loss derives from outcomes of the game that are fairer than the maximal fairness attainable under x. Material and fairness gains are defined similarly.

player game without Nature if $n = 3$ and Nature is passive. The purpose of this chapter is therefore twofold: First and foremost, it aims to account for the punishment patterns sustaining the power of delegation, which is an important economic phenomenon in its own right. Yet, a second purpose is to demonstrate that the model developed in Chapter 1, which performs well in the specific class of situations considered there, can be extended and successfully applied to more complex settings like the delegation games considered in this chapter.

Bartling and Fischbacher (2011) investigate how well existing behavioural theories perform in explaining their results on punishment patterns and find that these theories cannot account for all their evidence. For instance, *outcome-based theories* like the inequality-aversion model of Fehr and Schmidt (1999) cannot explain why punishment is targeted toward whichever player ultimately chooses the unfair option. The reason is that the principal and agent earn the same under the unfair option, which means that there is no difference between them from an outcome perspective. As a result, recipients have no systematic reason for punishing one player more than the other. Bartling and Fischbacher also propose their own model of *responsibility* for the unfair outcome to explain punishment choices, which they show to perform better than all existing approaches. In comparison to net-loss reciprocation, the responsibility model performs as well, but lacks the generality of net-loss reciprocation as a behavioural theory.

Besides Bartling and Fischbacher (2011), there exists a small number of experimental studies on delegation. In Hamman et al (2010), principals have the possibility of selecting agents for choosing on their behalf among different allocations involving a passive recipient. The authors find that principals prefer agents who choose more unfair allocations (benefiting the principal at the expense of the recipient) than principals choose themselves if they do not have the option of using an agent. In this sense, the power of delegation holds: Principals seek out agents in such a way that delegation yields them a higher payoff than choosing from the same set of options themselves. However, the reason behind this is not punishment, but principals' reluctance to make unfair choices, this reluctance disappearing if the same choices are made through agents. In contrast, this chapter focuses on the punishment behaviour of recipients, which helps sustain the power of delegation in settings with punishment opportunities. More in line with this chapter, Fershtman and Gneezy (2001) find that a principal's payoff in an ultimatum game is higher if she makes offers through an agent (who she can incentivise to make unfair offers). The reason is that responders punish unfair offers less if they have been made through the agent.

Coffmann (2011) also studies delegation and finds that delegation helps the principal avoid (third-party) punishment for unfair allocations. However, the power of delegation as defined in this chapter does not apply to his setup because the agent cannot choose from the same set of options as the principal. Rather, the principal first divides a fixed sum of payoff between a passive recipient and herself. The principal can then delegate to the agent, which gives the agent the option of taking between nothing and everything from the recipient of what the principal has previously left to the latter. Strikingly, even if the principal has left nothing for the agent to take away from the recipient by keeping the entire sum for herself, which means that delegation has no impact on the final allocation, delegation still leads to less punishment for the principal. Clearly, this result must be due to framing effects, which are beyond the scope of this chapter.

The remainder of this chapter is structured as follows. I first develop the model, which is followed by a detailed discussion of how the model can account for the punishment patterns constitutive of the power of delegation in Bartling and Fischbacher (2011) as well as a discussion of alternative approaches. All proofs are in Appendix A.

2.2 The Model: Net-Loss Reciprocation

The theoretical idea used in this chapter to explain the power of delegation is *net-loss reciprocation*. Net-loss reciprocation asserts that players' willingness to pay for imposing a net loss on others (e.g., by punishing them) increases in the net loss that they derive from these others' choice of strategy. Net losses are simply losses minus gains, where losses may loom larger. Before discussing derived and imposed net losses in turn, I first introduce the modelling framework (of which the delegation games considered below are special cases) as well as some elementary concepts.

I limit attention to finite-horizon multi-stage games. The set of players is I where $|I| = n \geq 3$ and one of the players is Nature (denoted N). A player's inactivity at a stage is modelled by the respective action set being singleton. At every stage, players have certainty about what happened at the previous stages, i.e., the non-terminal history up to that stage. Let H be the set of non-terminal histories, which contains the empty history (or root of the game) $\varnothing$. The *pure strategy* $s_i \in S_i$ of player $i \in I$ assigns to each history $h \in H$ an action $a_i \in A_i(h)$ available to i

at *h*, whereas *i*'s *behaviour strategy* $\sigma_i \in \Sigma_i$ assigns to each history a probability distribution on *i*'s available actions.[4] The set of pure and behaviour strategy profiles are given by $S = \prod_{i \in I} S_i$ and $\Sigma = \prod_{i \in I} \Sigma_i$, respectively. Outcomes $\hat{\pi}$ of the game are $(n-1)$-dimensional vectors of material payoffs.[5] The function $\hat{\pi} : S \to \mathbb{R}^{n-1}$ is the *outcome function*. It assigns to each pure strategy profile the payoff vector implemented by it. From $\hat{\pi}(s)$, we can derive $\pi : \Sigma \to \mathbb{R}^{n-1}$, which assigns to each profile of behaviour strategies the vector of expected payoffs implemented by it. The set $\Pi = \{\pi(\sigma) : \sigma \in \Sigma\}$ contains the game's feasible expected outcomes.[6]

Moreover, a few non-standard concepts are drawn on below. Firstly, $S_i(h) \subseteq S_i$ is the set of strategies of *i* that are *consistent* with history $h \in H$ in the following sense: If $h \neq \varnothing$, s_i is part of $S_i(h)$ if and only if it prescribes *i*'s actions contained in h. If $h = \varnothing$, we have $S_i(h) = S_i$. The set $\Sigma_i(h)$ is defined analogously: All actions in h must be prescribed with probability one for $\sigma_i \in \Sigma_i(h)$. Secondly, $s_i(s_i, h) \in S_i$ is the "update" of s_i that coincides with s_i except that it prescribes *i*'s actions contained in h and likewise for $\sigma_i(\sigma_i, h)$ where the actions in h are prescribed with probability one. Finally, $H(s_i) \subseteq H$ is the set of histories that are consistent with s_i in the sense that any $h \in H$ with $h \neq \varnothing$ is in $H(s_i)$ if and only if *i*'s actions contained in *h* are also actions prescribed by s_i. Moreover, $\varnothing \in H(s_i)$ for all $s_i \in S_i$.

To illustrate the model developed in this section, I draw on a simple delegation game corresponding to the main treatment in Bartling and Fischbacher (2011). The game has four players:[7] One principal (player A), one agent (player B) and two passive recipients, one of whom is called player C. The principal moves first. She can either implement a fair or unfair outcome directly or delegate this choice to the agent. The fair outcome yields 5 units of payoff to all parties, while the unfair outcome gives 9 units of payoff each to the principal and agent and 1 unit to each recipient. The punishment opportunities after A or A and B have made their choice are left out of the picture. This is done to simplify things, but does not affect the qualitative predictions of the model (see also the discussion in Section 2.3).

---

[4] All action sets are assumed finite.

[5] Outcomes specify a payoff for each player except Nature.

[6] Note that $S \subseteq \Sigma$ since all pure strategy profiles are degenerate behaviour strategy profiles. As a result, all pure strategy profiles are in the domain of π.

[7] Strictly speaking, Nature is the fifth player, who is however passive in this example.

2.2.1 Derived Net Losses

Player i's net loss from strategy s_j of player j consists of his loss minus gain from s_j where the two need not count for the same.[8] Player i assesses his loss and gain by comparing the *opportunity set* of outcomes created by s_j to the opportunity sets created by j's alternative strategies. In our example, suppose that C evaluates A's decision to delegate. He then compares the opportunity set created by A delegating to the opportunity sets created by A choosing the fair and unfair outcome directly.

Defining such opportunity sets raises several modelling issues. Firstly, the question arises which (if any) restrictions to place on the behaviour of third parties, i.e., on the other players besides i and j. Player A's decision to delegate is a case in point as its consequences depend on the behaviour of B, who is the third party in the relationship between A and C. In what follows, I assume that i considers the opportunity sets of *expected outcomes* created by s_j and its alternatives taking as given $\sigma_{-i,j} \in \prod_{k \in I \setminus \{i,j\}} \Sigma_k$, which is the profile of behaviour strategies of all other players including Nature and can be interpreted as i's belief about these players' average behaviour. The idea is that i, when assessing the opportunity sets created for him by s_j and its alternatives, has some sense of how third parties are likely to act, which affects his sense of opportunity.[9] In the delegation example, B is expected to choose the unfair outcome after delegation with a probability of 0.34 according to the beliefs measured by Bartling and Fischbacher (2011). The opportunity set of expected outcomes created by delegation is therefore

$$\left\{\left(0.34\cdot 9+0.66\cdot 5, 0.34\cdot 9+0.66\cdot 5, 0.34\cdot 1+0.66\cdot 5, 0.34\cdot 1+0.66\cdot 5\right)\right\}=\left\{\left(6.36, 6.36, 3.64, 3.64\right)\right\},$$

while A choosing the fair and unfair outcome directly entail $\left\{\left(5,5,5,5\right)\right\}$ and $\left\{\left(9,9,1,1\right)\right\}$, respectively.[10] As discussed before, these opportunity sets are singleton because C's punishment options are ignored.

---

[8] I refer to i as "he" and j as "she".

[9] Alternatively, $\sigma_{-i,j}$ could be interpreted as i's belief about j's belief about the other players, i.e., i's second-order belief.

[10] The first entry in payoff vectors refers to the payoff of A, the second to the payoff of B and the last two to the payoffs of the recipients.

A second issue is which perspective *i* adopts when assessing s_j and its alternatives given $\sigma_{-i,j}$. On the one hand, *i* could evaluate s_j from an *ex-ante* perspective meaning that he simply compares the opportunity set of expected outcomes created by s_j to the opportunity sets created by its alternatives taking as given $\sigma_{-i,j}$. On the other hand, he could adopt the perspective of some history $h \neq \varnothing$ of the game and evaluate *j*'s updated strategy $s_j\left(s_j, h\right)$ against its alternatives in $S_j\left(h\right)$ taking as given $\sigma_{-i,j}\left(\sigma_{-i,j}, h\right)$. The importance of such conditioning on histories can be seen by considering a second example, namely, C's evaluation of B choosing the unfair outcome after delegation, the alternative being choosing the fair outcome. From an *ex-ante* perspective, the opportunity sets of expected outcomes created by these two strategies depend on the belief about A, i.e., about how likely A is to delegate in the first place. In the most extreme scenario, where A is not believed ever to delegate, the two opportunity sets would be the same implying C's net loss from the two strategies is the same, namely, zero. This dependence of the evaluation of B on the beliefs about A seems implausible. Intuitively, given the information structure, C knows that B knows that A has delegated when B chooses the unfair outcome and C wants to hold B to account for this knowledge. This consideration can be captured if we condition on the history "A has delegated", which means that the likelihood of delegation is set to one. For this reason, I posit that *i* when evaluating s_j adopts the perspective of all histories in the set $H\left(s_j\right)$, which is the set of histories consistent with s_j. The idea is that *i* restricts attention to histories not ruled out by s_j, which has intuitive appeal.[11] Consequently, I first define *i*'s loss and gain from s_j for a given $h \in H\left(s_j\right)$ and then define *i*'s overall loss and gain as his maximal history-contingent loss and gain with respect to $H\left(s_j\right)$ as a whole.

Finally, when adopting the perspective of some $h \in H\left(s_j\right)$, the question arises if we should restrict *i*'s own behaviour to be in $S_i\left(h\right)$ in establishing the opportunity sets created by s_j and its alternatives. While the answer to this question is inconsequential in the setting considered below,[12] I include it for completeness: Limiting *i*'s behaviour to $S_i\left(h\right)$ is not fully convincing

---

[11] Limiting the conditioning to $H\left(s_j\right)$ is also required for making the definition of *i*'s loss and gain consistent with the definition given in Chapter 1 for a two-player setting, where there is no conditioning on histories ruled out by s_j. Consistency means that in any *n*-player game where the players in $I \setminus \{i, j\}$ are passive, *i*'s loss and gain from any given s_j is the same as his loss and gain from the corresponding s_j in the corresponding two-player game where the players in $I \setminus \{i, j\}$ are omitted.

[12] The reason is that the evaluating player C is passive implying that S_C is singleton

because i aims to assess the "elbow room" left for him by s_j and its alternatives. From this angle, restricting attention to $S_i(h)$ seems misguided. Intuitively, i holds j responsible for choosing s_j rather than its alternatives in $S_j(h)$ given that third parties behave according to $\sigma_{-i,j}(\sigma_{-i,j},h)$, but does not hold j responsible for ending up in h. Any part that j has played in bringing about h is dealt with by considering the rest of $H(s_j)$. Consequently, I define the opportunity set created by s_j from the perspective of h as $\left\{\pi\left(s_i,s_j,\sigma_{-i,j}\left(\sigma_{-i,j},h\right)\right): s_i \in S_i\right\}$ and likewise for the alternatives.

In a first step, I now define player i's *loss* and *gain* from $s_j \in S_j$ (the "status quo") relative to some alternative $\tilde{s}_j \in S_j$ without conditioning on histories. The sets Π^{s_j} and $\Pi^{\tilde{s}_j}$ are the opportunity sets of expected outcomes created by the two strategies. At this point, I only assume them to be non-empty without worrying about their precise definition, which is history-dependent and introduced at a later stage. For the sake of illustration, I continue to draw on our example, where we have $\Pi^D = \{(6.36, 6.36, 3.64, 3.64)\}$ for delegation, $\Pi^F = \{(5,5,5,5)\}$ for choosing the fair outcome and $\Pi^U = \{(9,9,1,1)\}$ for choosing the unfair outcome.

I begin with *losses*.[13] Player i derives a loss from s_j relative to $\tilde{s}_j$ only if what he can earn from $\tilde{s}_j$ exceeds what he can maximally earn from s_j. To capture this idea, let $\Pi^{\tilde{s}_j > s_j}$ be the set of outcomes in $\Pi^{\tilde{s}_j}$ that yield i a higher payoff than his highest payoff in Π^{s_j}. For i to derive a loss from s_j relative to $\tilde{s}_j$, $\Pi^{\tilde{s}_j > s_j}$ must be non-empty. The magnitude of i's loss is established by considering the different outcomes in $\Pi^{\tilde{s}_j > s_j}$. For each $\pi \in \Pi^{\tilde{s}_j > s_j}$, i calculates his *material loss*, which is the difference between his payoff from π and his highest payoff in Π^{s_j}, and his *fairness loss*, which is the extent to which π is fairer than the highest fairness attained by the outcomes in Π^{s_j}. While i's material loss is always positive, the fairness of π may or may not exceed the maximal fairness in Π^{s_j}. In the latter case, i derives no fairness loss because he has no fairness claim to π despite his higher earnings from it. All in all, i's loss from s_j relative to π is the weighted sum of his material and fairness loss, while his overall loss from s_j relative to $\tilde{s}_j$ is his maximal loss from s_j relative to all outcomes in $\Pi^{\tilde{s}_j > s_j}$.

To express these ideas more formally, let $\Pi_i^{s_j}$ be the payoffs to i contained in Π^{s_j}, $\bar{\pi}_i^{s_j} = max\, \Pi_i^{s_j}$ i's maximal payoff given s_j and $\Pi^{\tilde{s}_j > s_j} = \left\{\pi \in \Pi^{\tilde{s}_j} : \pi_i > \bar{\pi}_i^{s_j}\right\}$ the set of feasible outcomes given $\tilde{s}_j$ yielding i more payoff than $\bar{\pi}_i^{s_j}$. Fairness is measured by a fairness function,

[13] The definition of losses (and gains) is analogous to the definition for two-player games in Chapter 1.

isoquants of which are called *fairness curves*:

DEFINITION 1 The *fairness function* $f:\mathbb{R}^{n-1}\to\mathbb{R}$ is given by

$$f(\pi)=\underline{\pi}+\alpha\sum\nolimits_{i\in I\setminus\{N\}}(\pi_i-\underline{\pi})/(n-1)$$

with $\underline{\pi}=min\{\pi_i : i\in I\setminus\{N\}\}$ and $\alpha\in[0,1]$.

If $\alpha=1$, a one-unit increase of any player's payoff above $\underline{\pi}$, the minimal payoff among all players, increases fairness by $1/(n-1)$, which is the concomitant increase in average payoff. Fairness in this case reduces to material efficiency meaning for any π and $\tilde{\pi}$ we have $f(\pi)>f(\tilde{\pi})$ if and only if $\sum\nolimits_{i\in I\setminus\{N\}}\pi_i>\sum\nolimits_{i\in I\setminus\{N\}}\tilde{\pi}_i$. If $\alpha=0$, we have $f(\pi)>f(\tilde{\pi})$ if and only if $\underline{\pi}>\underline{\tilde{\pi}}$. Fairness in this case is identical to a concern for the least well off. For fairness to increase, all players with the least payoff must receive more. In general, the higher α, the more weight is put on efficiency and the smaller the regard for the least well-off. Notice that this formulation of fairness leaves out considerations of payoff equality (Fehr and Schmidt, 1999) whenever an increase in the lowest payoffs does not coincide with a reduction in inequality.

We can now define *i*'s loss from s_j relative to $\tilde{s}_j$. For each $\pi\in\Pi^{\tilde{s}_j>s_j}$, *i* determines the weighted sum of his material loss $\pi_i-\overline{\pi}_i^{s_j}>0$ and his fairness loss $max\{f(\pi)-\overline{f}^{s_j},0\}$ where $\overline{f}^{s_j}=max_{\pi\in\Pi^{s_j}}f(\pi)$ is the highest fairness attained in Π^{s_j}. Player *i*'s overall loss is the maximal weighted sum of this kind with respect to $\Pi^{\tilde{s}_j>s_j}$. These ideas are summarised in

DEFINITION 2 Player *i*'s *loss* from strategy $s_j\in S_j$ relative to strategy $\tilde{s}_j\in S_j$ is given by

$$l_i(s_j,\tilde{s}_j)=\begin{cases}max_{\pi\in\Pi^{\tilde{s}_j>s_j}}\left[\beta\left(\pi_i-\overline{\pi}_i^{s_j}\right)+(1-\beta)max\left\{f(\pi)-\overline{f}^{s_j},0\right\}\right] & \text{if }\Pi^{\tilde{s}_j>s_j}\neq\varnothing\\ 0 & \text{otherwise}\end{cases}$$

where $\beta\in[0,1]$.

In our example, the loss from delegation (D) relative to A choosing the unfair outcome (U) is zero because C can earn more after D, i.e., we have $\Pi^{U>D} = \varnothing$. But C could have earned more had A chosen the fair outcome (F) because $\Pi^D = \{(6.36, 6.36, 3.64, 3.64)\}$ and $\Pi^F = \{(5,5,5,5)\}$, which implies $\Pi^{F>D} = \{(5,5,5,5)\}$. Accordingly, we have a material loss from D relative to F of 1.36. The fairness of $(5,5,5,5)$ is $f = 5$, while the highest fairness in Π^D is $f = 3.64 + \alpha(2.72/4 + 2.72/4)$. This makes for a fairness loss of $1.36 - 1.36\alpha$, which is positive except if $\alpha = 1$, in which case fairness coincides with material efficiency and $(5,5,5,5)$ is only as fair as $(6.36, 6.36, 3.64, 3.64)$. In contrast, if $\alpha < 1$, $(5,5,5,5)$ is fairer because it contains a larger minimal payoff. All in all, $l_C(D,F) = \beta(1.36) + (1-\beta)(1.36 - 1.36\alpha)$.

I next turn to player *i*'s *gain* from s_j relative to $\tilde{s}_j$. A necessary condition for *i* deriving a gain is that $\Pi^{s_j > \tilde{s}_j}$ is non-empty, where $\Pi^{s_j > \tilde{s}_j}$ contains the feasible outcomes given s_j yielding *i* more payoff than what he can maximally earn given $\tilde{s}_j$. The level of *i*'s gain is determined by considering the different outcomes in $\Pi^{s_j > \tilde{s}_j}$. From a given $\pi \in \Pi^{s_j > \tilde{s}_j}$, *i* derives a *fairness gain*, which he adds to his positive *material gain*, if and only if π lies on a lower fairness curve than the highest curve reached in $\Pi^{\tilde{s}_j}$. Intuitively, if π instead lay on the same or a higher curve, *i* would consider his material gain to be well-deserved for leading to no decrease in fairness. These considerations motivate

DEFINITION 3 Player *i*'s gain from strategy $s_j \in S_j$ relative to strategy $\tilde{s}_j \in S_j$ is given by

$$g_i(s_j, \tilde{s}_j) = \begin{cases} max_{\pi \in \Pi^{s_j > \tilde{s}_j}} \left[\beta\left(\pi_i - \overline{\pi}_i^{\tilde{s}_j}\right) + (1-\beta) max\left\{ \overline{f}^{\tilde{s}_j} - f(\pi), 0 \right\} \right] & \text{if } \Pi^{s_j > \tilde{s}_j} \neq \varnothing \\ 0 & \text{otherwise} \end{cases}$$

where $\beta \in [0,1]$.

In our example, we have $g_C(D,F) = 0$ because C earns less after D than after F. Moreover, we have $g_C(D,U) = 2.64\beta$ because $\Pi^D = \{(6.36, 6.36, 3.64, 3.64)\}$ and $\Pi^U = \{(9,9,1,1)\}$ implying $\Pi^{D>U} = \{(6.36, 6.36, 3.64, 3.64)\}$. This makes for a material gain from D of 2.64, but no fairness gain because $(6.36, 6.36, 3.64, 3.64)$ is never less fair than $(9,9,1,1)$.

I now address *i*'s loss and gain from s_j at large. The two are established by considering all histories consistent with s_j, which are collected in $H(s_j)$.

DEFINITION 4 From the perspective of history $h \in H(s_j)$, player *i*'s *loss* and *gain* from $s_j \in S_j$ given that the players in $I \setminus \{i,j\}$ follow $\sigma_{-i,j} \in \Sigma_{-i,j}$ are given by

$$l_i(s_j,h) = max_{\tilde{s}_j \in S_j(h)} l_i(s_j,\tilde{s}_j) \text{ and } g_i(s_j,h) = max_{\tilde{s}_j \in S_j(h)} g_i(s_j,\tilde{s}_j)\text{, respectively,}$$

where $\Pi^{s_j} = \{\pi(s_i,s_j,\sigma_{-i,j}(\sigma_{-i,j},h)) : s_i \in S_i\}$ and likewise for each $\Pi^{\tilde{s}_j}$. Moreover, *i*'s overall loss and gain from s_j are given by

$$l_i(s_j) = max_{h \in H(s_j)} l_i(s_j,h) \text{ and } g_i(s_j) = max_{h \in H(s_j)} g_i(s_j,h).$$

Thus, given $\sigma_{-i,j}$, *i* assesses his loss and gain from s_j history-wise by considering each element in $H(s_j)$. Adopting the perspective of some such history, *i* determines his maximal loss and gain from s_j relative to its alternatives in $S_j(h)$ taking as given $\sigma_{-i,j}(\sigma_{-i,j},h)$.[14] Player *i*'s overall loss and gain from s_j are given by his maximal history-contingent loss and gain with respect to $H(s_j)$. In our example, $l_C(D) = \beta(1.36) + (1-\beta)(1.36 - 1.36\alpha)$ and $g_C(D) = 2.64\beta$ because the root of the game is the only history at which A is active, which implies that C's loss and gain from delegation conditional on the other histories consistent with delegation are zero.[15]

In what follows, I suppose that players react to the *net loss* imposed on them by others:

DEFINITION 5 Player *i*'s *net loss* from strategy $s_j \in S_j$ is given by

$$nl_i(s_j) = l_i(s_j) - \gamma g_i(s_j)$$

with $\gamma \in [0,1]$.

The case $\gamma < 1$ allows for the possibility that "losses loom larger than gains", which is key in the

[14] $S_j(h)$ is guaranteed to include s_j since *h* is taken from $H(s_j)$.

[15] The only non-terminal history besides the root consistent with *D* is "A has delegated". There is only one strategy of A consistent with this history, namely, *D* itself, implying that A's contingent strategy set is singleton. As a result, C's loss and gain from *D* conditional on "A has delegated" equal zero.

literature on loss aversion (Kahnemann and Tversky, 1979; Köszegi and Rabin, 2006).[16] The following is immediate:

LEMMA 1 If $S_j = \{s_j\}$, i.e., if player j is passive, we have $nl_i(s_j) = 0$.

The lemma follows from the fact that we have $l_i(s_j) = 0$ and $g_i(s_j) = 0$ if $S_j = \{s_j\}$.

2.2.2 Imposed Net Losses and Preferences

In a second step, I turn to imposed net losses. Faced with some profile s_{-i} of the other players' strategies, player i must choose an outcome from his opportunity set $\Pi^{s_{-i}} = \{\pi(s_i, s_{-i}) : s_i \in S_i\}$. Since each element in $\Pi^{s_{-i}}$ imposes a particular net loss on each other player, i's choice from $\Pi^{s_{-i}}$ can be thought of as his reciprocation to his net losses from s_{-i}. I now define the net loss that i imposes on j through his choice from $\Pi^{s_{-i}}$.[17] Let $\pi^c \in \Pi^{s_{-i}}$ be the outcome chosen by i and let $\Pi^{s_{-i}>c,j} = \{\pi \in \Pi^{s_{-i}} : \pi_j > \pi_j^c\}$ and $\Pi^{s_{-i}<c,j} = \{\pi \in \Pi^{s_{-i}} : \pi_j < \pi_j^c\}$ contain the outcomes in $\Pi^{s_{-i}}$ yielding player j more and less payoff than π^c, respectively. This leads to

DEFINITION 6 Player j's *loss* from $\pi^c \in \Pi^{s_{-i}}$ is

$$l_j(\pi^c, \Pi^{s_{-i}}) = \begin{cases} max_{\pi \in \Pi^{s_{-i}>c,j}} \left[\beta(\pi_j - \pi_j^c) + (1-\beta) max\{f(\pi) - f(\pi^c), 0\}\right] & \text{if } \Pi^{s_{-i}>c,j} \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Moreover, j's *gain* from $\pi^c \in \Pi^{s_{-i}}$ is

$$g_j(\pi^c, \Pi^{s_{-i}}) = \begin{cases} max_{\pi \in \Pi^{s_{-i}<c,j}} \left[\beta(\pi_j^c - \pi_j) + (1-\beta) max\{f(\pi) - f(\pi^c), 0\}\right] & \text{if } \Pi^{s_{-i}<c,j} \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Finally, j's *net loss* from $\pi^c \in \Pi^{s_{-i}}$ is

[16] In Chapter 1, I discuss why net-loss reciprocation is not a model of loss aversion. Broadly, loss aversion asserts that not only the baseline utility of outcomes, but also the deviation of this utility from some reference level determines decision-making (with negative deviations being weighted more). In contrast, net-loss reciprocation should be thought of as a model of baseline utility attached to outcomes.

[17] The definition is again analogous to that in Chapter 1.

$$nl_j\left(\pi^c,\Pi^{s_{-i}}\right)=l_j\left(\pi^c,\Pi^{s_{-i}}\right)-\gamma g_j\left(\pi^c,\Pi^{s_{-i}}\right).$$

I next characterise *i*'s preferences governing his choice from $\Pi^{s_{-i}}$.

ASSUMPTION 1 Player *i*'s preferences on the outcomes in $\Pi^{s_{-i}}$ are represented by

$$u_i\left(\pi,s_{-i}\right)=v\left(\pi_i\right)+\sum\nolimits_{j\in I\setminus\{i,N\}}r\left(nl_j\left(\pi,\Pi^{s_{-i}}\right),nl_i\left(s_j\right)\right)$$

where the continuous $v:\mathbb{R}\to\mathbb{R}$ and $r:\mathbb{R}^2\to\mathbb{R}$ satisfy $dv/d\pi_i>0$ and $\partial^2 r/\partial nl_j\partial nl_i>0$.

Thus, *i*'s utility from $\pi\in\Pi^{s_{-i}}$ can be separated into the utility from his own payoff and a reciprocation term for each other player that depends on the net loss that π imposes on that player as well as the net loss that *i* himself derives from that player's choice of strategy. As a result, $WTP=\partial r/\partial nl_j/dv/d\pi_i$, which is *i*'s willingness to pay for increasing *j*'s net loss, increases in the net loss that *i* himself derives from s_j.

Below, *i* must sometimes choose from a deterministic set of outcomes $\hat{\Pi}$ without having certainty about the strategy of all other players. Such a situation arises if one player's choice of action precludes some other player from having a move, which is therefore not observed by *i*. In such cases, I assume that $u_i=v\left(\pi_i\right)+\sum\nolimits_{j\in I\setminus\{i,N\}}r\left(nl_j\left(\pi,\hat{\Pi}\right),E_{\sigma_{ij}}\left[nl_i\left(s_j\right)\right]\right)$ where σ_{ij} is *i*'s belief about *j*'s choice of strategy. Preferences thus depend on expected net losses. In our example, if A chooses the fair or unfair outcome directly, C does not observe B's strategy and must rely on his belief about B to determine his expected net loss from B's behaviour.

Moreover, I follow McFadden (1974) and McKelvey and Palfrey (1995, 1996) in making

ASSUMPTION 2 The probability of player *i* choosing outcome $\pi\in\Pi^{s_{-i}}$ is given by

$$Pr\left(\pi,s_{-i}\right)=exp\left[u_i\left(\pi,s_{-i}\right)\right]/\sum\nolimits_{\breve{\pi}\in\Pi^{s_{-i}}}exp\left[u_i\left(\breve{\pi},s_{-i}\right)\right].$$

This quantal response structure (see Goeree et al, 2008, and Chapter 1 for details) allows us to interpret the punishment data considered in the next section because it implies that each subject chooses each available outcome with positive probability.

2.3 The Power of Delegation

In this section, I analyse in detail the experimental study of delegation by Bartling and Fischbacher (2011). Bartling and Fischbacher (BF) examine the following treatments: Their main treatment, entitled *D&P* (*Delegation and Punishment*), gives the principal (player A) the option of either choosing the fair or unfair outcome herself or delegating this choice to an agent (player B). The fair outcome yields 5 units of payoff to all parties, while the unfair outcome gives 9 units of payoff each to A and B and 1 unit to each of the two recipients. After A or A and B have made their choices, one of the two recipients (called player C) is given the following punishment opportunity: If he pays one unit of payoff, he can assign up to seven punishment points to his fellow players. Every punishment point assigned to a player reduces that player's payoff by one unit. BF find that A attracts most of punishment if she chooses the unfair outcome herself, but is largely spared if B does the same on her behalf, the latter now receiving most of the punishment. In line with this, delegation gives A an expected payoff of 5.93, while choosing the fair and unfair outcome directly yield her 4.80 and 4.73, respectively. These differences being statistically significant, the power of delegation holds in D&P.

The treatment *Random* replicates D&P except that B is now passive because A can only delegate to a random device ("die"). The die chooses the unfair outcome with a commonly known probability of 0.4, which corresponds to the rate at which subjects in the role of B choose the unfair outcome in D&P. BF find that A cannot evade punishment if delegation leads to the unfair outcome to the same degree as in D&P: If A has delegated her decision to the die and the unfair outcome results, she is punished significantly more than in the analogous situation in D&P. In fact, A now attracts more punishment than B, albeit less than if she implements the unfair outcome herself, to such an extent that the power of delegation continues to hold in Random: Delegation gives A an expected payoff of 5.07, while choosing the fair and unfair outcome directly give her 4.82 and 4.36, respectively.[18]

Two further treatments are considered as robustness checks: The treatment *Asymmetric* eliminates A's option of choosing the unfair outcome herself. Delegation appears more self-serving than in D&P because it is the only way for A to attain the unfair outcome. And indeed,

[18] These differences are significant at the 5% (delegation versus fair) and 10% level (delegation versus unfair).

if B chooses the unfair outcome after delegation, A is punished more and B less than in the corresponding situation in D&P.[19] Finally, in *NoD&P* (*No Delegation and Punishment*) there is no delegation option meaning that B is passive like in Random.

2.3.1 Negative Reciprocity

I now show that net-loss reciprocation can account for the punishment patterns sustaining the power of delegation in the experiments conducted by BF. In this sub-section, I focus on punishments for the unfair outcome. Punishments for the fair outcome (which barely exist) are discussed in the following sub-section.

