İrem Bozbay

Bargaining and Judgment Aggregation

İrem Bozbay

This book was typeset by the author using LaTeX
Cover design and illustration by: Caner Atakul
Printed by: Datawyse Maastricht

Bargaining and Judgment Aggregation

Dissertation

to obtain the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus, Prof. Mr. G.P.M.F. Mols,
in accordance with the decision of the Board of Deans,
to be defended in public on Thursday, April 26, 2012, at 16:00 hours

by

İrem Bozbay

Supervisors
Prof. dr. Hans J. M. Peters
Prof. dr. Franz Dietrich (University of East Anglia, Norwich, United Kingdom and University of Paris-Descartes, Paris, France)

Assessment Committee
Prof. dr. Rudolf Müller (chair)
Prof. dr. Clemens Puppe (Karlsruhe Institute of Technology, Karlsruhe, Germany)
Dr. A.J.A. Storcken

Acknowledgements

I would like to take the opportunity here to thank a number of people whose support helped me a lot during the making of this dissertation.

First and foremost, I would like to express my gratitude to my supervisors, Hans Peters and Franz Dietrich. I feel privileged to have had the opportunity to work with them. During our weekly meetings with Hans, I benefited from his excellent teaching, intellectual guidance and invaluable comments. No doubt, he has been a most important individual in the development and completion of my PhD. Simply, he has always been there for me, helping me out on countless occasions. I owe a great debt of gratitude for his patience, inspiration, encouragement and friendship.

Although Franz was in London, he dedicated his valuable time and expertise to the success of this dissertation. It has been an honor to be Franz's first Ph.D. student. The joy and enthusiasm he has for his research has been contagious and motivational for me. I have learned tremendously from working with him, from the many meetings and discussions we had over the phone and during my visit to LSE. I thank him wholeheartedly.

I'd like to thank the committee members Rudolf Müller, Clemens Puppe and Ton Storcken for the time and effort they have all put into scrutinizing my work and offering their feedback.

I thank my colleagues and fellow Ph.D. students at the Department of Quantitative Economics for providing a pleasant working environment. In particular, I am extremely grateful to Anita van den Berg (who has been too kind to offer her help with the Dutch summary), to Hande Karabıyık and Birol Yüceoğlu for their company. I thank Dries Vermeulen for taking interest in my work and giving very helpful comments. Many thanks to Haydeé Hallmanns and Karin van den Boorn for their administrative support.

Many influences played a role in bringing me to Maastricht for Ph.D., but one in particular stands out. I would like to express my unending gratitude to M. Remzi Sanver for his trust in me and for guiding me through pursuing an academic career. I am extremely indebted to Jean Laine for his support and encouragement during my master's study.

A very stimulating period during my Ph.D. was my visit to Centre for Philosophy of Natural and Social Science (CPNSS) at London School of Economics. I benefited from great discussions during seminars of the Choice Group. I thank CPNSS for hosting me.

My first year in Maastricht was the most difficult (though the most fun). I feel forever grateful to Çağatay Kayı for the role he played in my adaptation to Maastricht and Ph.D. life, for his comments on my work, for being a mentor and a good friend. Many thanks to Betigül Onay and Melida Tüzünoğlu for making Maastricht a lively and joyful place.

I thank my dearest friends in İstanbul; Berk, Elif, Esin, Gamze, Murat, Tuğçe and Zeynep, for making every homecoming a great pleasure. I have been comfortable in the knowledge that there is a city full of great friends waiting for me.

Many thanks to Özer, for his love, care, patience, encouragement, commitment and friendship. This dissertation would not have been possible without him in my life.

Finally, I owe my deepest gratitude to my parents Lütfiye and Seçkin Bozbay and my sister Hikmet Dinçer, for their endless support and love. I feel so grateful to be surrounded by such a caring, loving and fun family.