Player C's punishment decision after A or B has chosen the unfair outcome can be described as a choice from the set

$$P = \left\{ \left(p^A, p^B \right) : p^A, p^B \in \{0,1,\ldots,7\} \wedge p^A + p^B \leq 7 \right\}.^{20}$$

The option $\left(p^A = 0, p^B = 0 \right)$ corresponds to no punishment, which implements $(9,9,1,1)$.[21] The other punishment vectors capture C's decision to deduct up to seven units of payoff from A and B at a price of 1 to himself resulting in the final outcome $\left(9 - p^A, 9 - p^B, 0, 1 \right)$. Consequently, *P* defines the common set of outcomes $\hat{\Pi}$ from which punishing recipients choose in the situations considered in this sub-section.

Adopting the perspective of net-loss reciprocation, the elements in *P* map into net losses as follows:

$$nl_A\left(p^A, p^B \right) = \begin{cases} -\gamma\beta 7 & \text{if } p^A = 0 \\ \beta p^A + (1-\beta)\left(1 - \alpha 3/4 + \alpha\left(p^A + p^B\right)/4\right) - \gamma\beta\left(7 - p^A\right) & \text{if } p^A > 0 \end{cases}$$

---

[19] These differences are only significant in B's case. Although the power of delegation does not apply because A's and B's choice sets (net of delegation) differ, delegation maximises A's expected payoff.

[20] The experimental design of BF also allowed deductions from the other recipient. These are omitted because they (almost) never occurred.

[21] I omit C's option of paying one unit of payoff but not deducting any punishment points, which is dominated by not paying and is not reported to have been chosen by any recipient.

and likewise for player B. If $p^A = 0$, A bears no material loss and hence no overall loss. As for her gain, C could have reduced A's payoff by up to seven units. Yet, any such punishment would have resulted in a less fair outcome than $(9,9,1,1)$.[22] Consequently, A's gain is limited to her material gain of 7 and her net loss equals $-\gamma\beta 7$. If $p^A > 0$, A derives a material loss of p^A. Her fairness loss is the extent to which $(9-p^A, 9-p^B, 0, 1)$ is less fair than $(9,9,1,1)$. The fairness of the two is $f = 0 + \alpha(9-p^A)/4 + \alpha(9-p^B)/4 + \alpha/4$ and $f = 1 + 8\alpha/4 + 8\alpha/4$, respectively. The latter exceeds the former by $1 - 3\alpha/4 + \alpha(p^A + p^B)/4 > 0$ resulting in a total loss for A of $\beta p^A + (1-\beta)(1 - 3\alpha/4 + \alpha(p^A + p^B)/4)$. Moreover, A's gain is $\beta(7 - p^A)$ since C could have deducted up to seven punishment points.

In what follows, I make

ASSUMPTION 3 The following parameter restrictions are maintained throughout:

$\beta > 0$ and $\alpha < (4\beta + 4\beta\gamma)/(5 - 5\beta)$.

These *a priori* restrictions have intuitive appeal, as I now argue focusing again on A. Firstly, assuming $\beta > 0$, which implies that material factors play some role in the calculation of losses and gains, ensures $\partial nl_A / \partial p^A = \beta + (1-\beta)\alpha/4 + \gamma\beta > 0$ for all $p^A > 0$ meaning that A's net loss increases in the punishment points she receives.[23] Moreover, $\beta > 0$ guarantees that $\partial nl_A / \partial p^A > \partial nl_A / \partial p^B = (1-\beta)\alpha/4 \geq 0$ if $p^A > 0$. As a result, a one-unit increase in own punishment increases A's net loss more than a one-unit increase in B's.[24] Secondly, assuming $\alpha < (4\beta + 4\beta\gamma)/(5 - 5\beta)$ ensures that for any two punishment vectors the one imposing the larger punishment on A also imposes the larger net loss. More specifically, it guarantees that

---

[22] Since punishment costs the recipient one unit of payoff, both the minimal and average payoff are reduced by punishment, which implies that fairness decreases. Recall that fairness as defined in this paper does not value equality *per se* unlike in, e.g., Fehr and Schmidt (1999). Thus, reductions in inequality do not automatically lead to an increase in fairness.

[23] Also, since $\beta > 0$ implies $\beta + (1-\beta)(1 - \alpha 3/4 + \alpha(1 + p^B)/4) - \gamma\beta 6 > -\gamma\beta 7$, it ensures that A's net loss from $p^A = 1$ exceeds that from $p^A = 0$.

[24] The case $\partial nl_A / \partial p^B > 0$ corresponds to "fairness spillovers", which arise as follows: If p^B increases for a given $p^A > 0$, this leaves A's material loss constant, but makes the implemented outcome less fair (except if $\alpha = 0$ meaning that efficiency does not matter). This increases A's fairness loss (if $\beta < 1$).

A's net loss from $\left(p^A = 2, p^B = 0\right)$ exceeds that from $\left(p^A = 1, p^B = 6\right)$ and likewise for all higher punishment levels.

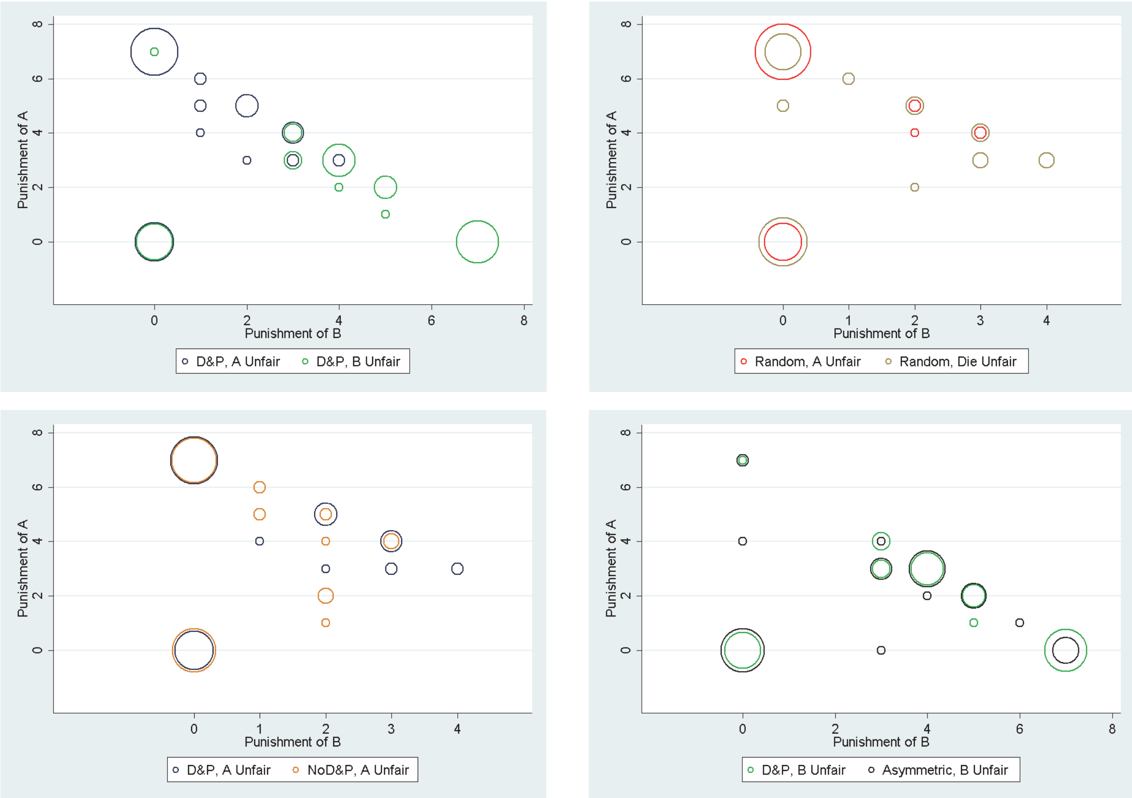

FIGURE 2.1 Allocations of Punishment In All Unfair Situations

Figure 2.1 shows recipients' punishment choices in the different treatments. The size of the circles corresponds to the frequency with which each punishment vector is chosen. I now present a series of propositions showing that my approach can account for the shifts in punishment across situations visualised in Figure 2.1. As argued above, these shifts help sustain the power of delegation in the experiment of BF. In deriving results, I follow BF in abstracting from the punishment stage. This means that the net loss that the punishing recipient (C) derives from A and B is determined by considering the game without punishment opportunities.[25]

[25] Otherwise, we would have to allow for the fact that the role of punisher is assigned randomly to one of the two recipients. The recipient selected to be punisher (player C) would have to form a belief about the likely punishment behaviour of the other recipient had she instead been selected. These beliefs are not measured by BF. If we used actual punishment choices, we would get similar results to those derived

Delegation to a Human Being

The top-left panel in Figure 2.1 shows the two "unfair" situations arising in D&P: A choosing the unfair outcome directly versus B choosing the unfair outcome after delegation. Here and in what follows, I say that the punishment of A *dominates* that of B if the (empirical or theoretical) distribution on A's punishment points first-order stochastically dominates the distribution on B's punishment points. We have

PROPOSITION 1 In D&P, if A chooses the unfair outcome directly, recipients derive $nl^A = 4(1-\alpha+\alpha\beta)$ and $nl^B = 1.36(1-\alpha+\alpha\beta) - 2.64\gamma\beta < nl^A$. As a result, the punishment of A dominates that of B. If B chooses the unfair outcome after delegation, recipients derive $nl^{A'} = 1.36(1-\alpha+\alpha\beta) - 2.64\gamma\beta = nl^B$ and $nl^{B'} = 4(1-\alpha+\alpha\beta) = nl^A > nl^{A'}$. As a result, the punishment of B dominates that of A. Moreover, the distribution of A's punishment in the first situation equals the distribution of B's punishment in the second situation and vice versa.

In the first situation, recipients derive a high net loss from A because she directly rules out the fair outcome, while recipients have to rely on their beliefs to assess B's behaviour. Given the beliefs measured by BF,[26] B chooses the fair outcome with a probability of 0.66, which implies that recipients derive a relatively low expected net loss from B. Consequently, they target their punishment mainly towards A. In the second situation, delegation leaves open the possibility that the fair outcome is chosen. Using measured beliefs, delegation implements the expected outcome $(6.36, 6.36, 3.64, 3.64)$, which yields the same low net loss as the expected net loss from B in the first situation. In contrast, B choosing the unfair outcome rules out the fair outcome, from which recipients derive the same high net loss as from A in the first situation.[27] Consequently, B is now the main punishment target. What is more, since the two situations are

below because punishment, which hurts fairness, mainly occurs after an unfair choice. Taking this fact into account would simply boost C's fairness loss from the unfair choice.

[26] BF conducted a separate, incentivised belief elicitation session where they asked subjects about the likelihood of the different moves of players A and B.

[27] This holds from the perspective of the history where A has delegated. From an *ex-ante* perspective, recipients derive a smaller net loss from B because A does not delegate for sure.

mirror images in terms of net losses derived by recipients, the punishment distributions for A and B are also mirror images.

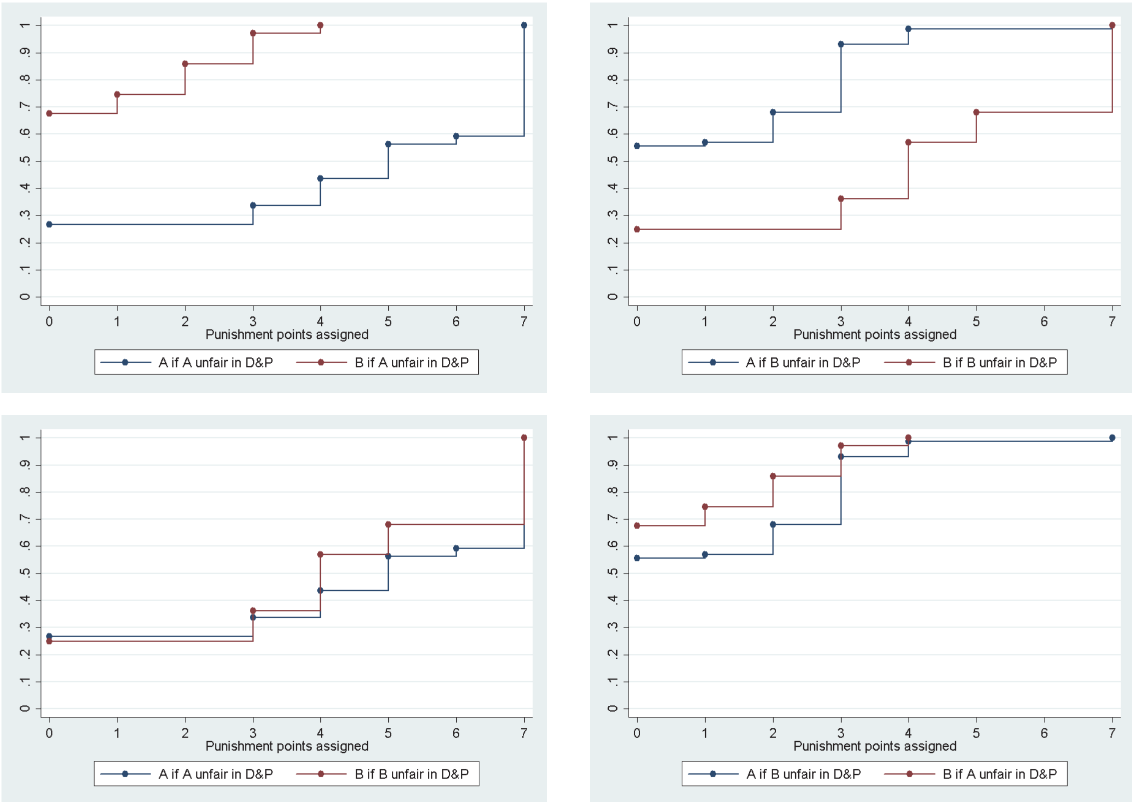

FIGURE 2.2 Punishment Distributions In the Unfair Situations in D&P

Proposition 1 presupposes no parameter restrictions beyond those contained in Assumption 3. In particular, the proposition is consistent with $\beta = 1$ meaning that only material factors matter. The intuition for this is that fairness and material concerns point in the same direction. Recipients' ranking of the different actions available to A and B is the same from a material and fairness perspective because recipients are the poorest players.

The experimental results of BF are consistent with Proposition 1: To see this, consider Figure 2.2, which contains the empirical punishment distributions in D&P, all of which are in line with Proposition 1. In particular, I can reject that the distribution pairs in the top panels are the same according to Kolmogorov-Smirnov (KS) tests ($p < 0.001$ in each case), while I fail to reject that the distribution pairs in the bottom panels are the same according to KS tests with $p = 0.481$ for the bottom-left and $p = 0.157$ for the bottom-right panel.

Delegation to a Random Device

The top-right panel in Figure 2.1 compares the two unfair situations in the treatment Random. In the first one, A has chosen the unfair outcome directly. In the second one, A has delegated her decision to a die choosing the unfair outcome with a commonly known probability of 0.4, while B is passive in both situations. As previously discussed, the power of delegation holds in Random. This suggest that the power of delegation is not limited to delegation to humans, but also works if there is delegation to an appropriately chosen random device. Regarding the punishment patterns sustaining this result, we have

PROPOSITION 2 Suppose that $\alpha < (1-1.5\gamma\beta)/(1-\beta)$ and $\gamma < 1/(1.5\beta)$. In Random, if A chooses the unfair outcome directly, recipients derive $nl^A = 4(1-\alpha+\alpha\beta)$ and $nl^B = 0 < nl^A$. As a result, the punishment of A dominates that of B. If A delegates to the die, recipients derive $nl^{A'} = 1.6(1-\alpha+\alpha\beta)-2.4\gamma\beta < nl^A$ and $nl^{B'} = 0 = nl^B < nl^{A'}$. As a result, A's punishment again dominates B's. Moreover, the punishment of A in the first situation dominates her punishment in the second situation.

Let us first focus on what happens across situations: Recipients' net loss from B equals zero in both because B is passive. As for A, delegation leaves open the possibility that the fair outcome materialises, while choosing the unfair outcome rules out the fair outcome for sure. We therefore have $nl^A > nl^{A'}$, which is guaranteed by $\beta > 0$ alone (from Assumption 3). Since we also have $nl^B = nl^{B'}$, A's punishment in the first situation dominates her punishment in the second situation (as Lemma A2 in the Appendix makes clear).

Let us now turn to the two within-situation comparisons: While $nl^A > 0$ follows directly from $\beta > 0$, $nl^{A'} > 0$ only holds given the additional restrictions on α and γ. The intuition for α is the following: For recipients' net loss from delegation to be positive, recipients' fairness loss must be sufficiently large. This is achieved by fairness leaning sufficiently towards a concern for the least well-off, i.e., towards $\alpha = 0$. If fairness instead coincided with material efficiency ($\alpha = 1$), there would be no fairness loss because delegation does not affect total payoffs. As for the restriction on γ, the formal reason is that $\alpha < (1-1.5\gamma\beta)/(1-\beta)$ can only be satisfied by $\alpha \geq 0$ if we have $\gamma < 1/(1.5\beta)$. Intuitively, if the gain from delegation is discounted too little, even $\alpha = 0$ would not generate high enough fairness losses for the net loss from delegation to be positive. The upper bound for γ falls in β because this problem becomes more acute as the weight on the material gain from delegation increases.

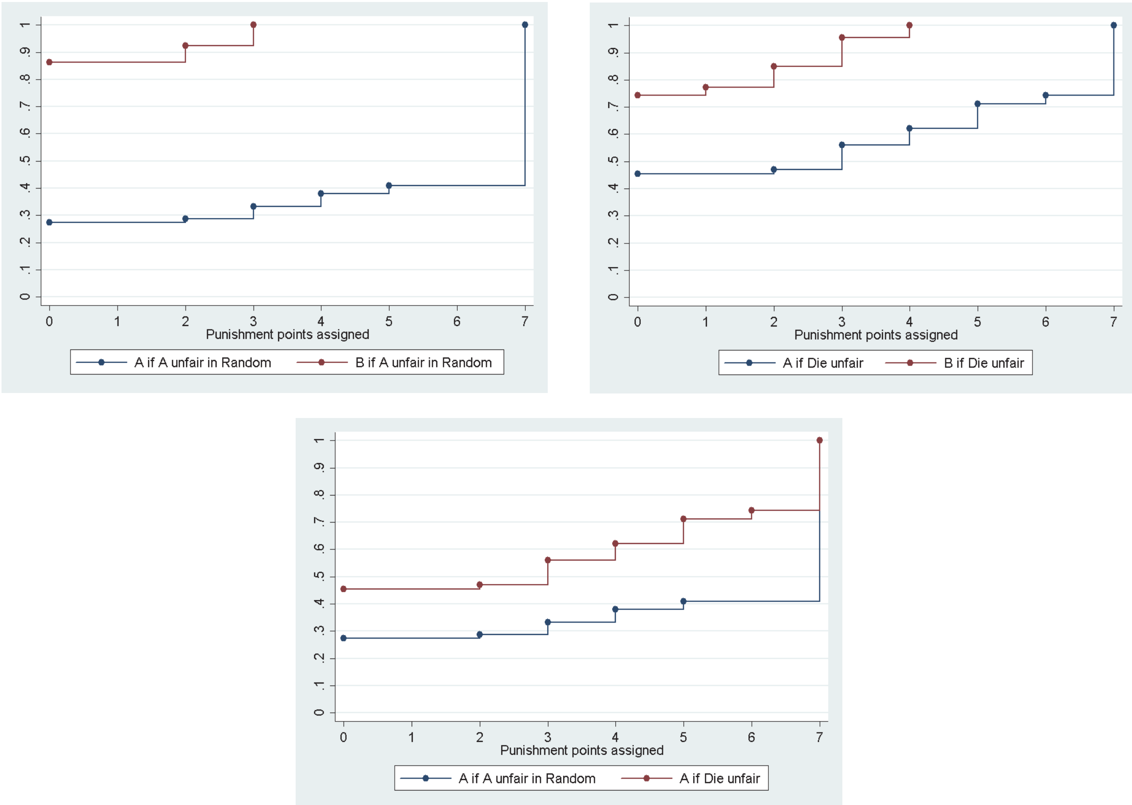

FIGURE 2.3 Punishment Distributions In the Unfair Situations in Random

Is Proposition 2 supported by BF's data? Figure 2.3 suggest this is the case. I can reject that the distribution pairs in it derive from the same data-generating process according to KS tests ($p < 0.01$ in each case).

Robustness of the Theory

I conclude this sub-section by carrying out two robustness checks for net-loss reciprocation. The first one explores the difference between being a passive player and being a player whose strategy has not been observed. Compare the following two situations: A choosing the unfair outcome in D&P and A choosing the unfair outcome in NoD&P. The difference between them is that B is passive in the latter and active in the former. Yet, as previously discussed, B's move in the first situation is not observed by recipients, who must therefore form a belief about how B would have acted had A chosen delegation. Given the beliefs measured by BF, we have

PROPOSITION 3 Suppose that $\alpha < (1.36 - 2.64\gamma\beta)/(1.36 - 1.36\beta)$ and $\gamma < 1.36/(2.64\beta)$. In D&P, if A chooses the unfair outcome, recipients derive $nl^A = 4(1-\alpha+\alpha\beta)$ and $nl^B = 1.36(1-\alpha+\alpha\beta) - 2.64\gamma\beta$. In NoD&P, if A chooses the unfair outcome, recipients derive $nl^{A'} = 4(1-\alpha+\alpha\beta) = nl^A$ and $nl^{B'} = 0 < nl^B$. Thus, B's punishment in the first situation dominates his punishment in the second.

The parameter restrictions ensure that the net loss that recipients derive from B in D&P is larger than zero and hence exceeds their net loss from B in NoD&P. The intuition is similar to that for Proposition 2: For recipients to derive a positive expected net loss from B, their loss from B choosing the unfair outcome after delegation must be sufficiently high, which is ensured by a sufficiently low α leading to a sufficiently large fairness loss. Furthermore, their material gain from B choosing the fair outcome must be sufficiently discounted. The top-left panel in Figure 2.4 contains evidence for Proposition 3 (see also the bottom-left panel in Figure 2.1). While the discrepancy between the two distributions has the direction suggested by the proposition, the difference is insignificant ($p = 0.803$, KS test).[28]

Finally, the bottom-right panel in Figure 2.1 compares B choosing the unfair outcome in D&P to B choosing the unfair outcome in Asymmetric. The net loss that recipients derive from B is the same in both situations, while the net loss from A is larger in Asymmetric. This is because the only alternative to delegation in Asymmetric is choosing the fair outcome. Unlike in D&P, there is therefore no gain from delegation. Also, recipients believe that agents are more likely to choose the unfair outcome in Asymmetric, which further boosts their net loss from delegation. These insights are summarised in

[28] Imposing the same parameter restrictions, the model also predicts that B's punishment if A has chosen the unfair outcome in D&P dominates that if A has chosen the unfair outcome in Random since B is also passive in the latter. The top-right panel in Figure 2.4 contains the associated empirical distributions, which are again consistent with my prediction. Yet, I fail to reject that the two distributions are the same ($p = 0.135$, KS test).

PROPOSITION 4 In D&P, if B chooses the unfair outcome, recipients derive $nl^A = 1.36(1-\alpha+\alpha\beta)-2.64\gamma\beta$ and $nl^B = 4(1-\alpha+\alpha\beta)$. In Asymmetric, if B chooses the unfair outcome, recipients derive $nl^{A'} = 1.56(1-\alpha+\alpha\beta) > nl^A$ and $nl^{B'} = 4(1-\alpha+\alpha\beta) = nl^B$. As a result, A's punishment in the second situation dominates her punishment in the first.

While the bottom panel in Figure 2.4 is broadly in line with Proposition 4, I cannot reject that the two distributions are the same ($p = 0.992$, KS test). I return to this issue below.

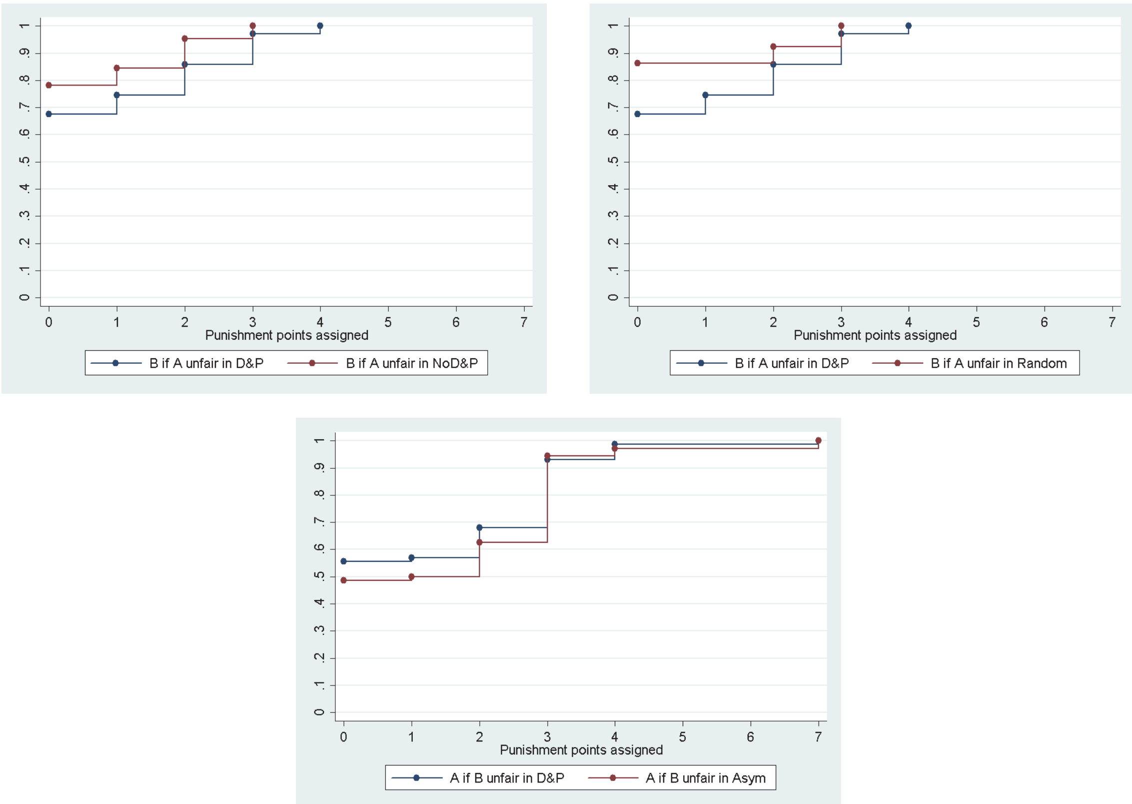

FIGURE 2.4 Punishment Distributions In Various Unfair Situations

Summing up, net-loss reciprocation can account for the most important punishment patterns sustaining the power of delegation in BF. I return to the problems encountered below. Before, I turn to positive reciprocity, i.e., the punishment behaviour of recipients if either A or B or the die have chosen the fair outcome.

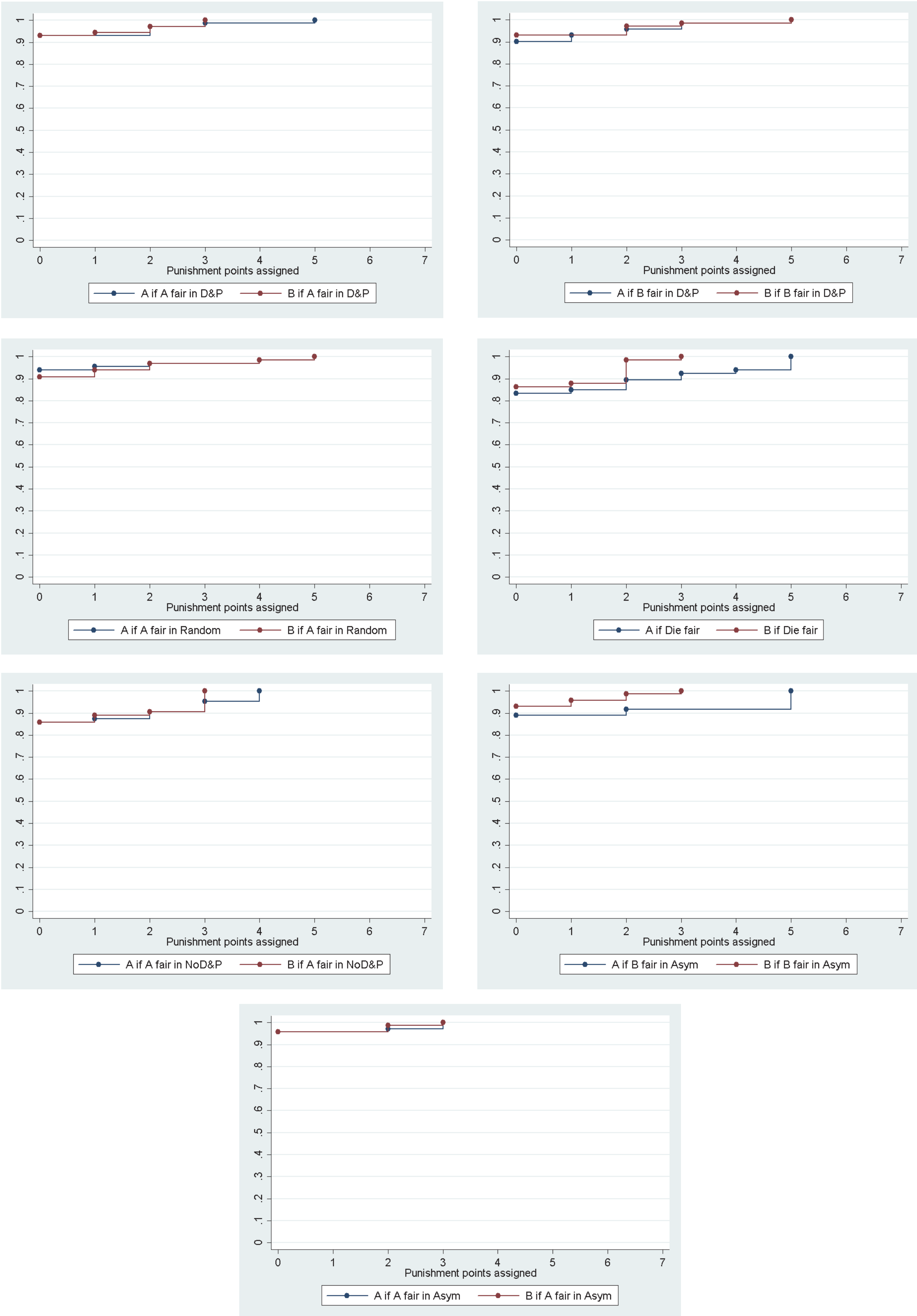

FIGURE 2.5 Punishment Distributions In All Fair Situations

2.3.2 Positive Reciprocity

Figure 2.5 displays the punishment of A and B in the "fair" situations examined by BF. Strikingly, all punishment distributions are very similar with the majority of recipients punishing neither player. This fact creates some problems for net-loss reciprocation: Consider for instance the right panel in the second row where the die has chosen the fair outcome after delegation. There is no significant difference between the punishments of A and B ($p = 0.925$, KS test), which points to $nl^A = 1.6(1-\alpha+\alpha\beta) - 2.4\gamma\beta = nl^B = 0$. This runs against Proposition 2, which uses $1.6(1-\alpha+\alpha\beta) - 2.4\gamma\beta > 0$ and was found well-supported by BF's data.

Recipients appear reluctant to punish whenever the fair outcome has materialised. In particular, they are reluctant to punish A for delegation if B or the die have ultimately chosen the fair outcome. BF term this the "no harm, no foul" principle. It means that recipients disregard the fact that A, by delegating, has left open the possibility that the unfair outcome results, which in my model translates into a loss for recipients from delegation relative to A choosing the fair outcome. Yet, recipients appear to pay no heed to this loss if B ends up choosing the fair outcome.

2.3.3 Discussion

Above, it was shown that the model of net-loss reciprocation defined in this chapter can by and large account for the punishment patterns constitutive of the power of delegation in the experimental study by BF. The following parameter restrictions are consistent with all results:

$$\beta > 0, \ \gamma < 1.36/(2.64\beta) \text{ and } \alpha < min\{(1.36-2.64\gamma\beta)/(1.36-1.36\beta), (4\beta+4\beta\gamma)/(5-5\beta)\}.$$

This merges Assumption 3 with the parameter restrictions in Proposition 3 and implies the restrictions in Proposition 2. The interpretation is that material factors must not be irrelevant, gains must be sufficiently discounted and fairness must not coincide with material efficiency, but make sufficient allowance for the least well-off. That said, the constraint on α becomes irrelevant as β approaches one. Indeed, given a not too large γ, we may even set $\beta = 1$ meaning that only material losses and gains matter. The intuition is that material and fairness considerations point into the same direction for recipients. Thus, as long as gains are sufficiently discounted, recipients' punishment choices can be explained by looking at material factors alone. In contrast, Chapter 1 has highlighted the importance of fairness losses and gains ruling out the case $\beta = 1$. At the end of this sub-section, I address in more detail how the preferred specification in Chapter 1 relates to the preferred specification in this chapter.

The main problems faced by net-loss reciprocation in this chapter are recipients' reluctance to punish the principal for delegation if the agent or die has chosen the fair outcome and the lack of clear-cut support for Propositions 3 and 4. An explanation for recipients' reluctance to punish is *collective responsibility*: Applied by player *i*, collective responsibility means that *i* initially treats all other players (including Nature) as one individual by assessing his net loss from s_{-i} relative to its alternatives in S_{-i} using analogous methods to those proposed above. Doing so leads to a single net loss value that *i* assigns to his fellow players (excluding Nature) and to which he reciprocates. Collective responsibility can account for the fact that recipients treat A and B the same by if one of them (or the die) has chosen the fair outcome, but appears a lesser force in the domain of negative reciprocity, where there are significant differences in how recipients treat A and B. This asymmetry can be accounted for if the net-loss value established under collective responsibility acts as a *cap* on the individual net-loss values. Recipients can then differentiate between A and B if A has delegated and B chosen the unfair outcome because the relatively low net loss from delegation does not reach the cap. In contrast, if B has chosen the fair outcome, the cap limits the net loss from delegation.

BF conduct an econometric comparison of how well existing behavioural theories explain their punishment results. They find that both outcome-based models of inequality aversion (Fehr and Schmidt, 1999) and intention-based models of reciprocity (Dufwenberg and Kirchsteiger, 2004) face difficulties in accounting for their data. As previously mentioned, the former cannot explain why recipients systematically target their punishment towards the player who ultimately chooses the unfair outcome, while the latter cannot explain why recipients do not punish the principal if she delegates and the agent chooses the fair outcome. The latter (but not the former) difficulty is shared by my approach.

BF also propose a model of *responsibility*, which they show to overcome the aforementioned difficulties. The main idea is the following: A player's punishment increases in his or her responsibility for the unfair outcome, where responsibility is defined as the extent to which the player's actions increase the probability of the unfair outcome materialising given recipients' beliefs about the behaviour of the other players. Two important additional features are the "no harm, no foul" principle, which asserts that no responsibility for the bad outcome exists on the part of the principal if the agent or die has chosen the fair outcome. As a result, recipients do not want to punish the principal in this event. Net-loss reciprocation achieves the same result by allowing for collective responsibility. A second noteworthy feature is that the agent's responsibility is deemed to equal zero if the principal has chosen the unfair outcome directly. In contrast, my model asserts that recipients rely on their beliefs about the agent in this case. As discussed in the context of Proposition 3, BF's punishment data point into direction suggested by my approach (without there being a significant difference in punishments). In general, the responsibility model lacks the generality of my approach as a behavioural theory.

I conclude this section with a comparison between the preferred parameterisation derived in this chapter and Chapter 1. Chapter 1 suggests $\alpha = (1-2\beta)/(2-2\beta)$, which may violate the caps on α derived in this chapter. At the same time, $\alpha = (1-2\beta)/(2-2\beta)$ is due to just one of the four applications considered in Chapter 1. The restrictions derived from the other applications are consistent with the preferred specification in this chapter, while putting more structure on the model: Fairness must play a sufficiently large role ($\beta < 0.5$) and gains must not be fully discounted ($\gamma > 0$). In the delegation settings, material and fairness considerations point into the same direction, which means that fairness can be fully disregarded ($\beta = 1$) provided that γ is low enough. We can also have $\gamma = 0$ meaning that only losses count.[29]

2.4 Conclusion

This chapter has analysed in depth the power of delegation in the experimental study by Bartling and Fischbacher (2011) by proposing a model of net-loss reciprocation for *n*-player games including Nature and showing the model's ability to account for the punishment patterns sustaining the power of delegation. The main problems encountered by the model relate to situations of positive reciprocity, where the low net loss that the principal and agent as a

[29] The possibility of $\gamma = 0$ emerges as the "second-best" specification in Chapter 1. This is in line with the view that positive reciprocity is less important than negative reciprocity (Charness and Rabin, 2002).

collective impose on recipients appears to act as a cap on higher individual net losses. Given the success of the model in explaining punishments, it may be worthwhile applying net-loss reciprocation to other n-player settings. In doing so, one may want to put more structure on the v- and r-terms in the utility function, which in this chapter have been left unspecified up to qualitative features.

2.5 Appendix A: Proofs

The following lemma is useful in what follows. I focus on the case $nl^A > nl^B$. An analogous result holds for $nl^B > nl^A$.

LEMMA A1 Consider one of the punishment situations after the unfair outcome has been chosen. Let $Pr\left(p^A \leq x\right)$ and $Pr\left(p^B \leq x\right)$ denote the probability of recipients assigning $x \in \{0,1,...,6\}$ or fewer punishment points to player A and B, respectively, and suppose that recipients derive the net losses nl^A and nl^B where $nl^A > nl^B$. We have

$$Pr\left(p^A \leq x\right) < Pr\left(p^B \leq x\right)$$

for all $x \in \{0,1,...,6\}$ meaning that the distribution on punishment points assigned to A first-order stochastically dominates the distribution on punishment points assigned to B.

PROOF Consider first $x = 0$ and let $\hat{\Pi}^{0,j}$ be the set of outcomes consistent with assigning zero punishment points to player $j \in \{A,B\}$. We have

$$Pr\left(p^A = 0\right) < Pr\left(p^B = 0\right) \Leftrightarrow \sum_{\pi^A \in \hat{\Pi}^{0,A}} Pr\left(\pi^A\right) < \sum_{\pi^B \in \hat{\Pi}^{0,B}} Pr\left(\pi^B\right).$$

Notice that on each side of the inequality there is an equal number of terms owing to the symmetry of punishment options. I establish the inequality by a matching of left-hand side terms to right-hand side ones establishing the inequality for each pair. Recall that there is a one-to-one mapping from outcomes to punishment vectors $\left(p^A, p^B\right)$. As a result, we can refer to outcomes by their associated punishment vectors. Consider the following matching procedure: To a given $\pi^A = \left(0, p^B\left(\pi^A\right)\right)$ where $p^B\left(\pi^A\right) \in \{0,1,...,7\}$ denotes the punishment of B implied by $\pi^A \in \hat{\Pi}^{0,A}$, the procedure assigns $\pi^B = \left(p^B\left(\pi^A\right), 0\right)$ where $\pi^B \in \hat{\Pi}^{0,B}$. As a result, the matching is such that π^A and π^B are mirror images in punishment space, which implies that $nl_A\left(\pi^A\right) = nl_B\left(\pi^B\right)$ and $nl_B\left(\pi^A\right) = nl_A\left(\pi^B\right)$. Notice that the outcome implemented by the punishment vector $(0,0)$ appears on both sides of the inequality and hence cancels out. We can therefore limit attention to outcome pairs for which $p^B\left(\pi^A\right) = p^A\left(\pi^B\right) > 0$.

By Assumptions 2 and 3,

$$Pr\left(\pi^{A}\right)<Pr\left(\pi^{B}\right)\Leftrightarrow$$

$$exp\left[v\left(\pi_{i}^{A}\right)\right]\cdot exp\left[\sum_{j\in\{A,B\}}r\left(nl_{j}\left(\pi^{A}\right),nl^{j}\right)\right]<exp\left[v\left(\pi_{i}^{B}\right)\right]\cdot exp\left[\sum_{j\in\{A,B\}}r\left(nl_{j}\left(\pi^{B}\right),nl^{j}\right)\right].$$

Since π^{A} and π^{B} are mirror images, they yield the same material payoff to recipient i (namely, zero). Taking logarithms, the above inequality simplifies to

$$r\left(nl_{A}\left(\pi^{A}\right),nl^{A}\right)+r\left(nl_{B}\left(\pi^{A}\right),nl^{B}\right)<r\left(nl_{A}\left(\pi^{B}\right),nl^{A}\right)+r\left(nl_{B}\left(\pi^{B}\right),nl^{B}\right)\Leftrightarrow$$

$$r\left(nl_{A}\left(\pi^{B}\right),nl^{A}\right)-r\left(nl_{A}\left(\pi^{A}\right),nl^{A}\right)>r\left(nl_{B}\left(\pi^{A}\right),nl^{B}\right)-r\left(nl_{B}\left(\pi^{B}\right),nl^{B}\right).$$

By Assumption 3, the net loss of each player tracks the punishments points he or she receives. We therefore have $nl_{A}\left(\pi^{B}\right)=nl_{B}\left(\pi^{A}\right)>nl_{A}\left(\pi^{A}\right)=nl_{B}\left(\pi^{B}\right)$ since $p^{A}\left(\pi^{B}\right)=p^{B}\left(\pi^{A}\right)>0$ while $p^{A}\left(\pi^{A}\right)=p^{B}\left(\pi^{B}\right)=0$. From this together with $nl^{A}>nl^{B}$ and our assumptions on r, it follows that the left-hand side is larger than the right-hand side.

Consider next $x=1$. We have

$$Pr\left(p^{A}\leq 1\right)<Pr\left(p^{B}\leq 1\right)\Leftrightarrow$$

$$\sum_{\pi^{A}\in\hat{\Pi}^{0,A}}Pr\left(\pi^{A}\right)+\sum_{\pi^{A'}\in\hat{\Pi}^{1,A}}Pr\left(\pi^{A'}\right)<\sum_{\pi^{B}\in\hat{\Pi}^{0,B}}Pr\left(\pi^{B}\right)+\sum_{\pi^{B'}\in\hat{\Pi}^{1,B}}Pr\left(\pi^{B'}\right).$$

Notice that some outcomes appear on both sides of the inequality and hence cancel out. In punishment terms, these are $(0,0)$, $(1,1)$, $(0,1)$ and $(1,0)$. For the remaining outcomes, let the matching procedure for a given $\pi^{A}\in\hat{\Pi}^{0,A}$ or $\pi^{A'}\in\hat{\Pi}^{1,A}$ be the same as above. For every π^{A} and matched $\pi^{B}\in\hat{\Pi}^{0,B}$, $Pr\left(\pi^{A}\right)<Pr\left(\pi^{B}\right)$ has already been established. For every $\pi^{A'}$ and matched $\pi^{B'}\in\hat{\Pi}^{1,B}$, $Pr\left(\pi^{A'}\right)<Pr\left(\pi^{B'}\right)$ because $nl_{A}\left(\pi^{B'}\right)=nl_{B}\left(\pi^{A'}\right)>nl_{A}\left(\pi^{A'}\right)=nl_{B}\left(\pi^{B'}\right)$ owing to the fact that $p^{A}\left(\pi^{B'}\right)=p^{B}\left(\pi^{A'}\right)>1$ while $p^{A}\left(\pi^{A'}\right)=p^{B}\left(\pi^{B'}\right)=1$ and $nl^{A}>nl^{B}$.

The proof for all higher punishment levels, where an increasing number of outcomes cancel out, is analogous.■

The following lemma characterises punishment probabilities across situations. It focuses on the case $nl^A < nl^{A'}$ and $nl^B = nl^{B'}$. An analogous result holds for $nl^A = nl^{A'}$ and $nl^B < nl^{B'}$.

LEMMA A2 Consider any pair of punishment situations after the unfair outcome has been chosen. Let nl^A, nl^B, $nl^{A'}$ and $nl^{B'}$ be the net losses that recipients derive from A and B in the first and second situation, respectively, with $nl^A < nl^{A'}$ and $nl^B = nl^{B'}$. Moreover, let $Pr(p)$ and $Pr'(p)$ denote the probabilities with which recipients choose the punishment level $p^A = p \in \{0,1,...,7\}$ in the first and second situation, respectively. The following relationship holds for any $x \in \{0,1,...,6\}$ and $y = x+1$:

$$Pr'(y)\big/Pr(y) > Pr'(x)\big/Pr(x).$$

PROOF Let $\hat{\Pi}^p$ be the set of outcomes consistent with recipients imposing $p^A = p$ punishment points on A. We have

$$\frac{Pr'(y)}{Pr(y)} > \frac{Pr'(x)}{Pr(x)} \Leftrightarrow \frac{\sum_{\pi'\in\Pi^y} Pr'(\pi')}{\sum_{\pi'\in\Pi^y} Pr(\pi')} > \frac{\sum_{\pi\in\Pi^x} Pr'(\pi)}{\sum_{\pi\in\Pi^x} Pr(\pi)} \Leftrightarrow$$

$$\sum_{\pi'\in\Pi^y} Pr'(\pi')\sum_{\pi\in\Pi^x} Pr(\pi) > \sum_{\pi'\in\Pi^y} Pr(\pi')\sum_{\pi\in\Pi^x} Pr'(\pi).$$

We can establish the above inequality by showing

$$Pr'(\pi')Pr(\pi) > Pr(\pi')Pr'(\pi) \Leftrightarrow \frac{Pr'(\pi')}{Pr'(\pi)} > \frac{Pr(\pi')}{Pr(\pi)}$$

for every $\pi' \in \hat{\Pi}^y$ and $\pi \in \hat{\Pi}^x$. By our assumption on preferences,

$$\frac{Pr'(\pi')}{Pr'(\pi)} = \frac{exp\left[v(\pi_i')\right]\cdot exp\left[\sum_{j\in\{A,B\}} r\left(nl_j(\pi'),nl^{j'}\right)\right]}{exp\left[v(\pi_i)\right]\cdot exp\left[\sum_{j\in\{A,B\}} r\left(nl_j(\pi),nl^{j'}\right)\right]}$$

and

$$\frac{Pr(\pi')}{Pr(\pi)} = \frac{exp\left[v(\pi_i')\right]\cdot exp\left[\sum_{j\in\{A,B\}} r\left(nl_j(\pi'),nl^{j}\right)\right]}{exp\left[v(\pi_i)\right]\cdot exp\left[\sum_{j\in\{A,B\}} r\left(nl_j(\pi),nl^{j}\right)\right]}.$$

Insertion and logarithmation yields

$$\left[r\left(nl_A\left(\pi'\right),nl^{A'}\right)-r\left(nl_A\left(\pi\right),nl^{A'}\right)\right]-\left[r\left(nl_A\left(\pi'\right),nl^{A}\right)-r\left(nl_A\left(\pi\right),nl^{A}\right)\right]>$$

$$\left[r\left(nl_B\left(\pi'\right),nl^{B}\right)-r\left(nl_B\left(\pi\right),nl^{B}\right)\right]-\left[r\left(nl_B\left(\pi'\right),nl^{B'}\right)-r\left(nl_B\left(\pi\right),nl^{B'}\right)\right].$$

Since $nl^B = nl^{B'}$, the right-hand side equals zero, while the left-hand side is positive because of $nl_A\left(\pi'\right) > nl_A\left(\pi\right)$, $nl^{A'} > nl^A$ and our assumptions on r. The inequality $nl_A\left(\pi'\right) > nl_A\left(\pi\right)$ follows from Assumption 3, which ensures that assigning y punishment points to A imposes a higher net loss than assigning $x = y-1$ points regardless of how much B is punished.■

PROOF OF PROPOSITION 1

Consider first the situation where A has delegated and B has chosen the unfair outcome. According to the beliefs measured by BF, B chooses the unfair outcome with a probability of 0.34. Delegation thus implements the expected outcome

$$\left(0.34\cdot 9+0.66\cdot 5,0.34\cdot 9+0.66\cdot 5,0.34\cdot 1+0.66\cdot 5,0.34\cdot 1+0.66\cdot 5\right)=\left(6.36,6.36,3.64,3.64\right),$$

the alternatives being $\left(9,9,1,1\right)$ and $\left(5,5,5,5\right)$. Relative to $\left(9,9,1,1\right)$, recipients derive a material gain from delegation of $3.64-1=2.64$ and no fairness gain. Relative to $\left(5,5,5,5\right)$, they derive a material loss of $5-3.64=1.36$ and a fairness loss of $5-3.64-2.72\alpha/2=1.36-1.36\alpha$ yielding an overall loss of $1.36\left(1-\alpha+\alpha\beta\right)$. As a result, $nl^{A'}=1.36\left(1-\alpha+\alpha\beta\right)-2.64\gamma\beta$.

Player B choosing the unfair outcome after delegation implements $\left(9,9,1,1\right)$, the sole alternative being $\left(5,5,5,5\right)$. Recipients derive no gain from this. Their material loss is $5-1=4$ and their fairness loss is $5-1-4\alpha/2=4-4\alpha$ yielding an overall loss of $4\left(1-\alpha+\alpha\beta\right)$. Since there is no gain, we also have $nl^{B'}=4\left(1-\alpha+\alpha\beta\right)$. We have $nl^{B'}>nl^{A'}$ because $\beta>0$ by Assumption 3. Lemma A1 then implies that B's punishment dominates that of A.

Consider next the situation where A chooses the unfair outcome directly. This implements $\left(9,9,1,1\right)$ where $\left(5,5,5,5\right)$ and $\left(6.36,6.36,3.64,3.64\right)$ would have been available. As a result, recipients derive no gain from $\left(9,9,1,1\right)$. Their loss consists of their maximal loss with respect to the alternatives. Both their material and fairness loss are larger relative to $(5,5,5,5)$. Their material loss is $5-1=4$ and their fairness loss is $5-1-4\alpha/2=4-4\alpha$. As a result, their overall and net loss equals $nl^A=4(1-\alpha+\alpha\beta)=nl^{B'}$. As for B, his strategy remains unobserved if A chooses the unfair outcome. Using the beliefs about B measured by BF, recipients derive an expected net loss of $nl^B=0.34\cdot 4(1-\alpha+\alpha\beta)+0.66\cdot(-4\gamma\beta)=1.36(1-\alpha+\alpha\beta)-2.64\gamma\beta=nl^{A'}$ where $4(1-\alpha+\alpha\beta)$ and $-\gamma\beta 4$ are their net losses from B choosing the unfair and fair outcome. We have $nl^A>nl^B$ because $\beta>0$. By Lemma A1, A's punishment therefore dominates B's.

Finally, I show that $Pr(p^A=x)=Pr'(p^B=x)$ for every $x\in\{0,1,...,7\}$ where $Pr(p^A=x)$ and $Pr'(p^B=x)$ are the probabilities of recipients assigning x punishment points to A and B in the first and second situation, respectively. For $Pr(p^B=x)=Pr'(p^A=x)$, the argument is analogous. Let $\hat{\Pi}^{x,j}$ be the set of outcomes consistent with assigning x punishment points to $j\in\{A,B\}$. For every x, we have

$$Pr(p^A=x)=Pr'(p^B=x)\Leftrightarrow\sum\nolimits_{\pi^A\in\hat{\Pi}^{x,A}}Pr(\pi^A)=\sum\nolimits_{\pi^B\in\hat{\Pi}^{x,B}}Pr'(\pi^B).$$

Notice that there is an equal number of terms on each side owing to the symmetry of punishment options. I establish the equality by a matching of left-hand side terms to right-hand side terms showing the equality for each pair. The matching procedure is the same as in the proof of Lemma A1. To a given $\pi^A=(x,p^B(\pi^A))$, the procedure assigns $\pi^B=(p^B(\pi^A),x)$, expressing outcomes as punishment vectors.

Clearly, $nl_A(\pi^A)=nl_B(\pi^B)$ as well as $nl_B(\pi^A)=nl_A(\pi^B)$. For every matched pair, we have

$$Pr(\pi^A)=Pr'(\pi^B)\Leftrightarrow$$

$$exp\left[v(\pi_i^A)+\sum\nolimits_{j\in\{A,B\}}r\left(nl_j(\pi^A),nl^j\right)\right]\Big/\sum\nolimits_{\pi\in\hat{\Pi}}exp\left[v(\pi_i)+\sum\nolimits_{j\in\{A,B\}}r\left(nl_j(\pi),nl^j\right)\right]=$$

$$exp\left[v(\pi_i^B)+\sum\nolimits_{j\in\{A,B\}}r\left(nl_j(\pi^B),nl^{j'}\right)\right]\Big/\sum\nolimits_{\pi\in\hat{\Pi}}exp\left[v(\pi_i)+\sum\nolimits_{j\in\{A,B\}}r\left(nl_j(\pi),nl^{j'}\right)\right].$$

The equality of the numerators follows from $nl_A(\pi^A)=nl_B(\pi^B)$ and $nl_B(\pi^A)=nl_A(\pi^B)$ together with $nl^A=nl^{B'}$ and $nl^B=nl^{A'}$ and the fact that the material payoff to the recipient i is the same from the two outcomes given the matching procedure. The equality of the denominators follows from applying our matching procedure to all outcomes in $\hat{\Pi}$. To a given $\pi = \left(p^A(\pi), p^B(\pi)\right)$, we assign $\pi' = \left(p^B(\pi), p^A(\pi)\right)$. Proceeding thus, we can establish the equality of the denominators by establishing the equality of each matched pair, which follows from $nl_A(\pi) = nl_B(\pi')$ and $nl_B(\pi) = nl_A(\pi')$ together with $nl^A = nl^{B'}$ and $nl^B = nl^{A'}$ and the fact that the material payoff to the recipient is the same from the matched outcomes.■

PROOF OF PROPOSITION 2

We have $nl^B = nl^{B'} = 0$ because B is passive. Moreover, $nl^{A'} = 1.6(1-\alpha+\alpha\beta) - 2.4\gamma\beta$ because delegation entails $(0.4\cdot 9+0.6\cdot 5, 0.4\cdot 9+0.6\cdot 5, 0.4\cdot 1+0.6\cdot 5, 0.4\cdot 1+0.6\cdot 5) = (6.6,6.6,3.4,3.4)$, the alternatives being $(9,9,1,1)$ and $(5,5,5,5)$. Recipients derive no loss from $(6.6,6.6,3.4,3.4)$ relative to $(9,9,1,1)$. Relative to $(5,5,5,5)$, their material loss is $5-3.4=1.6$, while their fairness loss is $5-3.4-3.2\alpha/2 = 1.6-1.6\alpha$ implying an overall loss of $1.6(1-\alpha+\alpha\beta)$. Recipients derive no gain from $(6.6,6.6,3.4,3.4)$ relative to $(5,5,5,5)$. Relative to $(9,9,1,1)$, their gain is limited to their material gain of $3.4-1=2.4$ because $(9,9,1,1)$ is never fairer.

Secondly, $nl^A = 4(1-\alpha+\alpha\beta)$ follows from the fact that A choosing the unfair outcome implements $(9,9,1,1)$. Recipients derive no gain from this. Their loss consists of their maximal loss with respect to the two alternatives, namely, A choosing the fair outcome and delegation. Both their material and fairness loss are larger relative to A choosing the fair outcome. Their material loss is $5-1=4$, whereas their fairness loss is $5-1-8\alpha/2 = 4-4\alpha$. Their overall and net loss thus equals $4(1-\alpha+\alpha\beta)$.

We have $nl^A > 0 = nl^B$ and $nl^A > nl^{A'}$ because $\beta > 0$. We have $nl^{A'} > 0 = nl^{B'}$ because of the additional assumptions on α and γ. To see this, note that $1.6(1-\alpha+\alpha\beta) - 2.4\gamma\beta > 0$ is equivalent to $\alpha < (1-1.5\gamma\beta)/(1-\beta)$, which can only be satisfied by $\alpha \geq 0$ if $\gamma < 1/1.5\beta$. By Lemma A1, A's punishment dominates that of B in both situations. Moreover, Lemma A2 implies that A's punishment in the first situation dominates her punishment in the second.■

PROOF OF PROPOSITION 3

The results for D&P have been shown in the proof of Proposition 1. We have $nl^{B'} = 0$ because B is passive in NoD&P. Moreover, $nl^{A'} = 4(1-\alpha+\alpha\beta)$ follows from the fact that A only has the choice between the unfair and fair outcome. Recipients derive no gain from the unfair outcome relative to the fair one, while their loss is $4(1-\alpha+\alpha\beta)$.

We have $nl^{B} > nl^{B'}$ because $\alpha < (1.36-2.64\gamma\beta)/(1.36-1.36\beta)$, which can only be satisfied by $\alpha \geq 0$ if $\gamma < 1.36/2.64\beta$. By Lemma A2, B's punishment in the first situation dominates his punishment in the second situation.■

PROOF OF PROPOSITION 4

The results for D&P have been established in the proof of Proposition 1. In Asymmetric, given the beliefs measured by BF, delegation implements

$$(0.39\cdot 9+0.61\cdot 5, 0.39\cdot 9+0.61\cdot 5, 0.39\cdot 1+0.61\cdot 5, 0.39\cdot 1+0.61\cdot 5)=(6.56, 6.56, 3.44, 3.44),$$

the sole alternative being $(5,5,5,5)$. Consequently, recipients derive a material loss of $5-3.44=1.56$ from delegation. To this, a fairness loss of $5-3.44-3.12\alpha/2$ must be added yielding an overall and net loss of $nl^{A'}=1.56(1-\alpha+\alpha\beta)$. B choosing the unfair outcome after delegation implements $(9,9,1,1)$, the sole alternative being $(5,5,5,5)$. Recipients thus derive a loss of $nl^{B'}=4(1-\alpha+\alpha\beta)=nl^{B}$. We have $nl^{A'} > nl^{A}$ because $\beta > 0$. By Lemma A2, A's punishment in the second situation dominates her punishment in the first.■

3 BELIEF CHOICE UNDER PRIOR REGRET

3.1 Introduction

A fundamental issue in economics is how people incorporate new information about the uncertain world around them into their decision making. The standard assumption in this regard is that people are Bayesians, who hold a prior belief about the world and use Bayes' Rule to update it in response to new information. Yet, this mechanical model of belief revision has been increasingly called into question. On the one hand, updating mechanics that differ from Bayes' Rule have been proposed (see, e.g., Rabin and Schrag, 1999). Another idea that has found its way into the economics literature is the possibility that beliefs are actually *objects of choice* for decision makers (see, e.g., Brunnermeier and Parker, 2005). The thought is that in the face of new information beliefs are selected according to certain well-defined criteria before playing their usual role of decision weights in action choice.

This chapter follows the second approach by proposing a model of belief choice under prior regret. Contrary to existing models, the choice procedure is entirely *instrumental* meaning that the value of a given belief depends entirely on the action(s) that the belief implements. The instrumental value of beliefs has two components: *Objective performance* and *regret avoidance*. Objective performance refers to the expected utility of the implemented action(s) given the correct Bayesian posterior containing the new information. The issue of regret avoidance is more subtle: New information prompts one to abandon one's prior belief. Yet, giving up one's prior is not a good idea in all states of the world. To see this, suppose there is a unique action implemented by the prior. Clearly, for each alternative action, there exists at least one state where this *reference action* performs better. Otherwise, the reference action would be dominated and hence not rationalised by the prior in the first place. As a result, whenever new information invites one to adjust one's belief *and* this new belief implements a different action than the prior, one faces the dilemma of there being at least one state where holding on to one's prior would have been preferable. In such states, there is scope for (*ex post*) regret from having abandoned one's prior. Our model builds on the idea that the avoidance of such regret is a second goal in belief choice next to objective performance, with the ultimately chosen belief representing the best compromise between objective performance and regret avoidance.

To illustrate our approach in more detail, consider the well-known "Monty Hall problem" (Friedman, 1998; Palacios-Huerta, 2003): A decision maker D participating in a game show has a choice of opening one of three doors behind one of which there is a prize. After D has made a preliminary choice, the show master, who knows the location of the prize, opens one of the remaining doors. In doing so, he avoids the door concealing the prize. After this, D is given the opportunity either to stick to her preliminary choice or switch to the door that the show master did not open. As is well known, switching is best for D because the show master's avoidance of the door with the prize turns his choice of which door to open into an informative signal about the prize's location. Conditional on this signal, it is more likely that the prize is behind the door that the show master did not open than behind D's preliminarily chosen door. Yet, subjects in experiments show a strong propensity to stick to their original choice even after familiarising themselves with the task and receiving ample feedback for learning the rational decision (Friedman, 1998).

In terms of our model, we can explain this propensity by the trade off between objective performance and regret avoidance outlined above. From the viewpoint of objective performance, those beliefs about the prize's location are best that cause D to switch doors since switching yields the highest expected utility given the Bayesian posterior. While one of these beliefs is the Bayesian posterior itself, other beliefs are also optimal owing to the coarseness of the action space. Regret avoidance points into a different direction: The reference action for D's regret is given by her preliminary choice of door.[1] As a result, all beliefs causing D to switch doors entail a loss in the state of the world where the prize is behind the preliminarily chosen door. This is because upholding her preliminary choice is a best response in this state and switching is not. In contrast, all beliefs maintaining D's preliminary choice cause no such feeling of regret and are thus superior from the point of view of regret avoidance. All in all, if D pays sufficient heed to her expected regret from switching, where the expectation is calculated

[1] We assume that D strictly prefers her preliminary choice of door under her prior. Similar results obtain if D is indifferent between the three doors given her prior, but the argument is more involved.

using again the correct Bayesian posterior, all beliefs causing her to stick to her preliminary choice are optimal despite their lower objective performance.[2]

The example sheds light on the role of the Bayesian posterior in our model. While we call into question the mechanical updating of beliefs posited by Bayes' rule, we maintain that the Bayesian posterior, i.e., the objectively correct probability given the new information, is an important attractor in the belief choice procedure by providing the weight put onto the different states of Nature. More specifically, both the expected utility and expected regret associated with any given belief are calculated using the Bayesian posterior. Thus, our model could be interpreted as saying that people know the correct probabilities given some new piece of information, but only have the additional goal of regret avoidance in choosing their new belief. While this literal reading is possible, we tend to view the model as an "as if" representation of belief formation.[3] An advantage of letting the Bayesian posterior enter the target function in the way just described is that the model nests Bayesian decision making as a special case. As we show below, the behaviour of a decision maker with no sensitivity to regret is indistinguishable from that of a Bayesian.[4]

The example also raises the question why it is belief choice that we model rather than action choice because the same explanation for people's reluctance to switch doors could be generated by a model of action choice where preferences incorporate regret relative to one's previous action, but belief formation is entirely Bayesian. While the action choice model yields the same prediction about behaviour as our model in a one-period set-up, predictions may differ in a dynamic framework.[5] More importantly, we are interested in explaining how and when *beliefs*

---

[2] Again, these beliefs include her prior, but are not restricted to the latter because of the coarseness of the action space.

[3] Also, our model is in no way committed to having the Bayesian posterior as an attractor next to regret. Other specifications are conceivable (e.g. allowing for updating errors as in Rabin and Schrag, 1999).

[4] The example also lends additional plausibility to our reference point. The preliminary choice of door induced by the prior is a salient reference action, which translates into a clear feeling of regret from abandoning it in the state where maintaining it would have been preferable.

[5] In a two-period framework, if the decision maker exhibits "strong regret" (see Section 3.2.4) under the belief choice model, she is committed to replicating her first-period belief and hence action in the second period. Under the action choice model, assuming that the first-period action is the reference point for the second period, there are levels of regret satisfying strong regret such that D is not committed to choosing her first period action again in the second period. This follows from different links between periods:

deviate from the Bayesian benchmark, for which the action choice model is not suited. Also, the action choice model faces the conceptual problem that it may lead to choices that are first-order stochastically dominated given the decision maker's belief. The belief choice model avoids this problem because the predicted action is always optimal given the chosen belief.

We use our model of belief choice under prior regret to address several anomalies pertaining to information processing behaviour that have been documented by the experimental literature. One such anomaly is *overconfidence*, which we define as holding an inflated belief relative to the belief held by a standard Bayesian agent. Likewise, *underconfidence* consists of maintaining a depressed belief relative to the Bayesian benchmark. In a dynamic framework, we show that our model can generate both over- and underconfidence and characterise the conditions under which each should be expected to occur.

We also investigate *conservative and asymmetric information processing*. Conservative information processing means that belief revision after good or bad news goes into the same direction as Bayesian updating, but less so. Asymmetry broadly means that good news carries more weight than bad news. More specifically, it holds if a decision maker who is confident in some bad state of the world revises his belief more in response to news indicating some alternative good state than does a decision maker who is equally confident in the good state in response to news (of the same precision) indicating the bad state. Such behaviour is inconsistent with Bayesian updating, which implies the same absolute belief change in each case. Möbius et al (2012) provide evidence for conservative and asymmetric information processing in an experiment where subjects receive signals about their intelligence. In our model, the range of priors where the decision maker sticks to her prior in response to bad news (causing her to revise her belief less than a Bayesian) exceeds the same range for good news.

Another application of our model is a *preference for consistency*. By this, we mean a tendency to replicate one's previous action choice regardless of any new information one has received in the meantime (Falk and Zimmermann, 2011). As we show below, if people's regret relative to their prior is strong enough, they have an incentive to repeat the action rationalised by their prior across time without paying heed to new information that has become available.

Under belief choice, the first-period belief provides the reference point, but also the prior on which Bayesian updating takes place in the second period. Under action choice, the first-period action choice provides the reference point, but updating is not affected by it.

Starting with Akerlof and Dickens (1982), a few papers build on the idea that beliefs are *objects of choice* (rather than being formed in the mechanical fashion envisaged by Bayes' Rule). The target function for belief choice takes different forms: Brunnermeier and Parker (2005) take it to include "anticipatory utility", which is defined as the expected utility of the action(s) implemented by each belief calculated using the belief itself. This favours beliefs about which the decision maker can feel good *ex ante*. We discuss anticipatory utility in more detail below. Yariv (2005) defines a property of beliefs called "directional confidence". In her model, the decision maker values increasing her level of confidence in the state of the world she regards as most likely. Both models have a tendency to create extreme beliefs (or beliefs that are severely tilted towards one state), which is at odds with the evidence that belief revision is conservative in the sense of going in the direction of Bayesian updating, but less so. Also, beliefs enter the target function directly in these models. In contrast, our model of belief choice is the first where the choice procedure is fully instrumental in that beliefs enter the target function for belief selection only via the action(s) that they implement. We take this to be a parsimonious specification where departures from the standard approach have been minimised.

A second strand of literature does not view beliefs as objects of choice, but develops alternative belief mechanics intended to replace Bayes' Rule. Rabin and Schrag (1999) propose a model of "confirmatory bias" where Bayes' Rule is applied successfully with a probability of less than one in cases where one's signal and prior point into different directions. While confirmatory bias can explain conservative information processing, it faces difficulties in accounting for asymmetry. We return to this issue below. In the model of Epstein et al (2010), the decision maker processes new information in such a way that her adopted belief is a weighted average of the Bayesian posterior and her prior. Our model can be viewed as providing a micro-foundation for the decision maker placing positive weight on her prior. In general, however, the foundational idea of our model is that information processing cannot be dissociated from the decision problem in which the information is (likely to be) employed. This differentiates our approach from mechanical, "context-free" models of belief revision.[6]

---

[6] In a similar vein, the literature on "base rate neglect" (Kahnemann and Tversky, 1973) finds that people sometimes miscalculate conditional probabilities by equating the probability of an event conditional on a signal with the probability of the signal conditional on the event (see the evidence in Dohmen et al, 2009).

Thirdly, several papers do not assume that beliefs are chosen, but examine the idea that beliefs (formed according to Bayes' Rule) enter the utility function for action choice directly. For example, Köszegi (2006) studies "ego utility", where one's belief about one's ability or health is a direct carrier of utility. The model of "psychological expected utility" introduced by Caplin and Leahy (2001) lends itself to a similar interpretation (see Köszegi, 2003, and Barigozzi and Levaggi, 2010, for applications). These models have been used to study anomalous attitudes to information, in particular, conditions under which less accurate information is preferred.[7] Yet, since these models assume standard belief formation, they cannot account for the evidence on non-standard beliefs reported, e.g., in Möbius et al (2012). Also, their range of application tends to be limited to decision problems involving one's ego. In contrast, our model can be applied to situations where people's ego is manifestly not at stake, e.g., if they must guess some abstract state of Nature like in the Monty Hall problem.[8]

Finally, building on the idea of regret relative to one's prior, our model situates itself within the large literature on reference-dependent choice.[9] Several formalisations of reference-dependence exist. One example is the classic literature on regret, where the reference point is not a fixed action, but rather actions not taken by the decision maker (Loomes and Sugden, 1982; Bell, 1982). Our model uses a fixed reference point not previously considered in the literature, namely, the action(s) induced by one's prior. We view the prior as a natural and salient reference point for belief choice because the prior can be thought of as the *status quo* to which individuals have developed some kind of attachment. The importance of the status quo is already recognised in Bell (1982): "The level of regret felt may sometimes be related to the original status quo no matter what the outcome of foregone alternatives." (p. 980)[10]

---

[7] Another model that has been applied to such contexts is Köszegi and Rabin (2009). In their model, utility depends on changes in (rational) beliefs rather than beliefs themselves because belief changes inform about present and future consumption changes.

[8] A critique of the explanatory power of belief-dependent utility models can be found in Eliaz and Spiegler (2006).

[9] All existing models of reference-dependence cover action choice. In contrast, for a given belief, action choice in our model is reference-independent. It is only belief choice that is reference-dependent.

[10] The status quo interpretation of our reference point is also related to the literature on loss aversion (Kahneman and Tversky, 1979) and status quo bias (Samuelson and Zeckhauser, 1988). In the second case, the reference point is the status quo consumption from which one dislikes moving away because the induced consumption losses loom larger than same-sized consumption gains.

The remainder of this chapter is structured as follows: The next section develops the model and discusses some of its general properties. We explore the limiting cases of no regret, "strong regret" and certainty in one state of the world. We then study two applications pertaining to information processing behaviour. The final section concludes. All proofs are in Appendix A. Appendix B contains an experimental procedure for eliciting an individual's regret parameter.

3.2 The Model

3.2.1 Setup

The set of periods is given by $T = \{1,...,\tau\}$ where $\tau \geq 1$. In each period $t \in T$, the decision-maker (D) first receives an informative signal s_t from some finite space S that is identical across periods. The signal allows D to make inferences about the prevailing state of Nature $x \in X$. The space X is finite. After receiving her period signal, D must take an action $a_t \in A$. Like S, the space A is finite and the same in all periods.[11] In general, the utility action $a \in A$ generates in a period if the state of Nature is $x \in X$ is given by $u(x,a)$.[12] Yet, instead of choosing directly from A, we model D as choosing in each period from the set of possible beliefs about the state of Nature. These beliefs are collected in Q. The belief that D chooses in a period then determines her action in that period.

We make a few basic assumptions, whose motivation is technical: Firstly, the set Q of admissible beliefs about the state of Nature only contains probability distributions with full support on X. Secondly, signals are not fully revealing, i.e., we have $Pr(s_t|x) \in (0,1)$ for every $x \in X$, $t \in T$ and $s_t \in S$. These two assumptions imply that also all updated beliefs have full support on *X*. This is important below. Secondly, conditional on $x \in X$, signals are identically and independently distributed across periods Finally, there is *scope for regret*: There exist at

---

[11] This assumption holds true in the applications considered below, but could easily be relaxed.

[12] Hence, there are no "spillovers" between periods at the level of actions in the sense that the consequences of future action choices do not depend on what is chosen today. Such a model is studied by Eyster (2002). Yet, as discussed below, there are spillovers at the level of beliefs.

least one belief $\overline{q} \in Q$ and signal $\overline{s} \in S$ such that we do *not* have $A^*(\overline{q}) \subseteq A^*(\tilde{q}(\overline{q},\overline{s}))$ where $A^*(q)$ are the optimal actions given $q \in Q$, i.e., $A^*(q) = arg\,max_{a\in A} \sum_x q(x)u(x,a)$, and $\tilde{q}(\overline{q},\overline{s})$ is the Bayesian posterior, i.e., the probability distribution on X implied by Bayes' Rule as applied to $\overline{q}$ and $\overline{s}$. Intuitively, scope for regret means that there is at least one prior belief and signal such that there is some tension at the level of implemented actions between adopting the Bayesian posterior and sticking to one's prior. Imposing scope for regret thus only rules out uninteresting cases where the trade-off we aim to capture in this model is absent.

3.2.2 Belief Choice Problem

Consider now any period $t \in T$ and suppose that D holds the prior belief $q_{t-1} \in Q$ and has received the signal $s_t \in S$. If $t < \tau$, the target function for D's belief choice $q_t \in Q$ is

$$V_t(q_t|q_{t-1},s_t) = P_t(q_t|q_{t-1},s_t) + F_t(q_t|q_{t-1},s_t).$$

Its two parts express the present as well as future payoff associated with q_t. The influence of q_t on the future derives from q_t serving as D's prior in the next period $t+1$, where it influences her choice of q_{t+1} by providing the reference action for her regret in that period.[13] Through the same channel, q_{t+1} influences D's belief choice in period $t+2$. As a result, q_t indirectly influences the belief choice in $t+2$ via its influence on period $t+1$ and so on for all subsequent periods. In contrast, if $t = \tau$, there is no future that can be influenced by q_t. In this case, the target function reduces to $V_t = P_t$.

As mentioned above, the present and future payoff of q_t has two components: On the one hand, *objective performance*, which refers to the expected utility of q_t given the correct Bayesian posterior. On the other hand, *regret avoidance*, which refers to the extent to which q_t minimises D's expected regret relative to her prior.

The first term of the target function formalises this idea for the current period t. It is given by

$$P_t(q_t|q_{t-1},s_t) = \sum_x \tilde{q}(x|q_{t-1},s_t)\left[U(x,q_t) + r\left(U(x,q_t) - U(x,q_{t-1})\right)\right]$$

[13] Thus, D is *sophisticated* in the sense that she realises that adopting some belief today will lead her to grow used to this belief, which entails regret when she abandons this belief tomorrow.

where $\tilde{q}(x|q_{t-1},s_t)$ is the probability assigned to state x by the Bayesian posterior, i.e., the probability of x according to Bayes' rule given the prior q_{t-1} and signal s_t. Moreover, $U(x,q_t)$ is the utility generated by q_t in the state of the world x. As a result, the first part of P_t, namely, $\sum_x \tilde{q}(x|q_{t-1},s_t)U(x,q_t)$, represents the expected utility of q_t under the Bayesian posterior and therefore the objective performance of q_t in period t.

At the same time, $U(x,q_t)$ remains to be defined. The idea here is that D's choice of belief q_t in period t pins down her action choice in that period, the link between the two being provided by standard expected utility maximisation, i.e., by the maximisation of $\sum_x q_t(x)u(x,a)$ with respect to $a \in A$.[14] Assuming that D randomises uniformly among her preferred actions, $U(x,q_t)$ assigns to q_t the implied level of (expected) utility in x. Formally, letting $A^*(q_t) = arg\,max_{a\in A} \sum_x q_t(x)u(x,a)$ denote the set of D's preferred actions given q_t, the probability that D chooses a if she adopts q_t is

$$\pi(a,q_t) = \begin{cases} 1/\left|A^*(q_t)\right| & \text{if } a \in A^*(q_t) \\ 0 & \text{otherwise} \end{cases}$$

Thus, $\pi(a,q_t)$ assigns a probability of $1/\left|A^*(q_t)\right| > 0$ if a is a maximiser of D's expected utility under q_t and zero probability otherwise. This allows us to define the utility generated by q_t in x as $U(x,q_t) = \sum_a u(x,a)\pi(a,q_t)$.

Since D trades off objective performance against regret avoidance, the second part of P_t expresses the expected regret associated with q_t where the reference point is provided by D's prior q_{t-1}. The regret function r is defined as

$$r(y) = \begin{cases} \lambda y & \text{if } y < 0 \\ 0 & \text{otherwise} \end{cases}$$

where $y \in \mathbb{R}$ is the relevant utility difference in state x and the parameter λ satisfies $\lambda \geq 0$.[15] Hence, D experiences regret from q_t in a given state x if and only if $U(x,q_t)$ falls short of

[14] Of course, other models than expected utility maximisation could be used at this point.

[15] For simplicity, we do not allow for any feelings of *rejoice* that D might experience. Including rejoice would not alter the conclusions from our model (as long as regret looms larger than rejoice). In the literature on loss aversion, several papers ignore gains focusing on losses only (see, e.g., Herweg and Mierendorff, 2013, and Herweg and Schmidt, 2013).

$U(x,q_{t-1})$, which is the utility generated in x by her prior q_{t-1}. The expected regret associated with q_t is calculated using again the Bayesian posterior. Clearly, setting $q_t = q_{t-1}$ yields zero expected regret in period t, while other beliefs may give rise to positive regret.

The future payoff function F_t repeats the same exercise for all future periods. Similar to before, D trades off two goals, namely, maximising expected utility given future posteriors (objective performance) versus minimising expected regret relative to future priors (regret avoidance). Formally, the future for D consists of the different future signal realisations. The set H_z collects the *future signal histories* from period $t+1$ till period $z \in \{t+1,...,\tau\}$. The future payoff of q_t is then given by

$$F_t\left(q_t \middle| q_{t-1}, s_t\right) =$$

$$\sum_{z\in\{t+1,...,\tau\}} \sum_{h_z \in H_z} Pr\left(h_z \middle| q_{t-1}, s_t\right) \sum_x \tilde{q}\left(x \middle| q_{t-1}, s_t, h_z\right) \left[U\left(x, q_t, h_z\right) + r\left(U\left(x, q_t, h_z\right) - U\left(x, q_t, \overleftarrow{h}\left(h_z\right)\right)\right)\right].$$

Like before, $\tilde{q}\left(x \middle| q_{t-1}, s_t, h_z\right)$ is the posterior probability assigned to state x, which now also conditions on the signals contained in the history $h_z \in H_z$. Secondly, $U\left(x, q_t, h_z\right)$ is the utility generated by q_t in x after h_z, whereas $U\left(x, q_t, \overleftarrow{h}\left(h_z\right)\right)$ is the same utility after $\overleftarrow{h}\left(h_z\right)$, which is the history immediately preceding h_z. Formally, the family of functions $\overleftarrow{h}$ assign to a given h_z the history yielded by eliminating the most distant signal in h_z. For example, if $h_z = \left(s_{t+1}, s_{t+2}\right)$, we have $\overleftarrow{h}\left(h_z\right) = s_{t+1}$. A special case arises if $z = (t+1)$. We then have $\overleftarrow{h}\left(h_z\right) = \varnothing$ for all $h_z \in H_z$ where $\varnothing$ is the *empty history*. Comparing $U\left(x, q_t, h_z\right)$ to $U\left(x, q_t, \overleftarrow{h}\left(h_z\right)\right)$ establishes D's regret in x after h_z via the regret function. D's regret is positive if and only if ignoring the last signal in h_z would have yielded her higher expected utility than acting upon it.

Similar to before, we define $U\left(x, q_t, h\right) = \sum_a u\left(x, a\right) \pi\left(a, q_t, h\right)$ where $\pi\left(a, q_t, h\right)$ expresses the probability of choosing a triggered by q_t after the history $h \in H_t = \bigcup_{z\in\{t+1,...,\tau\}} H_z \cup \{\varnothing\}$. As already mentioned, the influence of q_t on D's action choice after $h = h_z$ derives from q_t serving as D's prior in period $t+1$, where it pins down her belief choice given the signal s_{t+1} contained in h_z. The belief chosen by D in this situation then supplies the prior for her belief choice in $t+2$ given the signal s_{t+2} again taken from h_z, which pins down her belief choice after s_{t+1} and s_{t+2} and so on for all remaining signals in h_z. Through this mechanism, q_t ends up pinning down D's belief after h_z, which determines her action choice.

In this context, a subtle issue arises: Suppose that for a given prior q_{z-1} and signal s_z, D finds several beliefs optimal. How does she (predict she will) choose among them? At this stage, we do not specify how D chooses among her optimal beliefs $Q_z^*(q_{z-1},s_z)$, but simply assume that she follows some rule $f_z(q_{z-1},s_z)$ for selecting among them. Below, when studying applications, we put more structure on $f_z(q_{z-1},s_z)$.

We can now give a formal definition of $\pi(a,q_t,h)$. If $h=\varnothing$, we have $\pi(a,q_t,h)=\pi(a,q_t)$. This scenario only arises when considering period $t+1$, where we have $\overleftarrow{h}=\varnothing$. Letting $\pi(a,q_t,\varnothing)=\pi(a,q_t)$ simply means that the reference point for D's regret in $t+1$ is the action choice distribution implied by her current belief q_t. In contrast, if we have $h=h_z=(s_{t+1},...,s_z)$ for some $z\in\{t+1,...,\tau\}$, the definition of $\pi(a,q_t,h)$ is *recursive*. Let

$$\pi(a,q_t,s_{t+1},...,s_z)=w\Big(\big\{\pi(a,q_{t+1},s_{t+2},...,s_z):q_{t+1}\in Q_{t+1}^*(q_t,s_{t+1})\big\}\Big|f_{t+1}(q_t,s_{t+1})\Big)$$

where $Q_{t+1}^*(q_t,s_{t+1})=arg\,max_{q_{t+1}\in Q}V_{t+1}(q_{t+1}|q_t,s_{t+1})$ are D's preferred beliefs in period $t+1$ given q_t and the signal s_{t+1} from h_z. As a result, $\{\pi(a,q_{t+1},s_{t+2},...,s_z):q_{t+1}\in Q_{t+1}^*(q_t,s_{t+1})\}$ contains the different probabilities of choosing a after h_z consistent with D choosing an optimal belief in $t+1$ given q_t. Furthermore, w is a function that yields the weighted average of these probabilities given the rule $f_{t+1}(q_t,s_{t+1})$ for choosing among the optimal beliefs in $Q_{t+1}^*(q_t,s_{t+1})$. The definition is recursive because it builds on $\pi(a,q_{t+1},s_{t+2},...,s_z)$. We close it by defining

$$\pi(a,q_{z-1},s_z)=w\Big(\big\{\pi(a,q_z):q_z\in Q_z^*(q_{z-1},s_z)\big\}\Big|f_z(q_{z-1},s_z)\Big).$$

One then obtains $\pi(a,q_{z-2},s_{z-1},s_z)=w\Big(\big\{\pi(a,q_{z-1},s_z):q_{z-1}\in Q_{z-1}^*(q_{z-2},s_{z-1})\big\}\Big|f_{z-1}(q_{z-2},s_{z-1})\Big)$ etc.

In practice, whenever the period under consideration is not the final period, i.e., if $z\neq\tau$, one must first go back to the final period τ for establishing the target functions for earlier periods (*backward induction*).

3.2.3 No Regret

In this and the next two sub-sections, we discuss some general properties of our model. The first "polar" scenario that we consider is the case of "no regret", i.e., of $\lambda=0$. We have

PROPOSITION 1 Suppose that the decision maker has no regret, i.e., $\lambda = 0$. In every period $t \in T$ and for every signal $s_t \in S$ and prior $q_{t-1} \in Q$, it is optimal for the decision maker to choose the Bayesian posterior, i.e., $\tilde{q}(q_{t-1}, s_t) \in Q_t^*(q_{t-1}, s_t)$. Moreover, all other optimal beliefs must mimic $\tilde{q}(q_{t-1}, s_t)$ in the sense that $q_t \in Q_t^*(q_{t-1}, s_t)$ if and only if $A^*(q_t) \subseteq A^*(\tilde{q}(q_{t-1}, s_t))$ and $\{a \in A : \pi(a, q_t, h_z) > 0\} \subseteq A^*(\tilde{q}(q_{t-1}, s_t, h_z))$ for all $z \in \{t+1, ..., \tau\}$ and $h_z \in H_z$.

Proposition 1 confirms a desirable feature of our model. Intuitively, if D is not moved by regret, the behaviour induced by her belief choices should come as close as possible to the behaviour of a standard Bayesian agent who forms his belief according to Bayes' rule and chooses his actions so as to maximise his expected utility. Proposition 1 makes clear that this is the case: In any period t after any signal s_t and for any prior q_{t-1}, the Bayesian posterior $\tilde{q}(q_{t-1}, s_t)$ is an optimal belief for D if $\lambda = 0$. Moreover, any other optimal belief must mimic the Bayesian posterior in the sense that it triggers the same action choices (or a subset thereof) both in the current period and after all future signal histories.

The proof of Proposition 1 uses backward induction (see Appendix A). Consider therefore the final period. Clearly, adopting $\tilde{q}(q_{\tau-1}, s_\tau)$ is optimal there because it leads D to choose one of the actions in $A^*(\tilde{q}(q_{\tau-1}, s_\tau))$, which are nothing but the maximisers of $\tilde{U}(a|q_{\tau-1}, s_\tau)$, i.e., D's expected utility under the Bayesian posterior. Consequently, for other beliefs to be optimal, they must imitate the posterior by inducing D to choose only actions in $A^*(\tilde{q}(q_{\tau-1}, s_\tau))$. As a result, for any given $q_{\tau-1}$ and s_τ, D implements only maximisers of $\tilde{U}(a|q_{\tau-1}, s_\tau)$ regardless of which optimal belief she adopts.

Consider next period $\tau - 1$. For the same reason as before, the Bayesian posterior $\tilde{q}(q_{\tau-2}, s_{\tau-1})$ maximises D's payoff in $\tau - 1$, whereas her expected future payoff is given by

$$\sum_{s_\tau \in S} Pr(s_\tau | q_{\tau-2}, s_{\tau-1}) \sum_a \pi(a, q_{\tau-1}, s_\tau) \tilde{U}(a | q_{\tau-2}, s_{\tau-1}, s_\tau).$$

As just observed, D implements only maximisers of $\tilde{U}(a|q_{\tau-1}, s_\tau)$ after every $q_{\tau-1}$ and s_τ. Thus, adopting $q_{\tau-1} = \tilde{q}(q_{\tau-2}, s_{\tau-1})$ leads D to implement only maximisers of $\tilde{U}(a|\tilde{q}(q_{\tau-2}, s_{\tau-1}), s_\tau)$, which equals $\tilde{U}(a|q_{\tau-2}, s_{\tau-1}, s_\tau)$ because $\tilde{q}(\tilde{q}(q_{\tau-2}, s_{\tau-1}), s_\tau) = \tilde{q}(q_{\tau-2}, s_{\tau-1}, s_\tau)$ by Bayes' Rule. This establishes that $\tilde{q}(q_{\tau-2}, s_{\tau-1})$ also maximises D's future payoff and is hence optimal. Other optimal beliefs must imitate the posterior by implementing only actions in $A^*\left(\tilde{q}\left(q_{\tau-2},s_{\tau-1}\right)\right)$ and $A^*\left(\tilde{q}\left(q_{\tau-2},s_{\tau-1},s_\tau\right)\right)$ for every $s_\tau \in S$. The argument for all $t<\tau-1$ is analogous.

3.2.4 Strong Regret

The other polar scenario is what we refer to as "strong regret".

DEFINITION 1 The decision maker exhibits *strong regret* in period $t \in T$ if and only if for every $s_t \in S$ and $q_{t-1} \in Q$ we have $q_t \in Q_t^*\left(q_{t-1},s_t\right)$ if and only if $A^*\left(q_t\right)=A^*\left(q_{t-1}\right)$. If she exhibits strong regret in every period $t \in T$, she has *overall strong regret*.

Strong regret mandates that after every signal D finds exactly those beliefs optimal that are action-equivalent to her prior no matter what the latter may be. We have

PROPOSITION 2 For every period $t \in T$, there is a threshold level of regret $\bar{\lambda}(t)>0$ such that the decision maker exhibits strong regret in t for all $\lambda>\bar{\lambda}(t)$. Moreover, for all $t<\tau$, we have $\bar{\lambda}(t)=(1+\tau-t)\bar{\lambda}(\tau)$.

Thus, in any given period, a sufficiently large λ fulfils the strong regret requirement for that period. Moreover, the threshold $\bar{\lambda}(t)$ for any period $t<\tau$ is proportional to the last-period threshold $\bar{\lambda}(\tau)$, where the strong regret requirement is most easily fulfilled. In the second to last period, we have $\bar{\lambda}(\tau-1)=2\bar{\lambda}(\tau)$ meaning the threshold doubles compared to the last period, while it trebles in the third to last period and so on for all earlier periods. Consequently, the requirements on λ are highest for overall strong regret.

For an intuition, focus first on the case $t=\tau$. As Lemma A1 in the Appendix makes clear, all beliefs q_τ that satisfy $A^*\left(q_\tau\right) \neq A^*\left(q_{\tau-1}\right)$ entail regret in at least one state of the world. Hence, for a high enough λ, all beliefs satisfying $A^*\left(q_\tau\right) \neq A^*\left(q_{\tau-1}\right)$ are sub-optimal because regret avoidance becomes D's dominant concern. Consider next period $\tau-1$ and suppose that λ is high enough for strong regret to hold in period τ. Accordingly, D anticipates that she will choose any $q_{\tau-1}$ again in τ irrespective of s_τ thereby avoiding any regret in τ.[16] Since it will be

[16] Or, more precisely, she anticipates that she will adopt some action-equivalent belief to $q_{\tau-1}$.

chosen again, the absolute expected utility of a given $q_{\tau-1}$ equals $2\sum_x \tilde{q}\left(x|q_{\tau-2}, s_{\tau-1}\right)U\left(x, q_{\tau-1}\right)$. This means that every belief is now assigned twice as high absolute expected utility as in the case $t=\tau$. As a result, for a given prior to perform better than those beliefs that yield D higher absolute expected utility, λ must rise twice as high. Since this holds for every prior and signal, we have $\overline{\lambda}\left(\tau-1\right)=2\overline{\lambda}\left(\tau\right)$. An analogous argument applies to earlier periods. For example, $\overline{\lambda}\left(\tau-2\right)=3\overline{\lambda}\left(\tau\right)$ and in general $\overline{\lambda}\left(t\right)=\left(1+\tau-t\right)\overline{\lambda}\left(\tau\right)$.

Proposition 2 has the following

COROLLARY 1 Suppose that strong regret holds in period $t\in T$. For every prior $q_{t-1}\in Q$, the decision maker chooses an action from $A^*\left(q_{t-1}\right)$ in t and all future periods $z\in\left\{t+1,...,\tau\right\}$ irrespective of the signals that she receives.

The corollary follows from the relationship $\overline{\lambda}\left(t\right)=\left(1+\tau-t\right)\overline{\lambda}\left(\tau\right)$, which implies that the threshold for strong regret falls monotonically in the number of remaining periods and is hence highest in period t, second highest in $t+1$ etc. Thus, if strong regret holds in t, it holds in all future periods leading D to choose an action-equivalent belief to q_{t-1} from t onward.

As a result, strong regret provides a potential explanation for *preferences for consistency* (Falk and Zimmermann, 2011). More specifically, if $A^*\left(q_{t-1}\right)$ is singleton and strong regret holds in period t, D chooses the same action from t onward. Thus, our model identifies a new reason why people want to stick to previous behaviour in spite of receiving new information. This explanation is complementary to the explanations discussed in the literature like the desire to signal ability through consistent behaviour (see Falk and Zimmermann, 2011).

Proposition 2 assumes D to be perfectly forward-looking in the sense that she takes into account the impact of her current belief choice on all future periods. Yet, the logic that underpins it could also be applied to a slightly different analysis, namely, if D's level of foresight varies.[17] For example, if D is *myopic* meaning that she treats her decision problem in every period as a one-period problem, it is easiest for her to achieve strong regret in every period because the threshold for her regret is always as small as possible, namely, $\overline{\lambda}\left(\tau\right)$. An implication of this is that a preference for consistency, i.e., uniform behaviour across time, is

[17] We leave a formalisation of this for future work. We address benefits from myopia under regret in the following section.

most easily achieved under myopia and in general more easily with less foresight because it becomes easier to follow the lead of one's prior in any given period.

3.2.5 Certainty

While the bulk of this chapter restricts attention to beliefs with full support on the state space X and signals that are not fully revealing, we make an exception in this sub-section in order to explore a limiting case.[18] More specifically, we address what happens if D is *certain* of one particular state under the Bayesian posterior e.g. because of a fully revealing signal. We have

PROPOSITION 3 Suppose that in some period $t \in T$ the Bayesian posterior for some prior $q_{t-1} \in Q$ and signal $s_t \in S$ puts all probability mass on one state, i.e., there is a state $\overline{x} \in X$ such that $\tilde{q}\left(\overline{x} \middle| q_{t-1}, s_t\right) = 1$. In this case, the Bayesian posterior is optimal irrespective of the decision maker's regret, i.e., we have $\tilde{q}\left(q_{t-1}, s_t\right) \in Q_t^*\left(q_{t-1}, s_t\right)$ for all $\lambda \geq 0$.

Proposition 3 makes clear that D's prior being an attractor for her belief choice hinges on there being some lingering doubt about the state of the world. It is not a goal in itself for D to match her belief to her prior. Rather, she aims to avoid expected regret from giving up the latter. This goal is traded off against her second goal, namely, to achieve maximal expected utility under the Bayesian posterior. Yet, if she is certain of a particular state $\overline{x}$ under the posterior, the conflict between the two goals disappears. By maximising utility in $\overline{x}$, D both eliminates any regret she may have relative to her prior and maximises expected utility given the posterior. One way for D to do this is to adopt the posterior itself because this causes her to choose a maximiser of $1 \cdot u\left(\overline{x}, a\right)$.[19] In terms of the Monty Hall problem discussed in the Introduction, if the show master accidentally opens the door concealing the price, it is optimal for D to "believe her eyes" and assign probability one to the prize being where she perceives it to be.

---

[18] As stated above, the motivation for these assumptions is technical. Even maintaining them, we can approximate beliefs without full support and fully revealing signals arbitrarily closely.

[19] Of course, other beliefs are also optimal as long as they induce the same action(s) as the posterior.

The same logic applies if there are future periods: By choosing the Bayesian posterior (which puts all probability mass on $\bar{x}$) as her current belief, D ensures that her future decision making reflects her being certain of $\bar{x}$ under today's posterior. This is because her current belief serves as the next period's prior. By Bayes' Rule, if the prior puts all probability mass on $\bar{x}$, so does the posterior, which then provides the right incentives for D's future belief choices.

3.3 Applications

We now study applications of the model introduced in the preceding section. To keep the analysis tractable, we rely throughout on

ASSUMPTION 1 The state space is $X=\{0,1\}$ and the action space $A=\{0,1\}$ with $u(x=1,a=1)=1$, $u(x=1,a=0)=\alpha$, $u(x=0,a=0)=\beta$ and $u(x=0,a=1)=0$ where $0\leq\alpha<1$ and $0<\beta\leq 1$. The signal space is $S=\{0,1\}$. Signals are informative and symmetric meaning that $Pr(s=0|x=0)=Pr(s=1|x=1)=\sigma$ with $\sigma\in(0.5,1)$.

We call $x=1$ the *good state* and $x=0$ the *bad state*. The state corresponds to some feature of either the external world or D's own personality. In the second case, where D's "ego" is at stake, think of the state as something like D's ability, intelligence or health. Since the state space is binary, we can express beliefs as $q(x=1)=q$ and $q(x=0)=1-q$, the set of possible beliefs being $Q=(0,1)$. As for the available actions, we refer to $a=1$ as the *bet on the good state* and to $a=0$ as the *bet on the bad state*. In each state, the bet matching the state is preferred since $1>\alpha$ and $\beta>0$. Moreover, a bet that matches the state is at least as good in the good state as in the bad state (since $1\geq\beta$), while a bet that does not match the state is at least as bad in the bad state as in the good state (since $0\leq\alpha$).[20] In the limiting case $\alpha=0$ and $\beta=1$, payoffs are symmetrical, which means that D aims to "coordinate" with Nature without favouring one state over the other. Below, a key role is played by the belief $q\in Q$ for which the expected utility of betting on the good state equals that of betting on the bad state, i.e., the

[20] These utility differences can be driven by differences in underlying material consequences and/or state-dependent evaluation of the latter. Specific interpretations are given below.

threshold belief that makes D indifferent between the two actions in A. This belief is given by $\beta/(1-\alpha+\beta)$. In what follows, we denote it by q^A. Given Assumption 1, we have $0 < q^A < 1$.

Several interpretations of our framework are possible. For example, $x=1$ could be good weather and $x=0$ bad weather. In this case, the bet on the good state consists of D taking actions appropriate for good weather ("taking no umbrella"), while the bet on the bad state amounts to preparing for bad weather ("taking an umbrella"). If the weather is good, it is better to take no umbrella ($1>\alpha$), while the converse holds if the weather is bad ($\beta > 0$). Also, no umbrella in good weather is at least as good as an umbrella in bad weather ($1 \geq \beta$), while no umbrella in bad weather is at least as bad as an umbrella in good weather ($0 \leq \alpha$).

Secondly, $x=1$ could be high and $x=0$ low ability with the bet on the good state representing a human capital investment at some monetary cost (like obtaining a university degree) and the bet on the bad state corresponding to no investment.[21] Investing is normalised to yield a utility of one in the good state (where D is able and the investment hence worthwhile) and of zero in the bad state (where D is unable). The alternative of not investing and therefore not incurring its monetary cost yields α in the good state and β in the bad state. D prefers to invest if she is able ($1>\alpha$) and not to invest if she is unable ($\beta > 0$). Also, going beyond Assumption 1, $\alpha \geq \beta$ seems probable in this context because D is likely to rank not investing and being able at least as high as not investing and being unable.

Finally, the state of the world could be D's attractiveness to another agent E, which is high in the good and low in the bad state. The bet on the good state now consists of D approaching E, while the bet on the bad state corresponds to D abstaining from making advances to E. A conquest (D has approached E and her attractiveness is high) yields a utility of one, while rejection (D has approached E, but her attractiveness is low) generates zero utility. The alternative of not approaching E yields $\alpha < 1$ in the good and $\beta > 0$ in the bad state. Given that D's ego is at stake, $\alpha \geq \beta$ again seems plausible. In the following sub-section, we study the processing of "ego-sensitive" information pertaining to such things as attractiveness, ability or intelligence in more detail.

[21] This interpretation is inspired by the model studied in Möbius et al (2012), which represents a special case of Assumption 1.

3.3.1 Conservative and Asymmetric Information Processing

We now use our model to analyse the processing of ego-sensitive information. We aim to account for two co-existing phenomena: Firstly, *conservatism*, which means that individuals' belief revision in response to a good or bad (for their ego) signal has the same sign as under Bayesian updating, but is less pronounced. Secondly, *asymmetry*, which means that individuals with a given level of confidence in a good (for their ego) state react less, i.e., more conservatively, to a signal indicating an alternative bad state than individuals who are equally confident in the bad state react to a signal of the same precision indicating the good state. Asymmetry is at odds with Bayesian updating, which implies the same absolute belief change in the two cases. Clean experimental evidence for both phenomena is provided in the study of Möbius et al (2012), where subjects must process signals about their intelligence.

As no additional insight is generated by allowing for multiple periods, we restrict the analysis to a single period, i.e., $\tau = 1$. We use the modelling framework laid down in Assumption 1. For present purposes, we characterise the belief that D adopts from her optimal beliefs $Q_1^*(q_0, s_1)$. Having such a point prediction facilitates our subsequent analysis.[22]

ASSUMPTION 2 The decision maker's *adopted belief* $\hat{q}_1(q_0, s_1) \in Q_1^*(q_0, s_1)$ satisfies $\hat{q}_1 = \tilde{q}$ if $\tilde{q} \in Q_1^*$. If $\tilde{q} \notin Q_1^*$ and $q_0 \in Q_1^*$, we have $\hat{q}_1 = q_0$.

The assumption asserts that D uses a lexicographic tie-breaking rule for selecting among her optimal beliefs: If the Bayesian posterior is optimal, she adopts it. If the posterior is not optimal, but the prior is, the latter is adopted. The case where neither posterior nor prior is optimal does not arise below and need hence not be specified.

We first analyse the "good news scenario" where D has received the signal $s_1 = 1$. Let q_0 be the prior probability assigned to the good state and define

$$q_0^T = min\left\{\beta(1-\sigma)(1+\lambda)/\left((1-\alpha)\sigma+\beta(1-\sigma)(1+\lambda)\right), q^A\right\}$$

[22] The alternative approaches to belief choice discussed below also yield point predictions. Having a point prediction ourselves facilitates comparisons.

as D's *threshold prior* for $s_1 = 1$. It is also useful to define $q_0^B = \beta(1-\sigma)/((1-\alpha)\sigma + \beta(1-\sigma))$, which is the prior at which a Bayesian agent is indifferent between betting on the good and bad state. The belief choice of D is characterised in

LEMMA 1 Suppose that $\lambda > 0$ and $s_1 = 1$. If $q_0 < q_0^T$, the decision maker prefers beliefs making her bet on the bad state, i.e., $Q_1^* = (0, q^A)$. If $q_0 > q_0^T$, she prefers beliefs making her bet on the good state, i.e., $Q_1^* = (q^A, 1)$. Her adopted belief satisfies $\hat{q}_1 = q_0$ if $q_0^B \leq q_0 < q_0^T$ and $\hat{q}_1 = \tilde{q}$ if $q_0 < q_0^B$ or $q_0 > q_0^T$.

Thus, D adopts the Bayesian posterior except for priors between q_0^B and q_0^T, where she adopts her prior.[23] As a result, our model can account for conservatism because average belief revision across priors is positive, but smaller than what is implied by Bayesian updating.

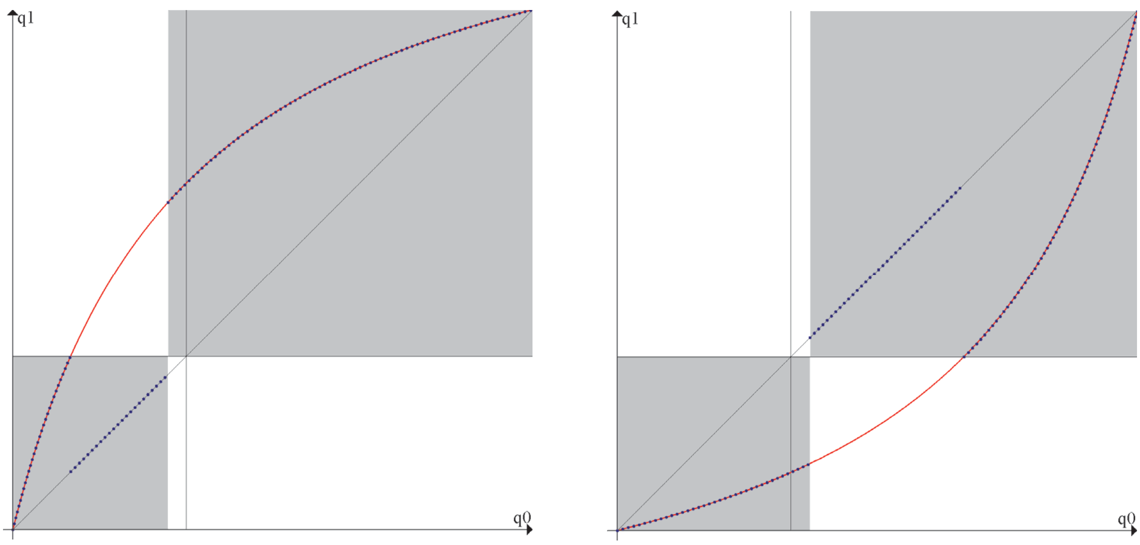

FIGURE 3.1: Optimal beliefs (grey areas) and adopted beliefs (blue dotted line) after $s_1 = 1$ and $s_1 = 0$ as a function of the prior q_0 given $\alpha = 0.2$, $\beta = 0.4$, $\sigma = 0.8$ and $\lambda = 2.4$

The intuition for Lemma 1 is the following: Recall that $q^A = \beta/(1-\alpha+\beta)$ is the probability assigned to the good state such that the expected utility of betting on the good state equals that of betting on the bad state. Hence, as long as $q_0 < q^A$, D's prior suggests the bet on the bad

[23] Notice that we ignore the case $q_0 = q_0^T$. This is to avoid tedious case distinctions that add nothing to our analysis.

state. Notice also that $q_0^B < q^A$. As a result, if $q_0 < q_0^B$, a Bayesian finds it optimal to bet on the bad state, which is also what D's prior points to. Clearly, D prefers betting on the bad state in this case. Since the beliefs inducing this bet include both her prior and the Bayesian posterior, Assumption 2 implies that D adopts the posterior. In contrast, if $q_0^B \le q_0 < q^A$, a Bayesian prefers the bet on the good state, while D's prior continues to suggest the bet on the bad state. For these priors, we have the trade-off between objective performance and regret avoidance that is at the heart of our model. The priors $q_0^B \le q_0 < q_0^T$, where $q_0^T \le q^A$ depending on λ, are the priors where regret avoidance dominates causing D to prefer the bet on the bad state.[24] Since the beliefs inducing this bet include her prior, but no longer the posterior, she adopts the prior. Finally, if $q_0 > q_0^T$, the force of regret has become too weak or is entirely absent (if $q_0 > q^A$). Accordingly, D finds the bet on the good state optimal, which leads her to adopt the posterior.

The left panel of Figure 3.1 illustrates Lemma 1 for a particular set of parameter values. The optimal belief sets for the different priors q_0 are given by the grey areas, while D's adopted belief is designated by the blue dotted line. The red curve represents the Bayesian posterior given $s_1 = 1$ as a function of q_0 . The black horizontal and vertical lines are each drawn at q^A . As a result, q_0^B is pinned down by the point of intersection between the Bayesian posterior and the horizontal line, while q_0^T is the prior where optimal beliefs change.

To address the issue of asymmetric information processing, we now compare the situation just studied to the mirror image situation where D has received the bad news $s_1 = 0$. In a first step, define

$$q_0^{T'} = max\left\{\beta\sigma/\left(\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\lambda\right)+\beta\sigma\right), q^A\right\}.$$

as the threshold prior for $s_1 = 0$. Moreover, define $q_0^{B'} = \beta\sigma/\left(\left(1-\alpha\right)\left(1-\sigma\right)+\beta\sigma\right)$ with $q_0^{B'}$ being the prior at which a Bayesian agent is indifferent between betting on the good and bad state after $s_1 = 0$. We have

LEMMA 2 Suppose that $\lambda > 0$ and $s_1 = 0$. If $q_0 < q_0^{T'}$, the decision maker prefers beliefs making her bet on the bad state, i.e., $Q_1^* = \left(0, q^A\right)$. If $q_0 > q_0^{T'}$, she prefers beliefs making her bet

[24] If λ is large enough that $q_0^T = q^A$, D prefers the bet on the bad state for all $q_0 < q^A$, i.e., for all priors suggesting this bet ("strong regret").

on the good state, i.e., $Q_1^* = \left(q^A, 1\right)$. Her adopted belief satisfies $\hat{q}_1 = q_0$ if $q_0^{T'} < q_0 \le q_0^{B'}$ and $\hat{q}_1 = \tilde{q}$ if $q_0 < q_0^{T'}$ or $q_0 > q_0^{B'}$.

Similar to before, D adopts the Bayesian posterior except if $q_0^{T'} < q_0 \le q_0^{B'}$. In this case, she adopts the prior. We again have conservatism because average belief revision is negative, but less pronounced than under Bayesian updating. The intuition for Lemma 2 is analogous to that for Lemma 1: For priors between $q_0^{T'}$ and $q_0^{B'}$, a Bayesian prefers to bet on the bad state, while the prior suggests the bet on the good state since $q_0^{T'} \ge q^A$. Because of her sensitivity to regret, D prefers to bet on the good state leading her to adopt the prior. The right panel in Figure 3.1 illustrates. The point $q_0^{B'}$ is defined by the intersection between the Bayesian posterior (the red curve) and the horizontal line at q^A, while $q_0^{T'}$ is where optimal beliefs change.

We now turn to asymmetry. As mentioned above, Möbius et al (2012) report that experimental subjects with a prior q_0 who receive the signal $s_1 = 1$ revise their belief significantly more in absolute terms than subjects with the inverse prior $1 - q_0$ who receive $s_1 = 0$. Under Bayesian updating, the absolute belief changes should be the same owing to the symmetry of priors and signals. Letting $\Delta = q_0^T - q_0^B$ and $\Delta' = q_0^{B'} - q_0^{T'}$, we can account for such asymmetry in belief revision if and only if $\Delta' > \Delta$. To see this, note there are two cases under our model where symmetry is *not* violated: If q_0 satisfies $q_0^B \le q_0 < q_0^T$ and $q_0^{T'} < 1 - q_0 \le q_0^{B'}$, symmetry is maintained because D adopts her prior both after good news given q_0 and after bad news given the counterfactual prior $1 - q_0$, which means that she does not revise her belief in either case. Similarly, if q_0 fulfils neither $q_0^B \le q_0 < q_0^T$ nor $q_0^{T'} < 1 - q_0 \le q_0^{B'}$, symmetry is preserved because the Bayesian posterior is adopted either way. It is only in the remaining two cases that symmetry is violated: If we have $q_0^B \le q_0 < q_0^T$, but not $q_0^{T'} < 1 - q_0 \le q_0^{B'}$, D maintains q_0 in response to good news, but adopts the Bayesian posterior in response to bad news assuming that $1 - q_0$ is her prior. Likewise, if we do not have $q_0^B \le q_0 < q_0^T$, but $q_0^{T'} < 1 - q_0 \le q_0^{B'}$, D adopts the posterior in response to good news, but maintains her prior in response to bad news. If $\Delta' = q_0^{B'} - q_0^{T'} > \Delta = q_0^T - q_0^B$, the second type of symmetry violation, which is the empirically observed type, occurs more often, i.e., for more priors.

And indeed, we have

PROPOSITION 4 For all $\lambda > 0$, the two ranges where the decision maker adopts her prior satisfy $\Delta' > \Delta$ if and only if $\alpha + \beta < 1$ and $\Delta' = \Delta$ if and only if $\alpha + \beta = 1$. We have $\partial(\Delta' - \Delta)/\partial\lambda > 0$ if $0 \le \lambda < (2\sigma - 1)/(1-\sigma)$ and $\partial(\Delta' - \Delta)/\partial\lambda = 0$ if $\lambda \ge (2\sigma - 1)/(1-\sigma)$.

Thus, our model can shed light on the findings of Möbius et al (2012) if we impose $\alpha + \beta < 1$ over and above Assumption 1, which is equivalent to $q^A < 0.5$. The intuition is the following: For D to strongly restrict her belief revision in response to bad news, which means that she wants to bet on the good state after bad news much longer than a Bayesian, but only mildly in response to good news, which means that she wants to bet on the bad state after good news not much longer than a Bayesian, betting on the good state must be relatively attractive. The parameter restriction takes care of this. Also, the proposition shows that the discrepancy between the "bad news range" Δ' and the "good news range" Δ and hence the problem of asymmetry increases in regret up to $\lambda = (2\sigma - 1)/(1-\sigma)$, where both ranges have attained their maximal width.[25]

We conclude this sub-section by addressing how existing models of belief formation analyse conservative and asymmetric information processing. We find that these alternative approaches face some difficulties in accounting for the evidence considered above.

Rabin and Schrag (1999) propose a model of *confirmatory bias*, which has the following implications: If the bad state is more likely under the prior ($q_0 < 0.5$), but the signal points to the good state ($s_1 = 1$), there is a probability $\mu > 0$ with which D erroneously perceives the signal as confirming her prior, i.e., as $s_1 = 0$, and likewise for the scenario where $q_0 > 0.5$ and $s_1 = 0$. If prior and signal point into the same direction, no misperception of the signal occurs. Given her perceived signal, D then updates her belief according to Bayes' Rule. The model can explain conservatism because there are values of μ such that average belief revision goes into the same direction as under Bayesian updating, but less so. However, it has a hard time accounting for asymmetry because the prior ranges affected by confirmatory bias in the good and bad news scenario are symmetric ($q_0 < 0.5$ and $q_0 > 0.5$).

Brunnermeier and Parker (2005) propose a model of belief choice where objective performance is traded off against *anticipatory utility* (rather than regret avoidance). In our single-period framework, they effectively posit two sub-periods: In the first sub-period, D

[25] Indeed, if $\lambda > (2\sigma - 1)/(1-\sigma)$, strong regret holds for our one-period problem (see Section 3.2.4).

receives her signal and chooses her belief. In the second sub-period, the chosen belief implements an action and utility is realised. When choosing her belief, D pursues two objectives: On the one hand, objective performance, i.e., the expected utility of the implemented action(s), which is calculated using the correct Bayesian posterior. On the other hand, D wants to feel good about her belief choice *ex ante*. This second objective is captured by the expected utility of the implemented action(s) *given the belief itself*. As we show in Appendix C, anticipatory utility can in principle account for (something reminiscent of) asymmetric information processing. However, it fails to capture conservatism: If $\alpha < \beta$, the model predicts extreme beliefs ($q_1 = 0$ or $q_1 = 1$) because D's anticipatory utility is maximised by assigning maximal probability to whichever state she is betting on. If $\alpha > \beta$, which is, as argued above, more plausible in the context of processing ego-sensitive information, D's belief revision exhibits a strong bias towards the good state. As we show in Appendix C, D never chooses a belief below her posterior in this case. This is incompatible with conservatism in the good news scenario and largely so in the bad news scenario, where D stays between her prior and posterior for only a small fraction of her prior range.[26]

3.3.2 Belief Dynamics

We next study how D's beliefs evolve over time in a two-period framework, i.e., assuming $\tau = 2$.[27] The goal is to reach a better understanding of the mechanics of our model and identify interesting inter-temporal effects. We remain in the framework of Assumption 1. Continuing to interpret $x = 1$ as the *good state*, we characterise D's optimal beliefs across time as *over-* or *underconfident* by saying that D is overconfident (underconfident) in a given situation if her optimal beliefs about the likelihood of the good state exceed (fall short of) the Bayesian posterior associated with the situation. This idea is formalised in

---

[26] The model offers somewhat more flexibility in the knife edge case $\alpha = \beta$ because every belief implementing the bet on the bad state, i.e., every belief between 0 and q^A, is then equally good, while $q_1 = 1$ is still best among the beliefs implementing the bet on the good state.

[27] The same qualitative insights hold for three periods. For tractability, we restrict the analysis to $\tau = 2$.

DEFINITION 2 Given the signal history $(s_1,\ldots,s_t)$ and the decision maker's initial prior q_0, the decision maker is *overconfident* after $(s_1,\ldots,s_t)$ if and only if $\inf Q_t^*(q_0,s_1,\ldots,s_t) > \tilde{q}(q_0,s_1,\ldots,s_t)$ where $Q_t^*(q_0,s_1,\ldots,s_t)$ is defined recursively as $\bigcup_{q_{t-1}\in Q_{t-1}^*(q_0,s_1,\ldots,s_{t-1})} Q_t^*(q_{t-1},s_t)$ and *underconfident* after $(s_1,\ldots,s_t)$ if and only if $\sup Q_t^*(q_0,s_1,\ldots,s_t) < \tilde{q}(q_0,s_1,\ldots,s_t)$.

The set $Q_t^*(q_0,s_1,\ldots,s_t)$ contains the beliefs that D finds optimal after $(s_1,\ldots,s_t)$ given that her initial prior is q_0. If all beliefs in it exceed the belief of a Bayesian who has started out with the same prior and faces the same signal history, D is overconfident. Likewise, if all optimal beliefs are below the Bayesian posterior, D is underconfident.[28]

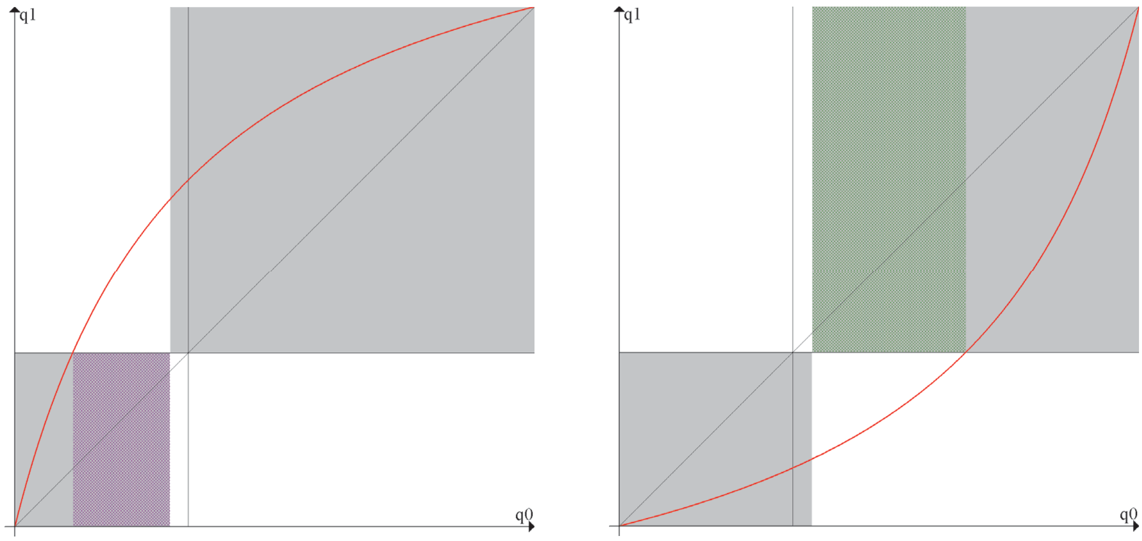

FIGURE 3.2: Optimal beliefs after $s_1 = 1$ and $s_1 = 0$ as a function of the prior q_0 given $\tau = 1$, $\alpha = 0.2$, $\beta = 0.4$, $\sigma = 0.8$ and $\lambda = 2.4$. Underconfident beliefs are highlighted violet, overconfident ones green.

For an illustration, consider Figure 3.2, which shows how over- and underconfidence play out in the one-period setting studied in the previous sub-section. Optimal beliefs are again represented by the grey areas. The left panel contains the situation after $s_1 = 1$. In this case, there is only underconfidence, the beliefs in question being highlighted violet. The intuition is that

[28] This notion is also drawn on in Möbius et al (2012). It differs from the conventional definition according to which we have e.g. overconfidence in a population if more than half the population believes to perform above average in terms of some desirable characteristic. Benoît and Dubra (2011) argue that this state of affairs can be consistent with Bayesian updating.

D's reference point, i.e., her information before receiving $s_1 = 1$, is relatively pessimistic about the good state. As a result, D's optimal beliefs are partly depressed relative to the Bayesian benchmark. The right panel shows what happens after $s_1 = 0$. Here, the inverse result of partial overconfidence obtains, the corresponding beliefs being shaded green. The priors associated with overconfidence after $s_1 = 0$ correspond to the "bad news range" from above, while the priors associated with underconfidence after $s_1 = 1$ coincide with the "good news range".

Furthermore, both over- and underconfidence are "action-relevant": In each case, the optimal beliefs are on one side of the horizontal line drawn at q^A and the Bayesian posterior on the other. As a result, q^A being the belief for which D is indifferent between the two actions, the optimal beliefs implement a different action than the posterior. Finally, the interval of priors entailing overconfidence after $s_1 = 0$ is wider than the interval entailing underconfidence after $s_1 = 1$ for the reasons of payoff asymmetry discussed above. In this sense, overconfidence is more prevalent whenever $\alpha + \beta < 1$.

We now turn to the two-period case, which consists of a two-fold repetition of the decision problem from Assumption 1. For the sake of simplicity, we restrict attention to beliefs implementing *pure strategies*. This means that D's choice set in every period is limited to the set of *admissible beliefs* implementing one action or one action plan with certainty. In the second period, the admissible beliefs comprise all $q_2 \in Q$ for which $A^*(q_2)$ is singleton. This holds for all beliefs except $q_2 = q^A$. In the first period, there are two requirements: Firstly, $A^*(q_1)$ must be singleton. Secondly, letting $Q_2^*(q_1, s_2)$ be the set of optimal admissible beliefs in the second period given q_1 and s_2, we must have $A^*(q_2) = A^*(q_2')$ for all $q_2, q_2' \in Q_2^*(q_1, s_2)$ and $s_2 \in S$, i.e., for each signal all optimal beliefs given q_1 must implement a common action. This rules out $q_1 = q^A$ and up to two more first-period beliefs (see the proof of Proposition 5). Definition 2 (over- or underconfidence) continues to apply once it is taken for granted that sets of optimal beliefs refer to sets of optimal admissible beliefs.

Letting $\lambda^m = \left(\sqrt{1 - 2\sigma + 9\sigma^2 - 12\sigma^3 + 4\sigma^4} - 1 - \sigma + 2\sigma^2\right) \Big/ 2\sigma(1-\sigma)$ and $\lambda^s = (2\sigma - 1)/(1-\sigma)$, we distinguish four scenarios for the regret parameter:

DEFINITION 3 The decision maker displays *mild regret* if $0 < \lambda \le \lambda^m$, *intermediate regret* if $\lambda^m < \lambda \le \lambda^s$, *strong regret in the last period* if $\lambda^s < \lambda \le 2\lambda^s$ and *overall strong regret* if $\lambda > 2\lambda^s$.

For an interpretation, consider mild regret first. Since we have $\lambda < \lambda^s$ in this case, where λ^s is the strong regret threshold for the second period (see Section 3.2.4), D does not want to follow her second-period prior q_1 irrespective of s_2. Four different *pure strategies* are therefore implementable: If $q_1 < q^A$, we have $a_1 = 0$ and $a_2 = 0$ after $s_2 = 0$ since there is no tension between s_2 and q_1 in this case. After $s_2 = 1$, $a_2 = 0$ is implemented for low and $a_2 = 1$ for high $q_1 < q^A$. Intuitively, since D does not always want to follow q_1, she prefers $a_2 = 1$ for sufficiently high $q_1 < q^A$. Likewise, if $q_1 > q^A$, we have $a_1 = 1$ and $a_2 = 1$ after $s_2 = 1$. After $s_2 = 0$, $a_2 = 1$ is implemented for high and $a_2 = 0$ for low $q_1 > q^A$.

Consider next intermediate and strong regret. As long as $\lambda < \lambda^s$, the same four pure strategies as under mild regret are implementable. What changes is that D no longer finds all of them optimal given some q_0 and s_1. For example, given $s_1 = 1$, no q_0 exists such that D wants to implement $a_1 = 1$ in the first period, but $a_2 = 0$ after $s_2 = 0$. Rather, if she prefers $a_1 = 1$, she prefers $a_2 = 1$ after both second-period signals because $a_2 = 0$ would lead to regret in the second period, which she wants to avoid if $\lambda > \lambda^m$. In contrast, if $\lambda \geq \lambda^s$, fewer pure strategies can be implemented because D wants to follow q_1 in the second period irrespective of s_2. If $q_1 < q^A$, $a_2 = 0$ is implemented after both signals, while $a_2 = 1$ is implemented if $q_1 > q^A$.[29] Finally, if $\lambda > 2\lambda^s$, strong regret holds in the first period and hence overall meaning that D follows q_0 when choosing q_1.

Now, let $H = \{s_1 = 1, s_1 = 0, (s_1 = 1, s_2 = 1), (s_1 = 1, s_2 = 0), (s_1 = 0, s_2 = 1), (s_1 = 0, s_2 = 0)\}$ be the set of non-empty signal histories in our two-period decision problem. We have

PROPOSITION 5 Suppose that $\tau = 2$ and $0 < \lambda < 2\lambda^s$. For every signal history $h \in H$, if the last signal in h points to the good (bad) state, there exists a non-empty interval of priors $Q_0^-(h)$ ($Q_0^+(h)$) for which the decision maker is underconfident (overconfident), while there are no priors for which she is overconfident (underconfident). Any over- or underconfidence is *action-relevant* meaning that the following holds for every $h \in H$: If the prior q_0 satisfies $q_0 \in Q_0^-(h)$ or $q_0 \in Q_0^+(h)$, $A^*(\tilde{q}(q_0, h)) \neq A^*(q_t)$ for every $q_t \in Q_t^*(q_0, h)$.

[29] As mentioned, we exclude $q_1 = q^A$ for entailing randomisation in the first (and possibly second) period.

COROLLARY 2 There is path dependence in belief choice meaning that $Q_2^*(q_0, s_1 = s, s_2 = s') \neq Q_2^*(q_0, s_1 = s', s_2 = s)$ for some $q_0 \in Q$ and $s, s' \in S$ where $s \neq s'$.

Action relevance puts additional structure on over- or underconfidence. Over- or underconfidence only says that the set of optimal beliefs $Q_t^*(q_0, h)$ does not contain the Bayesian posterior $\tilde{q}(q_0, h)$. Action relevance goes further by requiring that all optimal beliefs implement different actions than the posterior. The proposition makes clear that our model only generates this type of over- or underconfidence. The intuition is that the only reason D has for departing from the Bayesian posterior is to follow the lead of her prior in those cases where there is tension between the prior and the posterior at the level of implemented actions.

Furthermore, the proposition asserts that there is exactly one instance of over- or underconfidence after each signal history. If the last signal has been good (bad), some priors entail underconfidence (overconfidence) and none overconfidence (underconfidence). Thus, the basic pattern from the one-period problem is preserved: If the last signal has been good (bad), D's reference point is relatively pessimistic (optimistic), which causes her beliefs to be partly depressed (inflated) relative to the Bayesian benchmark. The proposition is limited to $\lambda < 2\lambda^s$ because some instances of over- or underconfidence disappear for higher levels of regret. We return to this issue below.

The corollary, which spells out an important difference between our model and Bayesian updating, follows from considering the histories $(s_1 = 1, s_2 = 0)$ and $(s_1 = 0, s_2 = 1)$. By Proposition 5, we have overconfidence after the first and underconfidence after the second history because of the different last signals. This together with the fact that the Bayesian posterior is the same in the two situations implies that optimal beliefs differ for some priors.

Three issues in connection with Proposition 5 merit further attention. Firstly, it is surprising that there is only one instance of over- or underconfidence in the first period because D's belief choice in this period is affected not only by current, but also by future (i.e., second-period) regret, which does not however lead to any additional over- or underconfidence in the first period. Secondly, we address how increasing λ influences the range of priors affected by over- or underconfidence after a given signal history. Intuitively, one would expect these ranges to grow in λ. Yet, there are exceptions to this rule, which we point out below. Finally, we briefly discuss potential welfare benefits of choosing myopically under regret.

In what follows, we draw heavily on the pure strategies implemented by D's first-period belief, which we label as follows: Consider mild regret. As mentioned before, four pure strategies are implementable in this case: If $q_1 < q^A$, we have $a_1 = 0$ and $a_2 = 0$ after $s_2 = 0$. After $s_2 = 1$, we have $a_2 = 0$ for low and $a_2 = 1$ for high $q_1 < q^A$. We refer to these belief intervals and strategies as I and II, respectively. If $q_1 > q^A$, we have $a_1 = 1$ and $a_2 = 1$ after $s_2 = 1$. After $s_2 = 0$, we have $a_2 = 0$ for low and $a_2 = 1$ for high $q_1 > q^A$, to which we refer as III and IV.[30]

The Impact of Regret on the First Period

To tackle the first issue, namely, why there is only one instance of over- or underconfidence in the first period, it is useful to focus on the case of mild regret and $s_1 = 1$. The intuition for higher levels of regret and/or $s_1 = 0$ is analogous. The top left panel of Figure 3.3 illustrates D's belief choice after $s_1 = 1$ as a function of q_0. For her lowest priors, D prefers I. As her prior increases, she eventually prefers II. With q_0 rising further, her preference switches to III. For her highest priors, she prefers IV. Consistent with Proposition 5, we have a single instance of underconfidence (shaded violet), which arises from D preferring II longer, i.e., for higher priors, than warranted by Bayesian updating. The mechanics of this should be familiar by now: II and III only differ in the action implemented in the first period with II entailing $a_1 = 0$ and III entailing $a_1 = 1$. As q_0 rises, the Bayesian posterior given $s_1 = 1$ (the red curve) eventually crosses the horizontal line at q^A. It is from this point onward that a Bayesian prefers $a_1 = 1$. However, as long as $q_0 < q^A$, which is given by the vertical line, the prior suggests $a_1 = 0$. Being sensitive to regret, D follows her prior for some time by preferring II when the posterior already suggests III.

Similar distortions from regret arise elsewhere, but do not lead to additional over- or underconfidence in the first period. Consider D's choice of I versus II at her lowest priors. The prior at which she is indifferent between I and II is larger than the prior at which a Bayesian agent is. To see this, notice that I and II only differ in what they entail after $(s_1 = 1, s_2 = 1)$ with I implementing $a_2 = 0$ and II implementing $a_2 = 1$. The Bayesian posterior for this signal history is given by the red curve in the middle left panel of Figure 3.3.

[30] The proof of Proposition 5 formally defines the four intervals.

Clearly, the prior at which this curve crosses the horizontal line is smaller than the prior at which D is indifferent between I and II. The reason is D's regret in the second period: Both I and II implement $a_1 = 0$, which is therefore the unique reference action for the second period. Since I implements $a_2 = 0$, D prefers it at priors where a Bayesian already prefers II. Yet, D preferring I longer than a Bayesian does *not* imply additional underconfidence in the first period because I is implementable by beliefs q_1 that are higher than the highest Bayesian posterior after $s_1 = 1$ such that $a_2 = 0$ is preferred after $(s_1 = 1, s_2 = 1)$. The reason is again D's regret in the second period, which gives her additional "slack" in implementing I. These two effects, namely, D preferring I at higher priors than a Bayesian and D being able to implement I through higher beliefs, exactly offset each other, so that D's optimal first-period beliefs are not affected by underconfidence in this region of the prior space. The same kind of "cancelling" obtains for D's choice between III and IV.[31] The lesson is that future sensations of regret determine present belief choices, but do not give rise to over- or underconfidence in the present (while they do lead to over- or underconfidence in the future, as is apparent from Figure 3.3).

The Impact of Increasing Regret

Secondly, we address the impact of increasing regret on over- or underconfidence. Intuitively, one would expect the affected prior ranges to increase in λ since over- and underconfidence are engendered by (sensitivity to) regret and should therefore become more acute as the latter grows. Yet, there are exceptions to this, as we now show. Again, we consider histories beginning with $s_1 = 1$. For brevity, we focus on the exceptions, which occur after $(s_1 = 1, s_2 = 0)$ and $s_1 = 1$. The message is that the comparative statics of over- or underconfidence can be surprising once beliefs are chosen dynamically under regret. All formal details are in the proof of Proposition 5. Figures 3.3-3.5 illustrate.

[31] There is no overconfidence in the first period although a Bayesian prefers IV to III if his prior exceeds q^A, while D prefers it sooner because her reference action is $a_1 = 1$ and IV conforms to it. However, IV is also implementable by beliefs that are below the lowest posterior after $s_1 = 1$ such that $a_2 = 1$ is preferred after $(s_1 = 1, s_2 = 0)$. These two effects cancel each other.

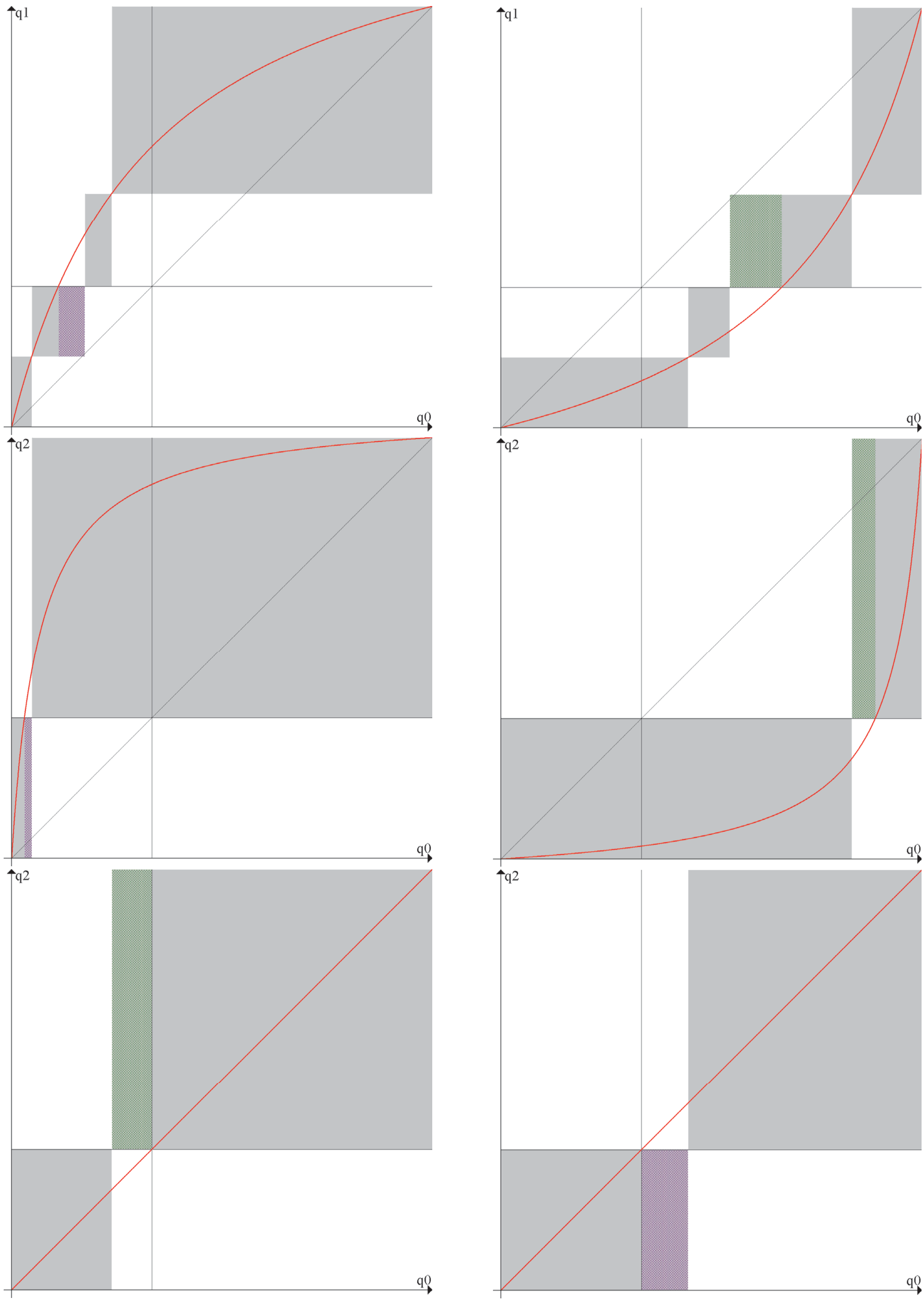

FIGURE 3.3: Optimal beliefs under mild regret given $\alpha = 0.2$, $\beta = 0.4$, $\sigma = 0.8$ and $\lambda = 0.6$ after the different possible signal histories. Histories beginning with $s_1 = 1$ are shown on the left-hand side and histories beginning with $s_1 = 0$ on the right-hand side.

Consider first D's overconfidence after $(s_1 = 1, s_2 = 0)$, where there is an unexpected reversal in the *marginal impact* of regret. To see this, let $q_0^{x\sim y}$ denote the prior at which D is indifferent between the belief intervals x and $y \neq x$ with $x, y \in \{I, II, III, IV\}$. If regret is mild, the lower bound of $Q_0^+(s_1 = 1, s_2 = 0)$ is $q_0^{III\sim IV}$, which falls in λ amounting to increasing overconfidence. If $\lambda^m < \lambda < \lambda^s$, III is no longer optimal, so that the lower bound becomes $q_0^{II\sim IV}$, which instead rises in λ.[32] The intuition for this reversal is the following: III and IV only differ after $(s_1 = 1, s_2 = 0)$ with III implementing $a_2 = 0$ and IV implementing $a_2 = 1$. Both implement $a_1 = 1$, which acts as the reference action for the second period. Thus, the higher λ, the earlier IV is preferred. In contrast, the choice between II and IV is *not* influenced by second-period regret because II implements $a_1 = 0$. Now, the higher λ, the longer II is preferred for conforming to D's first-period prior $q_0 < q^A$. This explains the differential impact of λ in the two cases. If $\lambda \geq \lambda^s$, the lower bound is $q_0^{I\sim IV}$, which again increases in λ to such an extent that D's overconfidence after $(s_1 = 1, s_2 = 0)$ has vanished once $\lambda \geq 2\lambda^s$, which is also the reason why Proposition 5 requires $\lambda < 2\lambda^s$. The intuition for the disappearance of overconfidence after $(s_1 = 1, s_2 = 0)$ is that there is no fundamental tension in this case between objective performance and the avoidance of prior regret since the two signals cancel each other. Thus, the Bayesian wants to follow his original prior q_0 after this history, which is also what D aims for (given overall strong regret).

Secondly, after $s_1 = 1$, there is an unexpected *drop* in underconfidence as regret becomes strong. If $\lambda^m < \lambda < \lambda^s$, the upper bound of $Q_0^-(s_1 = 1)$ is $q_0^{II\sim IV}$, while it is $q_0^{I\sim IV}$ if $\lambda \geq \lambda^s$ because only I and IV remain implementable. We have $q_0^{I\sim IV} < q_0^{II\sim IV}$ if $\lambda = \lambda^s$, which implies a discrete drop in underconfidence (compare Figures 3.4 and 3.5). The intuition is the following: The attraction of II is that it allows D to follow her prior $q_0 < q^A$ in the first period, while implementing $a_2 = 1$ after $(s_1 = 1, s_2 = 1)$, i.e., after two signals in favour of the good state, which represents an attractive mix between objective performance and regret avoidance. If regret is strong, I remains the only alternative to IV, which is less attractive in this regard because it commits D to $a = 0$ after all histories. As a result, D prefers IV around the prior where she is indifferent between II and IV.[33]

[32] The upper bound of $Q_0^+(s_1 = 1, s_2 = 0)$ is fixed. It is given by q^A.

[33] The marginal impact of regret is as expected because $q_0^{II\sim IV}$ and $q_0^{I\sim IV}$ increase in λ. Also, no drop or jump occurs as regret becomes intermediate.

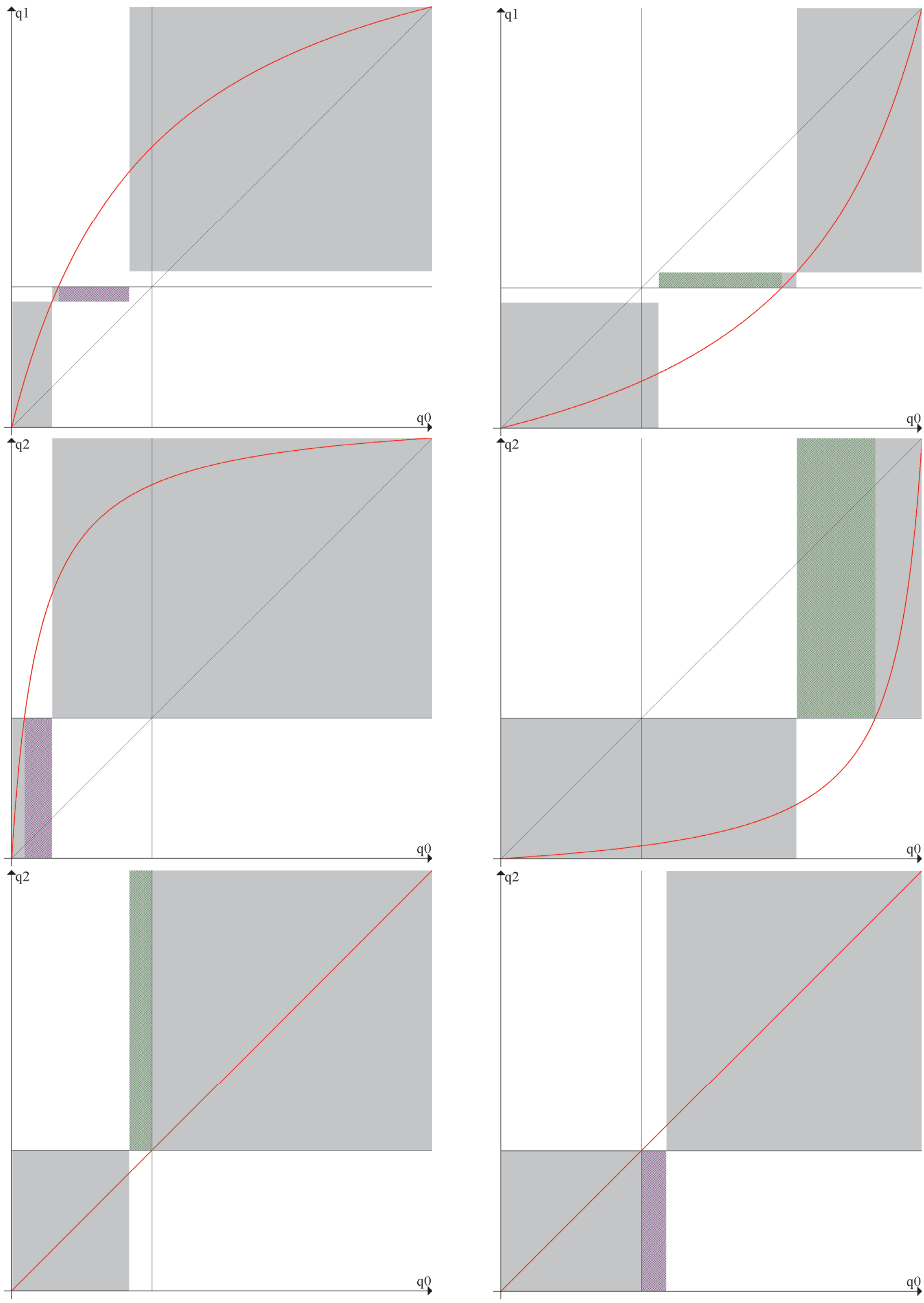

FIGURE 3.4: Optimal beliefs under intermediate regret given $\alpha = 0.2$, $\beta = 0.4$, $\sigma = 0.8$ and $\lambda = 2.4$ after the different possible signal histories. Histories beginning with $s_1 = 1$ are shown on the left-hand side and histories beginning with $s_1 = 0$ on the right-hand side.

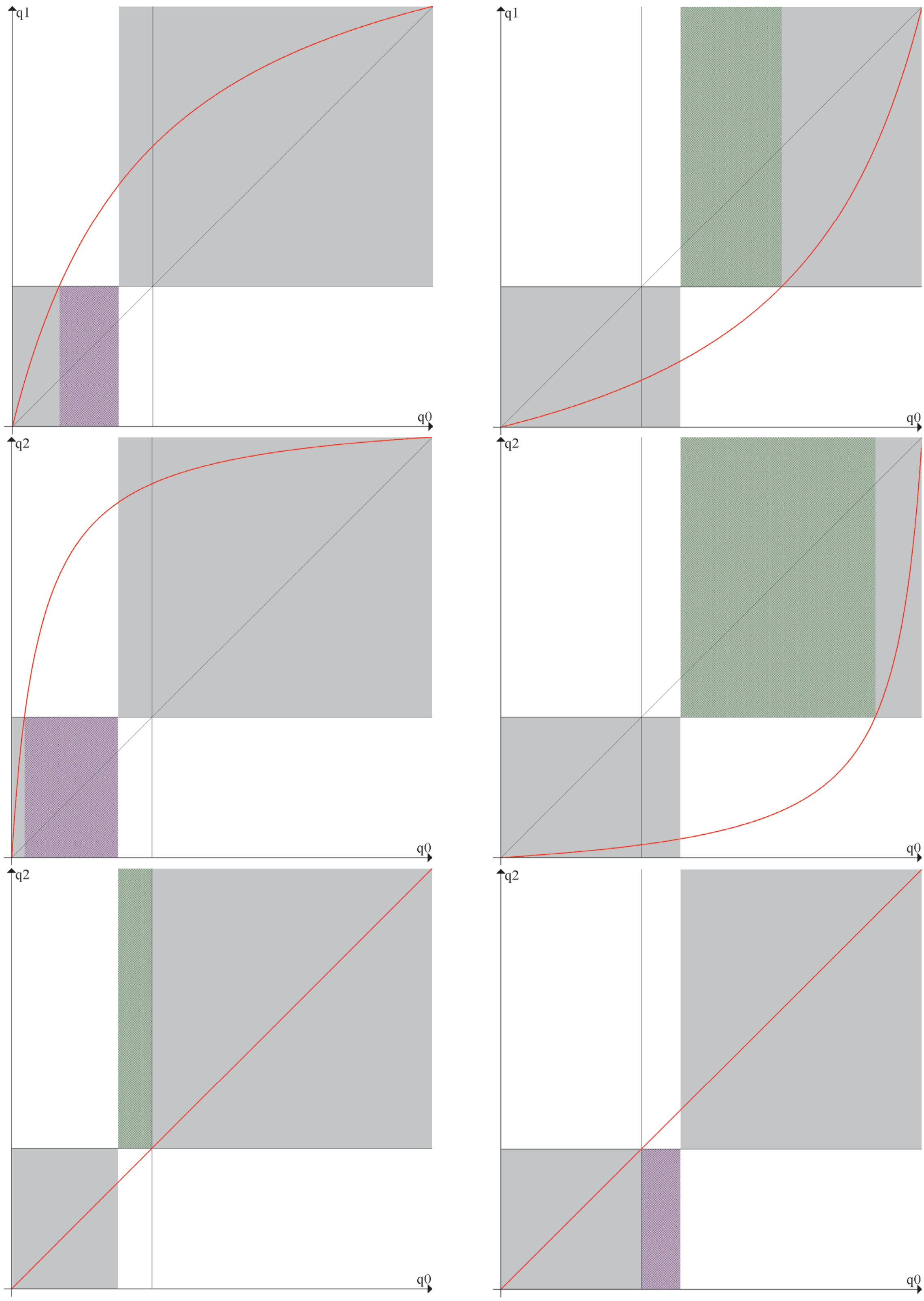

FIGURE 3.5: Optimal beliefs under strong regret in the last period given $\alpha = 0.2$, $\beta = 0.4$, $\sigma = 0.8$ and $\lambda = 3.2$ after the different possible signal histories. Histories beginning with $s_1 = 1$ are shown on the left-hand side and histories beginning with $s_1 = 0$ on the right-hand side.

The Benefits of Myopia under Regret

Finally, we address potential benefits of myopia under regret. As we now show, choosing beliefs myopically may improve welfare. This result is surprising because myopia is generally deemed to be welfare-reducing. However, in the present context, it may serve to mitigate the detrimental impact of regret. We say that D is *myopic* if she maximises $P_t\left(q_t\left|q_{t-1},s_t\right.\right)$ for every $t\in T$ and *forward-looking* if she maximises $P_t\left(q_t\left|q_{t-1},s_t\right.\right)+F_t\left(q_t\left|q_{t-1},s_t\right.\right)$ for all $t<\tau$ and $P_t\left(q_t\left|q_{t-1},s_t\right.\right)$ if $t=\tau$, as was assumed throughout (see Section 3.2.2). Moreover, we assume that D's *welfare* coincides with the preferences of a Bayesian agent with the same utility function as in Assumption 1. This is the only welfare standard worthy of further investigation. If welfare instead included regret, it is immediate that D cannot do worse being forward-looking.

In identifying the benefits of myopia under regret, which exist if $0<\lambda<\lambda^s$, we make life harder for myopia by assuming that D chooses every optimal belief with positive probability. This makes the drawback of myopia, namely, the lack of allowance for what current beliefs trigger in the future, particularly acute.

PROPOSITION 6 Suppose that $\tau=2$. If $0<\lambda<\lambda^s$, there exist priors $q_0\in Q$ and signals $s_1\in S$ such that the decision maker's welfare is higher if she is myopic than if she is forward-looking even if she adopts all optimal beliefs with positive probability. If $\lambda=0$, the decision maker's welfare is at least as high if she is forward-looking as if she is myopic for all $q_0\in Q$ and $s_1\in S$. The same holds if $\lambda\geq\lambda^s$ for all $q_0\neq q^A$.

An intuition for Proposition 6 can be obtained by letting $0<\lambda<\lambda^s$ and considering D's first-period choice given $s_1=1$ between the belief intervals I and II. As noted above, there are priors where a forward-looking D chooses I, but a Bayesian with the same prior prefers II. Suppose that D has such a prior. Since her welfare coincides with the Bayesian's preferences, D's welfare would rise if she increased the probability of choosing II at the expense of I. Being myopic achieves just that. To see this, consider the left panel of Figure 3.2, which can be interpreted as D's belief choice in the first period if she is myopic because she then treats her two-period problem as a one-period one. Clearly, at the low priors under consideration, a myopic D prefers the beliefs in $\left(0, q^A\right)$, which include I and II.[34] Since we assume every optimal belief is adopted with positive probability, it follows that being myopic lets D choose both I and II with positive probability, which increases her welfare relative to being forward-looking. At a more abstract level, this is an example of one bias (regret) being mitigated by a second (myopia). Adding a bias to an existing one can therefore be welfare-enhancing.[35]

In contrast, under no regret, D's preferences if she is forward-looking coincide with those of a Bayesian agent. Clearly, being myopic cannot do better in this case. As for the remaining levels of λ, consider first regret that is strong in the last period, but not overall. In this case, D replicates q_1 in the second period, which implies that two pure strategies remain implementable, namely, I and VI. If D is myopic, she treats her problem in the first period as a one-period problem. As a result, she follows $q_0 \neq q^A$, which causes her to implement I for all $q_0 < q^A$. In contrast, given $s_1 = 1$, a forward-looking D chooses VI for some $q_0 < q^A$, which is also what a Bayesian prefers. As a result, D is better off being forward-looking for these priors, while there is no difference otherwise. There is also no welfare difference if regret is strong overall because D follows $q_0 \neq q^A$ in this case no matter if she is myopic or forward-looking.[36]

Further Issues and Summary

One might also wonder what happens to over- or underconfidence as signals in support of a given state accumulate. Consider, for instance, overconfidence after $\left(s_1 = 0, s_2 = 0\right)$, which is illustrated for the case of mild regret in the middle right panel of Figure 3.3. Visibly, this overconfidence affects a smaller range of priors than D's overconfidence after $s_1 = 0$, as is apparent from the top right panel in Figure 3.3. From this, one could conclude that the problem of overconfidence (and underconfidence) diminishes in the number of signals in favour of a given state. However, this relationship is inverted for high levels of regret (see Figure 3.5).

---

[34] Strictly speaking, the myopic D does not prefer $\left(0, q^A\right)$, but the union of the intervals I and II because the point between them is not an admissible first-period belief.

[35] O'Donoghue and Rabin (2001) and Herweg and Müller (2011) show in the context of hyperbolic discounting that self-control problems may be mitigated by naiveté.

[36] Since we rule out beliefs implementing mixed strategies, we abstract for simplicity from the prior $q_0 \neq q^A$ suggesting randomisation in the first period.

More formally, let $\Delta^{+}(h) = \sup Q_0^{+}(h) - \inf Q_0^{+}(h)$ denote the width of the prior range $Q_0^{+}(h)$ entailing overconfidence after a given $h \in H$ and define $\Delta^{-}(h)$ analogously. Moreover, restrict attention to $\alpha + \beta = 1$. In this case, one can show that

$$\Delta^{+}(s_1 = 0) > \Delta^{+}(s_1 = 0, s_2 = 0) \text{ and } \Delta^{-}(s_1 = 1) > \Delta^{-}(s_1 = 1, s_1 = 1)$$

for all $0 < \lambda \leq \lambda^{m}$. However, the inverse result obtains if $\lambda \geq \lambda^{s}$. Intuitively, under strong regret, D is committed to replicating her first-period belief choice in the second period. Maintaining one's previous belief after an additional signal in favour of a given state aggravates one's over- or underconfidence.[37]

We conclude this sub-section by some remarks on the relative frequency of over- and underconfidence. Similar to the one-period case studied before, there is a sense in which there is more over- than underconfidence once we are willing to assume $\alpha + \beta < 1$. More precisely, $\alpha + \beta < 1$ can be shown to be equivalent to

$$\Delta^{+}(s_1 = 0) > \Delta^{-}(s_1 = 1) \text{ and to } \Delta^{+}(s_1 = 0, s_2 = 0) > \Delta^{-}(s_1 = 1, s_2 = 1)$$

for all $\lambda > 0$. To the extent that the likelihood of observing over- or underconfidence is monotonic in the width of the affected prior ranges, this fact makes overconfidence more likely. However, what works in the opposite direction is that $\alpha + \beta < 1$ is also equivalent to

$$\Delta^{-}(s_1 = 0, s_2 = 1) > \Delta^{+}(s_1 = 1, s_2 = 0)$$

for all $\lambda > 0$ such that these ranges are positive. While the last pair of histories can be thought of as only half as likely as $s_1 = 0$ and $s_1 = 1$, the "surplus" from above working in favour of overconfidence cannot be shown always to dominate this "deficit". Simulations suggest, however, that $\Delta^{+}(s_1 = 0) - \Delta^{-}(s_1 = 1) > 0.5\Delta^{-}(s_1 = 0, s_2 = 1) - 0.5\Delta^{+}(s_1 = 1, s_2 = 0) > 0$ is only violated for extreme parameter values that are unlikely to arise in practical applications.

---

[37] If $\lambda^{m} < \lambda < \lambda^{s}$, there is a $\overline{\lambda} > \lambda^{m}$ below (above) which the same occurs as for mild (strong) regret.

All in all, this sub-section has supplied an extensive characterisation of dynamic belief choice under regret in a two-period framework and generated novel, testable predictions: Over- and underconfidence exhibit a clear inter-temporal pattern by always depending on the last signal only. Moreover, over- or underconfidence is always action-relevant, which is a novel concept not previously considered in this context. The comparative statics of over- and underconfidence are largely intuitive, but exhibit some interesting deviations from what one would expect.

3.4 Conclusion

This chapter presents a model of instrumental belief choice under regret. Its core idea is that people attach value to maintaining their prior in the face of new information because they anticipate the regret that they will feel from giving up their prior if holding on to their prior turns out to be *ex post* optimal. Without the influence of such regret because of either insensitivity or certainty, our model reduces to the standard approach: Actions generated through belief choices are then identical to those of a standard Bayesian decision maker.

Under regret, we identify conditions under which people exhibit a preference for consistency meaning that their behaviour across time becomes completely unresponsive to the arrival of new information. We also explain why beliefs are updated conservatively and asymmetrically, for which clean empirical support exists. Finally, we study over- and underconfidence identifying conditions under which the two can be expected to materialise. We also introduce a new distinction between action-relevant and action-irrelevant over- and underconfidence and shed light on when belief formation exhibits path dependence meaning that the order of arrival of a given number of signals matters to adopted beliefs.

At a more abstract level, the foundational idea of this chapter is that the processing of new information is intrinsically context-dependent. It depends on the specific decision problem in which new information is (likely to be) exploited. This basic tenet of our model has rich empirical implications, which we aim to explore in future work.

3.5 Appendix A: Proofs

PROOF OF PROPOSITION 1

Consider first the case $t=\tau$ and suppose that D holds the prior $q_{\tau-1}\in Q$ and has received the signal $s_\tau \in S$. Since $\lambda=0$, the target function can be expressed as

$$V_\tau = \sum_a \pi(a,q_\tau)\sum_x \tilde{q}(x|q_{\tau-1},s_\tau)u(x,a) = \sum_a \pi(a,q_\tau)\tilde{U}(a|q_{\tau-1},s_\tau).$$

It is immediate that $\tilde{q}(q_{\tau-1},s_\tau)\in Q_\tau^*(q_{\tau-1},s_\tau)$ because choosing $q_\tau = \tilde{q}(q_{\tau-1},s_\tau)$ leads D to randomise on the actions in $A^*(\tilde{q}(q_{\tau-1},s_\tau)) = arg\,max_{a\in A}\tilde{U}(a|q_{\tau-1},s_\tau)$. Since $\tilde{q}(q_{\tau-1},s_\tau)$ maximises V_τ by implementing randomisation on $A^*(\tilde{q}(q_{\tau-1},s_\tau))$, any $q_\tau \neq \tilde{q}(q_{\tau-1},s_\tau)$ satisfies $q_\tau \in Q_\tau^*(q_{\tau-1},s_\tau)$ if and only if $A^*(q_\tau)\subseteq A^*(\tilde{q}(q_{\tau-1},s_\tau))$, i.e., if it implements randomisation on a subset of $A^*(\tilde{q}(q_{\tau-1},s_\tau))$. Hence, for every $q_{\tau-1}\in Q$ and $s_\tau\in S$, D implements only actions in $A^*(\tilde{q}(q_{\tau-1},s_\tau))$ with positive probability regardless of which belief from $Q_\tau^*(q_{\tau-1},s_\tau)$ she adopts.

Consider next the case $t=\tau-1$ for a given prior $q_{\tau-2}\in Q$ and signal $s_{\tau-1}\in S$. The target function is

$$V_{\tau-1} = \sum_a \pi(a,q_{\tau-1})\tilde{U}(a|q_{\tau-2},s_{\tau-1}) + \sum_{s_\tau\in S} Pr(s_\tau|q_{\tau-2},s_{\tau-1})\sum_a \pi(a,q_{\tau-1},s_\tau)\tilde{U}(a|q_{\tau-2},s_{\tau-1},s_\tau)$$

where $\pi(a,q_{\tau-1},s_\tau)$ is the probability of choosing action $a\in A$ induced by $q_{\tau-1}$ after $s_\tau\in S$. Clearly, for analogous reasons to above, setting $q_{\tau-1} = \tilde{q}(q_{\tau-2},s_{\tau-1})$ maximises the first term of $V_{\tau-1}$. As for the second term, recall that given any $q_{\tau-1}$ and s_τ D implements only actions with positive probability that are in $A^*(\tilde{q}(q_{\tau-1},s_\tau))$. Hence, by choosing $q_{\tau-1} = \tilde{q}(q_{\tau-2},s_{\tau-1})$, D implements after every $s_\tau\in S$ only actions in $A^*(\tilde{q}(\tilde{q}(q_{\tau-2},s_{\tau-1}),s_\tau))$. By Bayes' Rule, we have $\tilde{q}(\tilde{q}(q_{\tau-2},s_{\tau-1}),s_\tau) = \tilde{q}(q_{\tau-2},s_{\tau-1},s_\tau)$. Moreover, $A^*(\tilde{q}(q_{\tau-2},s_{\tau-1},s_\tau)) = arg\,max_{a\in A}\tilde{U}(a|q_{\tau-2},s_{\tau-1},s_\tau)$. Consequently, $q_{\tau-1} = \tilde{q}(q_{\tau-2},s_{\tau-1})$ maximises the second term and we have $\tilde{q}(q_{\tau-2},s_{\tau-1})\in Q_{\tau-1}^*(q_{t-2},s_{\tau-1})$ for every $q_{\tau-2}\in Q$ and $s_{\tau-1}\in S$.

Moreover, for any $q_{\tau-1}\neq \tilde{q}(q_{\tau-2},s_{\tau-1})$, we have $q_{\tau-1}\in Q_{\tau-1}^*(q_{t-2},s_{\tau-1})$ if and only if, firstly, $A^*(q_{\tau-1})\subseteq A^*(\tilde{q}(q_{\tau-2},s_{\tau-1}))$ and, secondly, the function $\pi(a,q_{\tau-1},s_\tau)$ only assigns positive probability to actions in $A^*(\tilde{q}(q_{\tau-2},s_{\tau-1},s_\tau))$, i.e., $\{a\in A:\pi(a,q_{\tau-1},s_\tau)>0\}\subseteq A^*(\tilde{q}(q_{\tau-2},s_{\tau-1},s_\tau))$ for every $s_\tau\in S$. The need to mimic the posterior for every $s_\tau\in S$ follows from every $s_\tau\in S$ arising with positive probability. Thus, given any $q_{\tau-2} \in Q$, D implements after every $s_{\tau-1} \in S$ only actions in $A^*\left(\tilde{q}\left(q_{\tau-2}, s_{\tau-1}\right)\right)$ and after every $s_{\tau-1} \in S$ and $s_\tau \in S$ only actions in $A^*\left(\tilde{q}\left(q_{\tau-2}, s_{\tau-1}, s_\tau\right)\right)$.

In period $\tau-2$, D's target function for a given $q_{\tau-3} \in Q$ and $s_{\tau-2} \in S$ is

$$V_{\tau-2} = \sum_a \pi\left(a, q_{\tau-2}\right) \tilde{U}\left(a \middle| q_{\tau-3}, s_{\tau-2}\right) + \sum_{s_{\tau-1} \in S} Pr\left(s_{\tau-1} \middle| q_{\tau-3}, s_{\tau-2}\right) \sum_a \pi\left(a, q_{\tau-2}, s_{\tau-1}\right) \tilde{U}\left(a \middle| q_{\tau-3}, s_{\tau-2}, s_{\tau-1}\right) +$$

$$\sum_{s_{\tau-1}, s_\tau \in S \times S} Pr\left(s_{\tau-1}, s_\tau \middle| q_{\tau-3}, s_{\tau-2}\right) \sum_a \pi\left(a, q_{\tau-2}, s_{\tau-1}, s_\tau\right) \tilde{U}\left(a \middle| q_{\tau-3}, s_{\tau-2}, s_{\tau-1}, s_\tau\right).$$

Clearly, $\tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right)$ is a maximiser of the first two terms. As for the third term, recall that for every $q_{\tau-2}$, $s_{\tau-1}$ and s_τ D only implements actions in $A^*\left(\tilde{q}\left(q_{\tau-2}, s_{\tau-1}, s_\tau\right)\right)$. If $q_{\tau-2} = \tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right)$, D therefore only implements actions in $A^*\left(\tilde{q}\left(\tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right), s_{\tau-1}, s_\tau\right)\right) = A^*\left(\tilde{q}\left(q_{\tau-3}, s_{\tau-2}, s_{\tau-1}, s_\tau\right)\right)$. Thus, $q_{\tau-2} = \tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right)$ maximises the third term implying $\tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right) \in Q^*_{\tau-2}\left(q_{t-3}, s_{\tau-2}\right)$. For any $q_{\tau-2} \neq \tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right)$, we have $q_{\tau-2} \in Q^*_{\tau-2}\left(q_{t-3}, s_{\tau-2}\right)$ if and only if three conditions are met: Firstly, $A^*\left(q_{\tau-2}\right) \subseteq A^*\left(\tilde{q}\left(q_{\tau-3}, s_{\tau-2}\right)\right)$. Secondly, $\left\{a \in A : \pi\left(a, q_{\tau-2}, s_{\tau-1}\right) > 0\right\} \subseteq A^*\left(\tilde{q}\left(q_{\tau-3}, s_{\tau-2}, s_{\tau-1}\right)\right)$ for every $s_{\tau-1} \in S$ and finally $\left\{a \in A : \pi\left(a, q_{\tau-2}, s_{\tau-1}, s_\tau\right) > 0\right\} \subseteq A^*\left(\tilde{q}\left(q_{\tau-3}, s_{\tau-2}, s_{\tau-1}, s_\tau\right)\right)$ for every $s_{\tau-1} \in S$ and $s_\tau \in S$. The need to mimic the posterior after every signal history arises from every history having positive probability.

The argument for all periods $t < \tau - 2$ is analogous.■

LEMMA A1 In any period *t*, for any belief $q_t \in Q$ that satisfies $A^*\left(q_t\right) \neq A^*\left(q_{t-1}\right)$, there always exists a state of the world $x \in X$ in which D regrets abandoning q_{t-1}, i.e., for which we have $U\left(x, q_t\right) = \sum_a u\left(x, a\right) \pi\left(a, q_t\right) < U\left(x, q_{t-1}\right) = \sum_a u\left(x, a\right) \pi\left(a, q_{t-1}\right)$.

PROOF Suppose that there is a belief $q_t \in Q$ with $A^*\left(q_t\right) \neq A^*\left(q_{t-1}\right)$ which for every state $x \in X$ satisfies $\sum_a u\left(x, a\right) \pi\left(a, q_t\right) \geq \sum_a u\left(x, a\right) \pi\left(a, q_{t-1}\right)$, which means that uniform randomisation on the actions in $A^*\left(q_t\right)$ yields the same or higher expected utility in every state than uniform randomisation on the actions in $A^*\left(q_{t-1}\right)$. This implies that under any belief $q \in Q$ the average expected utility of the actions in $A^*\left(q_t\right)$ is at least as high as the average expected utility of the actions in $A^*\left(q_{t-1}\right)$.

Suppose first that $A^*\left(q_t\right)$ is not a subset of $A^*\left(q_{t-1}\right)$ and consider the belief q_{t-1}. As just noted, the average expected utility of the actions in $A^*\left(q_t\right)$ under q_{t-1} is at least as high as the average expected utility of the actions in $A^*\left(q_{t-1}\right)$. By definition, the actions in $A^*\left(q_{t-1}\right)$ all yield the same expected utility under q_{t-1}, which represents the maximal level of expected utility attainable under q_{t-1}. Since $A^*(q_t)$ is not a subset of $A^*(q_{t-1})$, there is at least one action in $A^*(q_t)$ yielding less than maximal expected utility under q_{t-1}. For the actions in $A^*(q_t)$ to yield the same or higher average utility than the actions in $A^*(q_{t-1})$, this would have to be compensated by another action in $A^*(q_t)$ yielding more than the maximally attainable expected utility, a contradiction.

If $A^*(q_t)$ is a subset of $A^*(q_{t-1})$ and $\sum_a u(x,a)\pi(a,q_t) > \sum_a u(x,a)\pi(a,q_{t-1})$ for at least one $x \in X$, this implies that the average expected utility under q_{t-1} of the actions in $A^*(q_t)$ exceeds the average expected utility of the actions in $A^*(q_{t-1})$. Since the actions in $A^*(q_{t-1})$ attain the same, maximal expected utility under q_{t-1} and $A^*(q_t)$ is a subset of $A^*(q_{t-1})$, this is a contradiction.

If $A^*(q_t)$ is a subset of $A^*(q_{t-1})$ and $\sum_a u(x,a)\pi(a,q_t) = \sum_a u(x,a)\pi(a,q_{t-1})$ for all $x \in X$, the expected utility under q_t of the actions in $A^*(q_t)$, which represents the maximal level of expected utility attainable under q_t, equals the average expected utility of the actions in $A^*(q_{t-1})$. This can only be the case if either all actions that are in $A^*(q_{t-1})$, but not in $A^*(q_t)$, attain the same expected utility under q_t as the actions in $A^*(q_t)$, a contradiction, or if at least one action that is in $A^*(q_{t-1})$, but not in $A^*(q_t)$, reaches a higher than maximal expected utility under q_t, which is also a contradiction.■

PROOF OF PROPOSITION 2

Consider the case $t = \tau$, where the target function reduces to

$$V_\tau(q_\tau | q_{\tau-1}, s_\tau) = \sum_x \tilde{q}(x | q_{\tau-1}, s_\tau)\left[U(x,q_\tau) + r\left(U(x,q_\tau) - U(x,q_{\tau-1})\right)\right].$$

Lemma A1 makes clear that any belief $q_\tau \in Q$ that satisfies $A^*(q_\tau) \neq A^*(q_{\tau-1})$ entails regret in at least one state of the world. The probability put onto this state by the Bayesian posterior $\tilde{q}(q_{\tau-1}, s_\tau)$ is guaranteed to be positive owing to our assumptions on priors (which have full support on X) and signals (which are not fully revealing).

By the scope for regret assumption, at least one belief $q_\tau \in Q$ exists such that $A^*(q_\tau) \neq A^*(q_{\tau-1})$, namely, the Bayesian posterior, which satisfies $A^*(\tilde{q}(q_{\tau-1}, s_\tau)) \neq A^*(q_{\tau-1})$ and $A^*(q_{\tau-1}) \not\subset A^*(\tilde{q}(q_{\tau-1}, s_\tau))$ for some $q_{\tau-1} \in Q$ and $s_\tau \in S$. Fix such a $q_{\tau-1} \in Q$ and $s_\tau \in S$. Clearly, $\tilde{q}(q_{\tau-1}, s_\tau)$ is a maximiser of the absolute part $\sum_x \tilde{q}(x | q_{\tau-1}, s_\tau) U(x, q_\tau)$, which $q_{\tau-1}$ fails to maximise because it does not induce randomisation on a subset of $A^*(\tilde{q}(q_{\tau-1}, s_\tau))$. As a result, for $q_{\tau-1}$ or any other q_τ such that

$A^*(q_\tau) = A^*(q_{\tau-1})$ to perform better in terms of V_τ than $\tilde{q}(q_{\tau-1}, s_\tau)$, λ must rise sufficiently above zero that avoiding the expected regret $\sum_x \tilde{q}(x|q_{\tau-1}, s_\tau) r(U(x,\tilde{q}) - U(x,q_{\tau-1})) > 0$ inevitably associated with $\tilde{q}(q_{\tau-1}, s_\tau)$ becomes D's dominant concern. The same holds for any other q_τ with $A^*(q_\tau) \neq A^*(q_{\tau-1})$ that performs better in the absolute part than $q_{\tau-1}$.

Let $\bar{\lambda}_{t=\tau}(q_{\tau-1}, s_\tau)$ denote the threshold level of regret such that for a given $q_{\tau-1} \in Q$ and $s_\tau \in S$ we have $q_\tau \in Q^*_\tau(q_{\tau-1}, s_\tau)$ if and only if $A^*(q_\tau) = A^*(q_{\tau-1})$ for all $\lambda > \bar{\lambda}_{t=\tau}(q_{\tau-1}, s_\tau)$. This yields the threshold level $\bar{\lambda}(t=\tau) > 0$ as $\bar{\lambda}(t=\tau) = max_{q_{\tau-1}\in Q, s_\tau \in S}\, \bar{\lambda}_{t=\tau}(q_{\tau-1}, s_\tau)$.

Consider next $t = \tau - 1$ and suppose that $\lambda > \bar{\lambda}(t=\tau)$. Since λ is such that strong regret holds for $t = \tau$, D anticipates that she will follow the lead of $q_{\tau-1}$ in the subsequent period irrespective of s_τ thereby avoiding any regret. The target function for her belief choice is therefore

$$V_{\tau-1}(q_{\tau-1}|q_{\tau-2}, s_{\tau-1}) = \sum_x \tilde{q}(x|q_{\tau-2}, s_{\tau-1})\left[U(x,q_{\tau-1}) + r(U(x,q_{\tau-1}) - U(x,q_{\tau-2}))\right] +$$

$$\sum_{s_\tau \in S} Pr(s_\tau|q_{\tau-2}, s_{\tau-1}) \sum_x \tilde{q}(x|q_{\tau-2}, s_{\tau-1}, s_\tau)\left[U(x,q_{\tau-1})\right].$$

This simplifies to

$$V_{\tau-1}(q_{\tau-1}|q_{\tau-2}, s_{\tau-1}) = \sum_x \tilde{q}(x|q_{\tau-2}, s_{\tau-1})\left[2U(x,q_{\tau-1}) + r(U(x,q_{\tau-1}) - U(x,q_{\tau-2}))\right].$$

Since the absolute part is now $2\sum_x \tilde{q}(x|q_{\tau-2}, s_{\tau-1}) U(x,q_{\tau-1})$ and hence twice as large, we have

$$\bar{\lambda}_{t=\tau-1}(q_{\tau-2}, s_{\tau-1}) = 2\bar{\lambda}_{t=\tau}(q_{\tau-1} = q_{\tau-2}, s_\tau = s_{\tau-1})$$

for any given $q_{\tau-2} \in Q$ and $s_{\tau-1} \in S$ where analogously to before $\bar{\lambda}_{t=\tau-1}(q_{\tau-2}, s_{\tau-1})$ is defined such that for all $\lambda > \bar{\lambda}_{t=\tau-1}(q_{\tau-2}, s_{\tau-1})$ we have $q_{\tau-1} \in Q^*_{\tau-1}(q_{\tau-2}, s_{\tau-1})$ if and only if $A^*(q_{\tau-1}) = A^*(q_{\tau-2})$. Defining $\bar{\lambda}(t = \tau-1) = max_{q_{\tau-2}\in Q, s_{\tau-1}\in S}\, \bar{\lambda}_{t=\tau-1}(q_{\tau-2}, s_{\tau-1})$, it follows that $\bar{\lambda}(t=\tau-1) = 2\bar{\lambda}(t=\tau) > 0$. This also implies that our initial supposition $\lambda > \bar{\lambda}(t=\tau)$ is valid.

The argument for any $\tau > 2$ is analogous. As a result, we have $\bar{\lambda}(t = \tau-2) = 3\bar{\lambda}(t=\tau)$ and in general $\bar{\lambda}(t) = (1+\tau-t)\bar{\lambda}(t=\tau)$.■

PROOF OF PROPOSITION 3

Consider first the case $t=\tau$. The target function for q_τ is

$$V_\tau = 1\cdot\left[\sum_a u(\bar{x},a)\pi(a,q_\tau)+r\left(\sum_a u(\bar{x},a)\pi(a,q_\tau)-\sum_a u(\bar{x},a)\pi(a,q_{\tau-1})\right)\right].$$

It is immediate that $\tilde{q}(q_{\tau-1},s_\tau)\in Q_\tau^*(q_{\tau-1},s_\tau)$ because setting $q_\tau=\tilde{q}(q_{\tau-1},s_\tau)$ leads D to randomise on the actions in $A^*(\tilde{q}(q_{\tau-1},s_\tau))=arg\,max_{a\in A}1\cdot u(\bar{x},a)$ and $r(\cdot)$ is non-decreasing. Since $\tilde{q}(q_{\tau-1},s_\tau)$ maximises V_τ by only putting positive probability on actions maximising $u(\bar{x},a)$, any $q_\tau\neq\tilde{q}(q_{\tau-1},s_\tau)$ satisfies $q_\tau\in Q_\tau^*(q_{\tau-1},s_\tau)$ if and only if $A^*(q_\tau)\subseteq A^*(\tilde{q}(q_{\tau-1},s_\tau))$, i.e., if it entails randomisation on a subset of the actions maximising $u(\bar{x},a)$. As a result, for any $q_{\tau-1}\in Q$ and $s_\tau\in S$ such that $\tilde{q}(\bar{x}|q_{\tau-1},s_\tau)=1$, D chooses only actions with positive probability that are in $A^*(\tilde{q}(q_{\tau-1},s_\tau))$.

Consider next the case $t=\tau-1$ where $\tilde{q}(\bar{x}|q_{\tau-2},s_{\tau-1})=1$. The target function is

$$V_{\tau-1}=1\cdot\left[\sum_a u(\bar{x},a)\pi(a,q_{\tau-1})+r\left(\sum_a u(\bar{x},a)\pi(a,q_{\tau-1})-\sum_a u(\bar{x},a)\pi(a,q_{\tau-2})\right)\right]+$$

$$\sum_{s_\tau\in S}Pr(s_\tau|q_{\tau-2},s_{\tau-1})\cdot 1\cdot\left[\sum_a u(\bar{x},a)\pi(a,q_{\tau-1},s_\tau)+r\left(\sum_a u(\bar{x},a)\pi(a,q_{\tau-1},s_\tau)-\sum_a u(\bar{x},a)\pi(a,q_{\tau-1})\right)\right]$$

where $\pi(a,q_{\tau-1},s_\tau)$ is the probability of choosing action $a\in A$ entailed by $q_{\tau-1}$ after $s_\tau\in S$. For analogous reasons to above, setting $q_{\tau-1}=\tilde{q}(q_{\tau-2},s_{\tau-1})$ maximises the first term of $V_{\tau-1}$. As for the second term, recall from the preceding paragraph that for any $q_{\tau-1}\in Q$ and $s_\tau\in S$ such that $\tilde{q}(q_{\tau-1},s_\tau)$ puts all probability mass on $\bar{x}$, D only chooses actions in $A^*(\tilde{q}(q_{\tau-1},s_\tau))$ with positive probability. Consider $q_{\tau-1}=\tilde{q}(q_{\tau-2},s_{\tau-1})$. Since $\tilde{q}(q_{\tau-2},s_{\tau-1})$ puts all probability mass on $\bar{x}$, Bayes' Rule implies that $\tilde{q}(\tilde{q}(q_{\tau-2},s_{\tau-1}),s_\tau)=\tilde{q}(q_{\tau-2},s_{\tau-1},s_\tau)$ puts all probability mass on $\bar{x}$ for every $s_\tau\in S$. As a result, by setting $q_{\tau-1}=\tilde{q}(q_{\tau-2},s_{\tau-1})$, D puts positive probability after every $s_\tau\in S$ only on actions in $A^*(\tilde{q}(\tilde{q}(q_{\tau-2},s_{\tau-1}),s_\tau))=A^*(\tilde{q}(q_{\tau-2},s_{\tau-1},s_\tau))=arg\,max_{a\in A}1\cdot u(\bar{x},a)$. This also implies that D does not feel regret after any $s_\tau\in S$. Consequently, $q_{\tau-1}=\tilde{q}(q_{\tau-2},s_{\tau-1})$ maximises the second term of $V_{\tau-1}$. We therefore have $\tilde{q}(q_{\tau-2},s_{\tau-1})\in Q_{\tau-1}^*(q_{t-2},s_{\tau-1})$ for every $q_{\tau-2}\in Q$ and $s_{\tau-1}\in S$.

The argument for all periods $t<\tau-1$ is analogous to the proof of Proposition 1.■

PROOF OF LEMMA 1

Consider the case $q_0 < q_0^T$. We first characterise D's preferred beliefs. We have $q_0 < q^A$ since $q_0^T \leq q^A$. Thus, the bet on the bad state is D's reference action. For brevity, let π be the probability of betting on the good state. Given $s_1 = 1$, D's problem is to maximise

$$V_1 = q_0\sigma\left[\pi + (1-\pi)\alpha\right] + (1-q_0)(1-\sigma)\left[(1-\pi)\beta + \lambda\left((1-\pi)\beta - \beta\right)\right].$$

Since $dV_1/d\pi < 0 \Leftrightarrow q_0 < \beta(1-\sigma)(1+\lambda)/\left((1-\alpha)\sigma + \beta(1-\sigma)(1+\lambda)\right) \geq q_0^T$, we have $dV_1/d\pi < 0$ for all $q_0 < q_0^T$. Consequently, D prefers implementing the bet on the bad state, i.e., $Q_1^* = \left(0, q^A\right)$. As for adopted beliefs, we have $0 < \tilde{q} < q^A$ if $q_0 < q_0^B$. This implies $\hat{q}_1 = \tilde{q}$ by Assumption 2. In contrast, if $q_0^B \leq q_0 < q_0^T$, we have $\tilde{q} \geq q^A$. By Assumption 2, we now have $\hat{q}_1 = q_0$ since the posterior is no longer part of D's preferred beliefs.

If $q_0 > q_0^T$, two cases must be distinguished: Firstly, suppose $q_0^T = q^A$. We thus have $q_0 > q^A$ and

$$\hat{V}_1 = q_0\sigma\left[\pi + (1-\pi)\alpha + \lambda\left(\pi + (1-\pi)\alpha - 1\right)\right] + (1-q_0)(1-\sigma)\left[(1-\pi)\beta\right]$$

is appropriate. We have $d\hat{V}_1/d\pi > 0 \Leftrightarrow q_0 > \beta(1-\sigma)/\left((1-\alpha)\sigma(1+\lambda) + \beta(1-\sigma)\right) < q^A$. As a result, D prefers implementing the bet on the good state, i.e., $Q_1^* = \left(q^A, 1\right)$.

Secondly, suppose that we have $q_0^T < q^A$ implying $q_0^T = \beta(1-\sigma)(1+\lambda)/\left((1-\alpha)\sigma + \beta(1-\sigma)(1+\lambda)\right)$. In the case where $q_0^T < q_0 < q^A$, V_1 is the appropriate target function. We have $Q_1^* = \left(q^A, 1\right)$ because $dV_1/d\pi > 0 \Leftrightarrow$ $q_0 > \beta(1-\sigma)(1+\lambda)/\left((1-\alpha)\sigma + \beta(1-\sigma)(1+\lambda)\right) = q_0^T$. If $q_0 > q^A$, $\hat{V}_1$ is the appropriate target function, and $Q_1^* = \left(q^A, 1\right)$ follows from our previous discussion. Finally, if $q_0 = q^A$, we have

$$\bar{V}_1 = q_0\sigma\left[\pi + (1-\pi)\alpha + \lambda\left(\pi + (1-\pi)\alpha - 0.5(1+\alpha)\right)\right] + (1-q_0)(1-\sigma)\left[(1-\pi)\beta\right] \text{ if } \pi \leq 0.5$$

and

$$\bar{V}_1 = q_0\sigma\left[\pi + (1-\pi)\alpha\right] + (1-q_0)(1-\sigma)\left[(1-\pi)\beta + \lambda\left((1-\pi)\beta - 0.5\beta\right)\right] \text{ if } \pi > 0.5$$

with $\overline{V}_1$ being continuous in π. In the first case, $d\overline{V}_1/d\pi > 0$ follows from our discussion of $\hat{V}_1$. In the second case, $d\overline{V}_1/d\pi > 0$ follows from our discussion of V_1. We thus always have $Q_1^* = \left(q^A, 1\right)$ if $q_0 > q_0^T$. Since $q^A < \tilde{q} < 1$ holds for these priors, we have $\hat{q}_1 = \tilde{q}$ by Assumption 2.■

PROOF OF LEMMA 2

Consider first $q_0 > q_0^{T'}$. Since $q_0^{T'} \geq q^A$, betting on the good state is D's reference action. Taking as given $s_1 = 0$ and letting π denote the probability of the bet on the good state, D maximises

$$V_1 = q_0\left(1-\sigma\right)\left[\pi + \left(1-\pi\right)\alpha + \lambda\left(\pi + \left(1-\pi\right)\alpha - 1\right)\right] + \left(1-q_0\right)\sigma\left[\left(1-\pi\right)\beta\right].$$

We have $dV_1/d\pi > 0 \Leftrightarrow q_0 > \beta\sigma/\left(\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\lambda\right)+\beta\sigma\right) \leq q_0^{T'}$. As a result, D prefers implementing the bet on the good state, i.e., $Q_1^* = \left(q^A, 1\right)$. Since $\tilde{q} \leq q^A$ if $q_0^{T'} < q_0 \leq q_0^{B'}$ and $q^A < \tilde{q} < 1$ if $q_0 > q_0^{B'}$, we have $\hat{q}_1 = q_0$ in the former case and $\hat{q}_1 = \tilde{q}$ in the latter.

Consider next the case $q_0 < q_0^{T'}$. Suppose first that $q_0^{T'} = q^A$. In this case, betting on the bad state is D's reference action, which means she maximises

$$\hat{V}_1 = q_0\left(1-\sigma\right)\left[\pi + \left(1-\pi\right)\alpha\right] + \left(1-q_0\right)\sigma\left[\left(1-\pi\right)\beta + \lambda\left(\left(1-\pi\right)\beta - \beta\right)\right].$$

We have $d\hat{V}_1/d\pi < 0 \Leftrightarrow q_0 < \beta\sigma\left(1+\lambda\right)/\left(\left(1-\alpha\right)\left(1-\sigma\right)+\beta\sigma\left(1+\lambda\right)\right) > q^A$. Thus, D prefers the bet on the bad state, i.e., $Q_1^* = \left(0, q^A\right)$.

Suppose next that $q_0^{T'} > q^A$ implying $q_0^{T'} = \beta\sigma/\left(\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\lambda\right)+\beta\sigma\right)$. In this case, if $q^A < q_0 < q_0^{T'}$, V_1 is the appropriate target function, and we have $dV_1/d\pi < 0$ because this is equivalent to $q_0 < \beta\sigma/\left(\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\lambda\right)+\beta\sigma\right) = q_0^{T'}$. Secondly, if $q_0 < q^A$, $\hat{V}_1$ is appropriate and $d\hat{V}_1/d\pi < 0$ follows from our discussion above. Finally, consider the case $q_0 = q^A$. We have

$$\overline{V}_1 = q_0\left(1-\sigma\right)\left[\pi + \left(1-\pi\right)\alpha + \lambda\left(\pi + \left(1-\pi\right)\alpha - 0.5\left(1+\alpha\right)\right)\right] + \left(1-q_0\right)\sigma\left[\left(1-\pi\right)\beta\right] \text{ if } \pi \leq 0.5$$

and

$$\overline{V}_1 = q_0\left(1-\sigma\right)\left[\pi + \left(1-\pi\right)\alpha\right] + \left(1-q_0\right)\sigma\left[\left(1-\pi\right)\beta + \lambda\left(\left(1-\pi\right)\beta - 0.5\beta\right)\right] \text{ if } \pi > 0.5$$

with $\overline{V}_1$ being continuous in π. In the first case, $d\overline{V}_1/d\pi < 0$ follows from our discussion of V_1. In the second case, $d\overline{V}_1/d\pi < 0$ follows from our discussion of $\hat{V}_1$. As a result, we always have $Q_1^* = (0, q^A)$ if $q_0 < q_0^{T'}$. Since $0 < \tilde{q} < q^A$ in this case, $\hat{q}_1 = \tilde{q}$ holds by Assumption 2.■

PROOF OF PROPOSITION 4

Notice first that $q_0^T = q_0^{T'} = q^A$ if $\lambda = (2\sigma - 1)/(1-\sigma)$. As a result, any further increases of λ do not affect the width of the prior ranges. Consequently, $\partial(\Delta' - \Delta)/\partial\lambda = 0$ if $\lambda \geq (2\sigma - 1)/(1-\sigma)$. If $0 \leq \lambda < (2\sigma - 1)/(1-\sigma)$, we have

$$\Delta = \frac{\beta(1-\sigma)(1+\lambda)}{(1-\alpha)\sigma + \beta(1-\sigma)(1+\lambda)} - \frac{\beta(1-\sigma)}{(1-\alpha)\sigma + \beta(1-\sigma)}$$

and

$$\Delta' = \frac{\beta\sigma}{(1-\alpha)(1-\sigma) + \beta\sigma} - \frac{\beta\sigma}{(1-\alpha)(1-\sigma)(1+\lambda) + \beta\sigma},$$

which satisfy $\partial(\Delta' - \Delta)/\partial\lambda > 0$.

Comparing Δ' and Δ, it is also sufficient to limit attention to $0 \leq \lambda < (2\sigma - 1)/(1-\sigma)$. We have

$$\Delta' > \Delta \Leftrightarrow$$

$$\frac{\beta\sigma}{(1-\alpha)(1-\sigma) + \beta\sigma} - \frac{\beta\sigma}{(1-\alpha)(1-\sigma)(1+\lambda) + \beta\sigma} > \frac{\beta(1-\sigma)(1+\lambda)}{(1-\alpha)\sigma + \beta(1-\sigma)(1+\lambda)} - \frac{\beta(1-\sigma)}{(1-\alpha)\sigma + \beta(1-\sigma)} \Leftrightarrow$$

$$\frac{(\alpha-1)\beta\lambda(\sigma-1)\sigma}{(1+\alpha(\sigma-1)+(\beta-1)\sigma)(1+\lambda+\alpha(1+\lambda)(\sigma-1)-\sigma+\beta\sigma-\lambda\sigma)} >$$

$$\frac{(\alpha-1)\beta\lambda(\sigma-1)\sigma}{(\beta(\sigma-1)+(\alpha-1)\sigma)\beta(1+\lambda(\sigma-1)+(\alpha-1)\sigma)} \Leftrightarrow$$

$$(1+\alpha(\sigma-1)+(\beta-1)\sigma)(1+\lambda+\alpha(1+\lambda)(\sigma-1)-\sigma+\beta\sigma-\lambda\sigma) -$$

$$(\beta(\sigma-1)+(\alpha-1)\sigma)\beta(1+\lambda(\sigma-1)+(\alpha-1)\sigma) < 0 \Leftrightarrow$$

$\left(1-2\alpha+\alpha^2-\beta^2\right)\left(1+\lambda\left(\sigma-1\right)^2-2\sigma\right)<0$.

We have $1+\lambda\left(\sigma-1\right)^2-2\sigma<0$ because it is equivalent to $\lambda<\left(2\sigma-1\right)/\left(1-\sigma\right)^2>\left(2\sigma-1\right)/\left(1-\sigma\right)$. Moreover, $1-2\alpha+\alpha^2-\beta^2>0 \Leftrightarrow \alpha+\beta<1$ and $1-2\alpha+\alpha^2-\beta^2=0 \Leftrightarrow \alpha+\beta=1$.■

PROOF OF PROPOSITION 5

Consider mild regret first ($0<\lambda\leq\lambda^m$). We characterise D's choice set in the first period using backward induction. Consider therefore the second period. If $s_2=1$, the second-period prior where D is indifferent between implementing $a_2=0$ and $a_2=1$ is

$$q_1^T=\beta\left(1-\sigma\right)\left(1+\lambda\right)/\left(\left(1-\alpha\right)\sigma+\beta\left(1-\sigma\right)\left(1+\lambda\right)\right)$$

which satisfies $0<q_1^T<q^A$ given that $0<\lambda\leq\lambda^m$. The point of indifference after $s_2=0$ is

$$q_1^{T'}=\beta\sigma/\left(\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\lambda\right)+\beta\sigma\right)$$

which satisfies $q^A<q_1^{T'}<1$. As a result, the set of beliefs implementing pure strategies can be partitioned as follows

$$\left\{I,II,III,IV\right\}$$

where $I=\left(0,q_1^T\right)$, $II=\left(q_1^T,q^A\right)$, $III=\left(q^A,q_1^{T'}\right)$ and $IV=\left(q_1^{T'},1\right)$. We abstract from $q_1=q_1^T$, $q_1=q^A$ and $q_1=q_1^{T'}$, which implement mixed strategies and were ruled out above.

Consider now D's first-period problem given q_0 and $s_1=1$. The case $s_1=0$ is analogous and therefore omitted. Supposing $q_0<q^A$, the expected payoff of the pure strategies can be expressed as

$$V\left(I\right)=q_0\sigma\alpha+\left(1-q_0\right)\left(1-\sigma\right)\beta+q_0\sigma^2\alpha+\left(1-q_0\right)\left(1-\sigma\right)^2\beta+q_0\sigma\left(1-\sigma\right)\alpha+q_0\left(1-\sigma\right)\sigma\beta$$
$$V\left(II\right)=q_0\sigma\alpha+\left(1-q_0\right)\left(1-\sigma\right)\beta+q_0\sigma^2+\left(1-q_0\right)\left(1-\sigma\right)^2\left(-\lambda\beta\right)+q_0\sigma\left(1-\sigma\right)\alpha+q_0\left(1-\sigma\right)\sigma\beta$$
$$V\left(III\right)=q_0\sigma+\left(1-q_0\right)\left(1-\sigma\right)\left(-\lambda\beta\right)+q_0\sigma^2+q_0\sigma\left(1-\sigma\right)\left(\alpha+\lambda\left(\alpha-1\right)\right)+q_0\left(1-\sigma\right)\sigma\beta$$
$$V\left(IV\right)=q_0\sigma+\left(1-q_0\right)\left(1-\sigma\right)\left(-\lambda\beta\right)+q_0\sigma^2+q_0\sigma\left(1-\sigma\right).$$

Equating these expressions yields the various points of indifference. In particular,

$$q_0^{I\sim II} = \frac{\beta(1+\lambda)(1-\sigma)^2}{(1-\alpha)\sigma^2 + \beta(1+\lambda)(1-\sigma)^2}$$

$$q_0^{II\sim III} = \frac{\beta(1-\sigma)(1+\lambda\sigma)}{(1-\alpha)\sigma(1-\lambda(1-\sigma)) + \beta(1-\sigma)(1+\lambda\sigma)}$$

$$q_0^{III\sim IV} = \frac{\beta}{(1-\alpha)(1+\lambda)+\beta}.$$

We have $q_0^{I\sim II} < q_0^{II\sim III} < q_0^{III\sim IV}$ unless $\lambda = \lambda^m$, in which case $q_0^{I\sim II} < q_0^{II\sim III} = q_0^{III\sim IV}$. Additionally, we have $q_0^{I\sim II} < q_0^{I\sim III} < q_0^{II\sim III}$, $q_0^{I\sim II} < q_0^{I\sim IV} < q_0^{III\sim IV}$ and $q_0^{II\sim III} < q_0^{II\sim IV} < q_0^{III\sim IV}$ unless $\lambda = \lambda^m$, in which case $q_0^{II\sim III} = q_0^{II\sim IV} = q_0^{III\sim IV}$. Consequently, D prefers I if $q_0 < q_0^{I\sim II}$, II if $q_0^{I\sim II} < q_0 < q_0^{II\sim III}$, III if $q_0^{II\sim III} < q_0 < q_0^{III\sim IV}$ (unless $\lambda = \lambda^m$) and IV if $q_0 > q_0^{III\sim IV}$. Also, since $q_0^{III\sim IV} < q^A$, our initial supposition $q_0 < q^A$ is valid.

We now characterise the prior ranges $Q_0^-(s_1=1)$, $Q_0^-(s_1=1, s_2=1)$ and $Q_0^+(s_1=1, s_2=0)$ and show that these are the only instances of over- or underconfidence, which is moreover action-relevant. Consider $s_1 = 1$ first and define

$$Q_0^-(s_1=1) = \left[\beta(1-\sigma)/((1-\alpha)\sigma+\beta(1-\sigma)), q_0^{II\sim III}\right).$$

Since we have $q_0^{I\sim II} < \beta(1-\sigma)/((1-\alpha)\sigma+\beta(1-\sigma))$, D uniquely prefers $II = (q_1^T, q^A)$ for all $q_0 \in Q_0^-(s_1=1)$, while the Bayesian posterior given $s_1 = 1$ satisfies $\tilde{q} \geq q^A$ for these priors because it equals q^A if $q_0 = \beta(1-\sigma)/((1-\alpha)\sigma+\beta(1-\sigma))$ and increases in q_0. As a result, D exhibits action-relevant underconfidence for all $q_0 \in Q_0^-(s_1=1)$. There is neither over- nor underconfidence elsewhere. D prefers $I = (0, q_1^T)$ if $q_0 < q_0^{I\sim II}$ and both $I = (0, q_1^T)$ and $I = (q_1^T, q^A)$ if $q_0 = q_0^{I\sim II}$. Since $\tilde{q} = q_1^T$ if $q_0 = q_0^{I\sim II}$, D is not over- or underconfident for these priors as well as for $q_0^{I\sim II} < q_0 < \beta(1-\sigma)/((1-\alpha)\sigma+\beta(1-\sigma))$. Further, D prefers $III = (q^A, q_1^{T'})$ if $q_0^{II\sim III} \leq q_0 < q_0^{III\sim IV}$, both $III = (q^A, q_1^{T'})$ and $IV = (q_1^{T'}, 1)$ if $q_0 = q_0^{III\sim IV}$ and $IV = (q_1^{T'}, 1)$ if $q_0 > q_0^{III\sim IV}$. Since $\tilde{q} > q^A$ if $q_0 = q_0^{II\sim III}$ and $\tilde{q} = q_0^{T'}$ if $q_0 = q_0^{III\sim IV}$, there is again no over- or underconfidence.

Consider next $(s_1 = 1, s_2 = 1)$ and define

$$Q_0^-\left(s_1=1,s_2=1\right)=\left[\beta\left(1-\sigma\right)^2\Big/\left(\left(1-\alpha\right)\sigma^2+\beta\left(1-\sigma\right)^2\right),q_0^{I\sim II}\right).$$

Suppose that $q_0\in Q_0^-\left(s_1=1,s_2=1\right)$. Since $q_0<q_0^{I\sim II}$, D uniquely prefers beliefs after $\left(s_1=1,s_2=1\right)$ that implement $a_2=0$, which are collected in $\left(0,q^A\right)$. The lower bound of $Q_0^-\left(s_1=1,s_2=1\right)$ is the prior at which the Bayesian posterior given $\left(s_1=1,s_2=1\right)$ satisfies $\tilde{q}=q^A$. Thus, D is underconfident for all priors in $Q_0^-\left(s_1=1,s_2=1\right)$ and neither over- nor underconfident elsewhere because she prefers $\left(q^A,1\right)$ if $q_0\geq q_0^{I\sim II}$ and $\left(0,q^A\right)$ if $q_0<\beta\left(1-\sigma\right)^2\Big/\left(\left(1-\alpha\right)\sigma^2+\beta\left(1-\sigma\right)^2\right)$.

Finally, consider $\left(s_1=1,s_2=0\right)$ and define

$$Q_0^+\left(s_1=1,s_2=0\right)=\left(q_0^{III\sim IV},q^A\right].$$

Suppose that $q_0\in Q_0^+\left(s_1=1,s_2=0\right)$. Since $q_0>q_0^{III\sim IV}$, D uniquely prefers beliefs after $\left(s_1=1,s_2=0\right)$ that implement $a_2=1$, which are collected in $\left(q^A,1\right)$. The Bayesian posterior given $\left(s_1=1,s_2=0\right)$ satisfies $\tilde{q}=q_0$, which implies that it attains q^A at the upper bound of $Q_0^+\left(s_1=1,s_2=0\right)$. Consequently, D is overconfident for all priors in $Q_0^+\left(s_1=1,s_2=0\right)$ and neither over- nor underconfident otherwise because she prefers $\left(0,q^A\right)$ if $q_0\leq q_0^{III\sim IV}$ and $\left(q^A,1\right)$ if $q_0>q^A$.

We next address the case $\lambda^m<\lambda<\lambda^s$. We now have $q_0^{II\sim III}>q_0^{III\sim IV}$, which implies that no prior exists where D finds III optimal. D's choice between II and IV is governed by

$$q_0^{II\sim IV}=\frac{\beta\left(1-\sigma\right)\left(1+\sigma+\lambda\sigma\right)}{\left(1-\alpha\right)\left(2-\sigma\right)\sigma+\beta\left(1-\sigma\right)\left(1+\sigma+\lambda\sigma\right)}<q^A.$$

D prefers I if $q_0<q_0^{I\sim II}$, II if $q_0^{I\sim II}<q_0<q_0^{II\sim IV}$ and IV if $q_0>q_0^{II\sim IV}$. Accordingly, we can now define

$$Q_0^-\left(s_1=1\right)=\left[\beta\left(1-\sigma\right)\big/\left(\left(1-\alpha\right)\sigma+\beta\left(1-\sigma\right)\right),q_0^{II\sim IV}\right).$$

The Bayesian posterior given $s_1=1$ satisfies $\tilde{q}>q_1^{T'}$ at $q_0^{II\sim IV}$. This together with our previous discussion implies that there are no priors outside $Q_0^-\left(s_1=1\right)$ exhibiting over- or underconfidence. Regarding the second period, things are unchanged for $\left(s_1=1,s_2=1\right)$. As for $\left(s_1=1,s_2=0\right)$, define

$Q_0^+\left(s_1=1,s_2=0\right)=\left(q_0^{II\sim IV},q^A\right]$, Our conclusions remain the same as before because II and III both implement $a_2 = 0$ after $(s_1 = 1, s_2 = 0)$. As a result, D prefers $(0, q^A)$ if $q_0 \leq q_0^{II \sim IV}$ and uniquely $(q^A, 1)$ otherwise.

Finally, consider the case $\lambda^s \leq \lambda < 2\lambda^s$. If $q_1 < q^A$, D prefers to implement $a_2 = 0$ regardless of s_2. Likewise, she prefers $a_2 = 1$ if $q_1 > q^A$, while $q_1 = q^A$ is abstracted from because it implements a mixed strategy. As a result, two pure strategies remain implementable, namely, I and IV.

Regarding over- and underconfidence, define

$$Q_0^-(s_1 = 1) = \left[\beta(1-\sigma)/\left((1-\alpha)\sigma + \beta(1-\sigma)\right), q_0^{I \sim IV}\right)$$

where

$$q_0^{I \sim IV} = \frac{\beta(1-\sigma)(2+\lambda)}{(1-\alpha)2\sigma + \beta(1-\sigma)(2+\lambda)} < q^A.$$

If $q_0 < q_0^{I \sim IV}$, D prefers I, which translates into her uniquely preferring $(0, q^A)$ in the first period. If $q_0 \geq q_0^{I \sim IV}$, D prefers $(q^A, 1)$. The posterior $\tilde{q}$ attains q^A at $q_0 = \beta(1-\sigma)/\left((1-\alpha)\sigma + \beta(1-\sigma)\right)$. Thus, $Q_0^-(s_1 = 1)$ is the only instance of over- or underconfidence. As for the second-period, defining

$$Q_0^-(s_1 = 1, s_2 = 1) = \left[\beta(1-\sigma)^2/\left((1-\alpha)\sigma^2 + \beta(1-\sigma)^2\right), q_0^{I \sim IV}\right)$$

and

$$Q_0^+(s_1 = 1, s_2 = 0) = \left(q_0^{I \sim IV}, q^A\right]$$

yields the same kind of conclusion for $(s_1 = 1, s_2 = 1)$ and $(s_1 = 1, s_2 = 0)$ since the corresponding $\tilde{q}$ attains q^A at $q_0 = \beta(1-\sigma)^2/\left((1-\alpha)\sigma^2 + \beta(1-\sigma)^2\right)$ and $q_0 = q^A$, respectively.■

For the sake of completeness, we give the over- and underconfidence ranges for the signal histories beginning with $s_1 = 0$, which were omitted from the proof of Proposition 5. If $0 < \lambda \leq \lambda^m$, we have

$$\Delta^+(s_1 = 0) = \left(q_0^{II \sim III, s_1 = 0}, \beta\sigma/\left((1-\alpha)(1-\sigma) + \beta\sigma\right)\right]$$

$$\Delta^{+}\left(s_1=0,s_2=0\right)=\left(q_0^{III\sim IV,s_1=0},\beta\sigma^2\Big/\left(\left(1-\alpha\right)\left(1-\sigma\right)^2+\beta\sigma^2\right)\right]$$

$$\Delta^{-}\left(s_1=0,s_2=1\right)=\left[q^A,q_0^{I\sim II,s_1=0}\right)$$

where $q_0^{I\sim II,s_1=0}=\dfrac{\beta\left(1+\lambda\right)}{\left(1-\alpha\right)+\beta\left(1+\lambda\right)}$, $q_0^{II\sim III,s_1=0}=\dfrac{\beta\sigma\left(1-\lambda\left(1-\sigma\right)\right)}{\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\lambda\sigma\right)+\beta\sigma\left(1-\lambda\left(1-\sigma\right)\right)}$ and

$q_0^{III\sim IV,s_1=0}=\dfrac{\beta\sigma^2}{\left(1-\alpha\right)\left(1+\lambda\right)\left(1-\sigma\right)^2+\beta\sigma^2}$.

Secondly, if $\lambda^m<\lambda<\lambda^s$, we have

$$\Delta^{+}\left(s_1=0\right)=\left(q_0^{I\sim III,s_1=0},\beta\sigma\Big/\left(\left(1-\alpha\right)\left(1-\sigma\right)+\beta\sigma\right)\right]$$

$$\Delta^{+}\left(s_1=0,s_2=0\right)=\left(q_0^{III\sim IV,s_1=0},\beta\sigma^2\Big/\left(\left(1-\alpha\right)\left(1-\sigma\right)^2+\beta\sigma^2\right)\right]$$

$$\Delta^{-}\left(s_1=0,s_2=1\right)=\left[q^A,q_0^{I\sim III,s_1=0}\right)$$

Where $q_0^{I\sim III,s_1=0}=\dfrac{\beta\left(2-\sigma\right)\sigma}{\left(1-\alpha\right)\left(1-\sigma\right)\left(1+\sigma+\lambda\sigma\right)+\beta\left(2-\sigma\right)\sigma}$.

Finally, if $\lambda\geq\lambda^s$, we have

$$\Delta^{+}\left(s_1=0\right)=\left(q_0^{I\sim IV,s_1=0},\beta\sigma\Big/\left(\left(1-\alpha\right)\left(1-\sigma\right)+\beta\sigma\right)\right]$$

$$\Delta^{+}\left(s_1=0,s_2=0\right)=\left(q_0^{I\sim IV,s_1=0},\beta\sigma^2\Big/\left(\left(1-\alpha\right)\left(1-\sigma\right)^2+\beta\sigma^2\right)\right]$$

$$\Delta^{-}\left(s_1=0,s_2=1\right)=\left[q^A,q_0^{I\sim IV,s_1=0}\right)$$

where $q_0^{I\sim IV,s_1=0}=\dfrac{\beta 2\sigma}{\left(1-\alpha\right)\left(1-\sigma\right)\left(2+\lambda\right)+\beta 2\sigma}$.

PROOF OF PROPOSITION 6

Throughout, we consider the case $s_1 = 1$. The argument for $s_1 = 0$ is analogous. Suppose that $0 < \lambda < \lambda^s$. We now construct an example where D attains higher welfare being myopic. Let $q_0 \in \left(\beta(1-\sigma)^2 \big/ \left((1-\alpha)\sigma^2 + \beta(1-\sigma)^2\right), q_0^{I \sim II}\right)$. A Bayesian agent prefers II in this case, while a forward-looking D prefers I where the belief intervals implementing these strategies satisfy $I = \left(0, q_1^T\right)$ and $II = \left(q_1^T, q^A\right)$ (see the proof of Proposition 5). Suppose now that D chooses her belief myopically in the first period. D in this case prefers beliefs implementing $a_1 = 0$, which are given by $\left(0, q_1^T\right)$ and $\left(q_1^T, q^A\right)$, for all

$$q_0 < q_0^T = \beta(1-\sigma)(1+\lambda) \big/ \left((1-\alpha)\sigma + \beta(1-\sigma)(1+\lambda)\right).$$

Since $q_0^{I \sim II} < q_0^T$, D thus prefers $I = \left(0, q_1^T\right)$ and $II = \left(q_1^T, q^A\right)$. As D puts positive probability on all preferred beliefs, she implements both I and II with positive probability. A Bayesian, who prefers II to I, prefers any mixture of II and I to implementing I for sure. As a result, D's welfare is higher if she is myopic.

Consider now the case $\lambda = 0$. D's points of indifference if she is forward-looking are given by $q_0^{I \sim II} = \beta(1-\sigma)^2 \big/ \left((1-\alpha)\sigma^2 + \beta(1-\sigma)^2\right)$, $q_0^{II \sim III} = \beta(1-\sigma) \big/ \left((1-\alpha)\sigma + \beta(1-\sigma)\right)$ and $q_0^{III \sim IV} = q^A$, which coincide with those of a Bayesian and satisfy $q_0^{I \sim II} < q_0^{II \sim III} < q^A$. In contrast, if D is myopic, she prefers I and II if $q_0 < q_0^{II \sim III}$, while she prefers III and IV if $q_0 > q_0^{II \sim III}$ and is indifferent if $q_0 = q_0^{II \sim III}$. As a result, D cannot do better being myopic regardless of how she chooses among optimal beliefs.

Next, let $\lambda^s \leq \lambda < 2\lambda^s$. In this case, two pure strategies remain implementable, namely, I and IV, via $q_1 \in \left(0, q^A\right)$ and $q_1 \in \left(q^A, 1\right)$. If forward-looking, D is indifferent between I and IV at $q_0^{I \sim IV} = \beta(1-\sigma)(2+\lambda) \big/ \left((1-\alpha)2\sigma + \beta(1-\sigma)(2+\lambda)\right) < q^A$. In contrast, a Bayesian is indifferent between I and IV if $q_0 = \beta(1-\sigma) \big/ \left((1-\alpha)\sigma + \beta(1-\sigma)\right) < q_0^{I \sim IV}$. If D is myopic, she treats her first-period problem as a single- or last-period one. As a result, she prefers $\left(0, q^A\right)$ if $q_0 < q^A$ and $\left(q^A, 1\right)$ if $q_0 > q^A$. For all $q_0 < q_0^{I \sim IV}$, D prefers $\left(0, q^A\right)$ under both myopia and regret, which implies the same welfare. If $q_0 = q_0^{I \sim IV}$, D is indifferent between $\left(0, q^A\right)$ and $\left(q^A, 1\right)$ if forward-looking and prefers $\left(0, q^A\right)$ if myopic. Since a Bayesian prefers IV, forward-lookingness performs not worse. If

$q_0^{I\sim IV} < q_0 < q^A$, forward-lookingness performs better because it entails IV for sure, while myopia entails I. Finally, if $q_0 > q^A$, IV is chosen either way.

If $\lambda \geq 2\lambda^s$, a forward-looking D chooses $\left(0, q^A\right)$ for all $q_0 < q^A$ and $\left(q^A, 1\right)$ for all $q_0 > q^A$ and thus the same beliefs as under myopia. As a result, there is no difference between being myopic and forward-looking.

3.6 Appendix B: Eliciting Individual Regret Sensitivities

In this appendix, we show how to elicit an individual's regret parameter using an experimental choice task. Consider the following elicitation procedure for λ: A subject D is endowed with a prior q_0 on some state space $X = \{0,1\}$ where $q_0 = Pr\left(x=1\right) < 0.5$ and chooses once from the action set $A = \{0,1\}$. Ultimately, D receives 1 money unit if her action choice coincides with the state, i.e., if $x = 0$ and $a = 0$ or $x = 1$ and $a = 1$, and no money otherwise. Since the state does not engage D's ego, we plausibly have $u\left(x=1, a=1\right) = u\left(x=0, a=0\right) = 1$ as well as $u\left(x=1, a=0\right) = u\left(x=0, a=1\right) = 0$. This together with $q_0 < 0.5$ implies that $a = 0$ is D's reference action.

Using the strategy method, D is asked to make a contingent action choice before receiving a symmetric signal $s \in \{0,1\}$ on the state. Importantly, her choice not only conditions on the signal, but also on the *signal precision*, which is not revealed to D *ex ante*. More specifically, D is asked to indicate the signal precision making her indifferent between the two actions after observing $s = 1$. Letting $\tilde{\sigma} \in \left(0.5, 1\right)$ be the stated precision, D is paid according to the following rule: Given that $s = 1$ has materialised, action 1 (action 0) is implemented if and only if $\sigma > \tilde{\sigma}$ ($\sigma < \tilde{\sigma}$), i.e., if the actual signal precision exceeds (falls short of) the stated threshold, while the two actions are implemented with equal probability if $\sigma = \tilde{\sigma}$.

To see why $\tilde{\sigma}$ pins down D's regret parameter λ, express D's target function directly in terms of the probability of choosing action 1 denoted by π. For a given σ, D's problem is to maximise

$$V\left(\pi\right) = \frac{q_0\sigma}{q_0\sigma + \left(1-q_0\right)\left(1-\sigma\right)}\left(\pi\right) + \frac{\left(1-q_0\right)\left(1-\sigma\right)}{q_0\sigma + \left(1-q_0\right)\left(1-\sigma\right)}\left(1-\pi+\lambda\left(1-\pi-1\right)\right)$$

where $q_0\sigma \big/ \left(q_0\sigma + \left(1-q_0\right)\left(1-\sigma\right)\right)$ is the Bayesian posterior probability assigned to $x = 1$ given $s = 1$. Notice that D only feels regret in state 0 because $q_0 < 0.5$ implies that $a = 0$ is her reference action.

We have the following indifference condition:

$$dV(\pi)/d\pi = 0 \Leftrightarrow q_0\sigma + (1-q_0)(1-\sigma)(-1-\lambda) = 0 .$$

Solving for σ yields the precision making D indifferent as $\sigma = (1-q_0)(1+\lambda)/(1+\lambda(1-q_0))$.

PROPOSITION B1 In the elicitation procedure, it is a weakly dominant strategy for D to tell the truth, i.e., to choose $\tilde{\sigma} = (1-q_0)(1+\lambda)/(1+\lambda(1-q_0)) = \sigma(\lambda)$.

PROOF Consider any lying strategy $\tilde{\sigma}' < \sigma(\lambda)$. If the actual signal precision σ satisfies $\sigma > \sigma(\lambda)$, $\tilde{\sigma}' < \sigma(\lambda)$ performs as well as the truth-telling strategy $\tilde{\sigma} = \sigma(\lambda)$ because both strategies implement action 1. If $\sigma = \sigma(\lambda)$, $\tilde{\sigma}'$ again performs as well as $\tilde{\sigma}$ because D is now indifferent between the two actions and $\tilde{\sigma}'$ implements action 1, while $\tilde{\sigma}$ yields each action with equal probability. In contrast, if $\tilde{\sigma}' \le \sigma < \sigma(\lambda)$, $\tilde{\sigma}$ performs better than $\tilde{\sigma}'$: The former implements action 0, which is preferred by D since $\sigma < \sigma(\lambda)$, while the latter implements either action 1 or both actions with equal probability. Finally, if $\sigma < \tilde{\sigma}'$, both strategies again perform the same because both lead to action 0. All in all, any strategy $\tilde{\sigma}' < \sigma(\lambda)$ is therefore weakly dominated by $\tilde{\sigma} = \sigma(\lambda)$. An analogous argument can be constructed for all strategies $\tilde{\sigma}'' > \sigma(\lambda)$. Again, these are weakly dominated by $\tilde{\sigma} = \sigma(\lambda)$. As a result, it is weakly dominant for D to choose $\tilde{\sigma} = \sigma(\lambda)$.■

3.7 Appendix C: Anticipatory Utility and Asymmetry

In this section, we derive the predictions of Brunnermeier and Parker's (2005) model of "optimal expectations" about asymmetric information processing. The target function for belief selection in our one-period setting is given by

$$\breve{V}_1(q_1|q_0,s_1) = \sum_x \tilde{q}(x|q_0,s_1)U(x,q_1) + \sum_x q_1(x)U(x,q_1) .$$

Allow beliefs without full support and consider the case $\alpha < \beta$. Letting $\tilde{q}(x=1|q_0,s_1) = \tilde{q}$, we have

$\breve{V}_1 = \tilde{q} + q_1$ if $q_1 > q^A = \beta/(1-\alpha+\beta)$

$\breve{V}_1 = \tilde{q}\alpha + (1-\tilde{q})\beta + q_1\alpha + (1-q_1)\beta$ if $q_1 \leq q^A$

assuming for analytical convenience that D chooses $a_1 = 0$ if $q_1 = q^A$. Clearly, within these two intervals, the extreme values $q_1 = 1$ and $q_1 = 0$ are optimal, respectively. Inserting them, we have

$$\tilde{q} + 1 > \tilde{q}\alpha + (1-\tilde{q})\beta + \beta \Leftrightarrow \tilde{q} > q^A - (1-\beta)/(1-\alpha+\beta).$$

Thus, as soon as $\beta < 1$, D prefers $q_1 = 1$ before the posterior $\tilde{q}$ attains q^A. The intuition is that by believing in the good state she realises an anticipatory utility of 1, while believing in the bad state only yields her $\beta < 1$. Thus, D prefers $q_1 = 1$ implementing $a_1 = 1$ at priors where the objective performance of $a_1 = 1$ falls short of that of $a_1 = 0$.

If $q^A - (1-\beta)/(1-\alpha+\beta) < 0.5$, the model implies an asymmetry in information processing. To see this, suppose first that $q^A - (1-\beta)/(1-\alpha+\beta) = 0.5$. In this case, D chooses $q_1 = 0$ after $s_1 = 1$ if $q_0 < 1-\sigma < 0.5$ and $q_1 = 1$ if $q_0 > 1-\sigma$, while she chooses $q_1 = 1$ after $s_1 = 0$ if $q_0 > \sigma > 0.5$, which amounts to symmetry. This changes if $q^A - (1-\beta)/(1-\alpha+\beta) < 0.5$: After $s_1 = 1$, D now prefers $q_1 = 1$ if $q_0 > 1-\sigma-x$ for some $x > 0$. After $s_1 = 0$, she prefers $q_1 = 1$ if $q_0 > \sigma - y$ for some $y > 0$. Thus, if $q_0 \in (1-\sigma-x, 1-\sigma+y)$, D chooses $q_1 = 1$ both after $s_1 = 1$ and after $s_1 = 0$ given the counterfactual prior $1-q_0$. This can be interpreted as the good signal carrying more weight.

Consider next the case $\alpha > \beta$. Within the two belief intervals, $q_1 = 1$ and $q_1 = q^A$ are now optimal, respectively. Intuitively, since $\alpha > \beta$, D wants to assign maximal probability to the good state (subject to implementing a given action) in order to flatter her ego. Inserting these values, we have

$$\tilde{q} + 1 > \tilde{q}\alpha + (1-\tilde{q})\beta + q^A \Leftrightarrow \tilde{q} > q^A - (1-\alpha)/(1-\alpha+\beta)^2 .$$

Now, the model generates an asymmetry in information processing that is more extreme than in the previous case. In fact, chosen beliefs always exceed the Bayesian posterior because D prefers $q_1 = q^A$ if $\tilde{q} < q^A - (1-\alpha)/(1-\alpha+\beta)^2 < q^A$ and $q_1 = 1$ if $\tilde{q} > q^A - (1-\alpha)/(1-\alpha+\beta)^2$. Belief revision is therefore tilted towards the good state.

BIBLIOGRAPHY

Akerlof, George A. and William T. Dickens (1982). “The economic consequences of cognitive dissonance”. *The American Economic Review*, 72(3), 307-319.

Battigalli, Pierpaolo and Martin Dufwenberg (2007). “Guilt in Games”. *The American Economic Review (P&P)*, 97(2), 170-176.

Battigalli, Pierpaolo and Martin Dufwenberg (2009). “Dynamic psychological games”. *Journal of Economic Theory*, 144(1), 1-35.

Barigozzi, Francesca and Rosella Levaggi (2010). “Emotional decision-makers and anomalous attitudes towards information”. *Journal of Risk and Uncertainty*, 40(3), 255-280.

Bartling, Björn and Urs Fischbacher (2011). “Shifting the Blame. On Delegation and Responsibility”. *Review of Economic Studies*, 79(1), 67-87.

Bell, David E. (1982). “Regret in decision making under uncertainty”. *Operations Research*, 30(5), 961-981.

Benoît, Jean Pierre and Juan Dubra (2011). “Apparent overconfidence”. *Econometrica*, 79(5), 1591-1625.

Brandts, Jordi and Carles Solá (2001). “Reference Points and Negative Reciprocity in Simple Sequential Games”. *Games and Economic Behavior*, 36(2), 138-157.

Brunnermeier, Markus K. and Jonathan A. Parker (2005). “Optimal Expectations”. *The American Economic Review*, 95(4), 1092-1118.

Caplin, Andrew and John Leahy (2001). “Psychological expected utility theory and anticipatory feelings”. *The Quarterly Journal of Economics*, 116(1), 55-79.

Charness, Gary and Matthew Rabin (2002). “Understanding Social Preferences With Simple Tests”. *The Quarterly Journal of Economics*, 117(3), 817-869.

Coffman, Lucas C. (2011). “Intermediation Reduces Punishment (and Reward)”. *American Economic Journal. Microeconomics*, 3(4), 77-106.

Cox, James C. (2004). “How to identify trust and reciprocity”. *Games and Economic Behavior*, 46(2), 260-281.

Cox, James C., Daniel Friedman and Vjollca Sadiraj (2008). “Revealed Altruism”. *Econometrica*, 76(1), 31-69.

Cox, James C., Maroš Servátka and Radovan Vadovič (2010). “Saliency of outside options in the lost wallet game”. *Experimental Economics*, 13(1), 66-74.

Dhaene, Geert and Jan Bouckaert (2010). “Sequential reciprocity in two-player, two-stage games. An experimental analysis”. *Games and Economic Behavior*, 70(2), 289-303.

Dohmen, Thomas, Armin Falk, David Huffman, Felix Marklein and Uwe Sunde (2009). "The non-use of Bayes rule: representative evidence on bounded rationality", Working Paper RM09038.

Dufwenberg, Martin and Uri Gneezy (2000). “Measuring beliefs in an experimental lost wallet game”. *Games and Economic Behavior*, 30(2), 163-182.

Dufwenberg, Martin and Georg Kirchsteiger (2004). “A theory of sequential reciprocity”. *Games and Economic Behavior*, 47(2), 268-298.

Eliaz, Kfir and Ran Spiegler (2006). “Can anticipatory feelings explain anomalous choices of information sources?”. *Games and Economic Behavior*, 56(1), 87-104.

Epstein, Larry G., Jawwad Noor and Alvaro Sandroni (2010). “Non-bayesian learning”. *The BE Journal of Theoretical Economics*, 10(1).

Eyster, Erik (2002). “Rationalizing the past: A taste for consistency”. Working Paper Nuffield College Oxford.

Falk, Armin, Ernst Fehr and Urs Fischbacher (2003). “On the nature of fair behavior”. *Economic Inquiry*, 41(1), 20-26.

Falk, Armin and Urs Fischbacher (2006). “A theory of reciprocity”. *Games and Economic Behavior*, 54(2), 293-315.

Falk, Armin and Michael Kosfeld (2006). “The hidden costs of control”. *The American Economic Review*, 96(5), 1611-1630.

Falk, Armin and Florian Zimmermann (2011). “Preferences for consistency”. IZA Working Paper No.5840.

Fehr, Ernst and Klaus M. Schmidt (1999). “A theory of fairness, competition, and cooperation”. *The Quarterly Journal of Economics*, 114(3), 817-868.

Fehr, Ernst and Klaus M. Schmidt (2006). “The economics of fairness, reciprocity and altruism – experimental evidence and new theories”. *Handbook on the Economics of Giving, Reciprocity and Altruism*, 1, 615-691.

Fershtman, Chaim and Uri Gneezy (2001). “Strategic delegation. An experiment”. *RAND Journal of Economics*, 32(2), 352-368.

Friedman, Daniel (1998). “Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly.” *The American Economic Review*, 88(4), 933-946.

Gächter, Simon and Ernst Fehr (2000). "Cooperation and punishment in public goods experiments". *The American Economic Review*, 90(4), 980-994.

Geanakoplos, John, David Pearce and Ennio Stacchetti (1989). "Psychological games and sequential rationality". *Games and Economic Behavior*, 1(1), 60-79.

Goeree, Jacob K., Charles A. Holt and Thomas R. Palfrey (2008). "Quantal response equilibria". In Steven N. Durlauf and Lawrence E. Blume ed.s., *The New Palgrave Dictionary of Economics (Second Edition)*, Palgrave Macmillan.

Hamman, John R., George Loewenstein and Roberto A. Weber (2010). "Self-interest through delegation. An additional rationale for the principal-agent relationship". *The American Economic Review*, 100(4), 1826-1846.

Herweg, Fabian and Konrad Mierendorff (2013). "Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs". *Journal of the European Economic Association*, 11, 399–432.

Herweg, Fabian and Daniel Müller (2011). "Performance of Procrastinators: On the Value of Deadlines". *Theory and Decision*, 70(3), 329-366.

Herweg, Fabian and Klaus M. Schmidt (2013). "A Theory of Ex-Post Inefficient Renegotiation". SFB TR/15 Discussion Paper 390.

Kahneman, Daniel and Amos Tversky (1973). "On the psychology of prediction". *Psychological Review*, 80(4), 237-251.

Kahneman, Daniel and Amos Tversky (1979). "Prospect theory. An analysis of decision under risk". *Econometrica*, 47(2), 263-291.

Kőszegi, Botond (2003). "Health anxiety and patient behavior". *Journal of Health Economics*, 22(6), 1073-1084.

Köszegi, Botond (2006). “Ego utility, overconfidence, and task choice”. *Journal of the European Economic Association*, 4(4), 673-707.

Köszegi, Botond and Matthew Rabin (2006). “A model of reference-dependent preferences”. *The Quarterly Journal of Economics*, 121(4), 1133-1165.

Köszegi, Botond and Matthew Rabin (2009). “Reference-Dependent Consumption Plans”. *The American Economic Review*, 99(3), 909-36.

Loomes, Graham and Robert Sugden (1982). “Regret theory: An alternative theory of rational choice under uncertainty”. *The Economic Journal*, 92(368), 805-824.

Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green (1995). *Microeconomic Theory*. New York: Oxford University Press.

McFadden, Daniel (1974). “Conditional logit analysis of qualitative choice behavior”. In P. Zarembka ed., *Frontiers in Econometrics*, Academic Press, New York, 105-142.

McCabe, Kevin A., Mary L. Rigdon and Vernon L. Smith (2003). “Positive reciprocity and intentions in trust games”. *Journal of Economic Behavior & Organization*, 52(2), 267-275.

McKelvey, Richard D. and Thomas R. Palfrey (1995). “Quantal response equilibria for normal form games”. *Games and Economic Behavior*, 10(1), 6-38.

McKelvey, Richard D. and Thomas R. Palfrey (1996). “A Statistical Theory of Equilibrium in Games”. *Japanese Economic Review*, 47(2), 186-209.

Möbius, Markus M., Muriel Niederle, Paul Niehaus and Tanya S. Rosenblat (2012). “Managing Self-Confidence: Theory and Experimental Evidence”. Working Paper Stanford University.

O'Donoghue, Ted and Matthew Rabin (2001). "Choice and procrastination". *The Quarterly Journal of Economics*, 116(1), 121-160.

Palacios-Huerta, Ignacio (2003). "Learning to open Monty Hall's doors". *Experimental Economics*, 6(3), 235-251.

Rabin, Matthew (1993). "Incorporating fairness into game theory and economics". *The American Economic Review*, 83(5), 1281-1302.

Rabin, Matthew and Joel L. Schrag (1999). "First impressions matter: A model of confirmatory bias". *The Quarterly Journal of Economics*, 114(1), 37-82.

Rabin, Matthew (2002). "A perspective on psychology and economics". *European Economic Review*, 46(4), 657-685.

Servátka, Maroš and Radovan Vadovic (2009). "Unequal Outside Options in the Lost Wallet Game". University of Canterbury Working Paper.

Samuelson, William and Richard Zeckhauser (1988). "Status quo bias in decision making". *Journal of Risk and Uncertainty*, 1, 7-59.

Shalev, Jonathan (2000). "Loss aversion equilibrium". *International Journal of Game Theory*, 29(2), 269-287.

Von Siemens, Ferdinand (2011). "Intention-Based Reciprocity and the Hidden Costs of Control". Tinbergen Institute Discussion Paper.

Yariv, Leeat (2005). "I'll See It When I Believe It - A Simple Model of Cognitive Consistency". Working Paper Caltech.