İrem Bozbay
Maastricht, March 2012

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis is a collection of essays on two different research areas in the field of economic theory; namely, bargaining and judgment aggregation. The first part of the thesis addresses bargaining problems and consists of Chapter 2. The second part deals with the theory of judgment aggregation and consists of Chapter 3, Chapter 4 and Chapter 5.

Part I: Bargaining

A bargaining problem is a problem of understanding how people cooperate for their mutual benefit when non-cooperation is bad for all of them. In the classical two-person bargaining situation, two individuals try to reach an agreement regarding how to share some good. If they do not reach an agreement, they end up receiving some 'fixed' allocation which is worse than any allocation they would get if they agreed. Following Nash (1950), a classical bargaining problem is defined by a utility possibility set and a disagreement outcome, which, as the name suggests, is the resulting allocation if there is no agreement. We, however, consider bargaining problems in which no exogenous disagreement is given. We assume that the disagreement outcome is determined endogenously, namely by the bargaining solution. A bargaining solution assigns a pair of outcomes – the compromise outcome and the disagreement outcome – to every bargaining problem. This is different from the classical approach where a solution assigns only the compromise outcome.

The disagreement outcome in the classical bargaining problem serves as a reference point and enables comparisons of different allocations. From a positive point of view, bargaining solutions aim to predict the outcome of a bargaining process. From a normative point of view, bargaining solutions should propose a reasonable or fair outcome like an arbitrator or mediator would do. In our framework where the disagreement outcome is endogenously determined, it would then be proposed by the mediator (or arbitrator) and operate as a reference point or threat. According to the positive interpretation of our problem, the compromise outcome is the predicted outcome of bargaining process in case of agreement (without mediation or arbitration) and the disagreement outcome is the predicted outcome in case agreements fail. Some of these interpretations come close to so-called *Alternative Dispute Resolution* (ADR), an increasingly common form of negotiation which aims to avoid court cases by helping parties reach an agreement beforehand, commonly used in Anglo-Saxon common law systems. Different forms of ADR processes correspond to bargaining problems with endogenous disagreement from the perspective of the normative interpretation. In Chapter 2, we present a detailed discussion of how ADR processes and bargaining problems with endogenous disagreement outcome are related.

Vartiainen (2007) studies bargaining problems with endogenous disagreement outcome. He extends the classical Nash bargaining solution to this framework and axiomatically characterizes the *extended* Nash bargaining solution. For this framework we propose and study an extension of the classical Kalai-Smorodinsky bargaining solution. We identify the (large) domain on which this solution is single-valued, and present two axiomatic characterizations on subsets of this domain. Our first characterization theorem is based on an axiom called *Independence of Non-Utopia information* (INU), which states that the compromise and disagreement outcomes in two different problems should be the same if the associated utopia and anti-utopia points coincide. This is a relatively strong condition, and, in the second characterization this axiom is replaced by three axioms with direct counterparts in classical bargaining theory. This chapter has led to the paper called *Bargaining with endogenous disagreement: The extended Kalai-Smorodinsky solution* which is published in Games and Economic Behavior, 74 (2012), 407–417.

Part II: Judgment aggregation

We address judgment aggregation problems in the second part. A judgment aggregation problem arises whenever a group needs to make a collective yes/no judgment on several (possibly interconnected) propositions based on group members' judgments on these propositions. A simple example is the problem of a jury in a court trial having to reach a collective judgment on whether the defendant has broken the contract and whether the contract is legally valid. Following a generally recognized legal doctrine of common law systems, the defendant is judged to be guilty if and only if both propositions are collectively accepted. There are three issues in this example, and a yes judgment on each of the first two issues and a no judgment on the last issue are inconsistent. A consistent judgment is free from all logical inconsistencies. The theory of judgment aggregation substantially deals with the question of whether (and when) it is possible to reach consistent group judgments which are *fair* to group members. Another question that immediately arises is whether and how we can reach *true* group judgments. The procedural approach in judgment aggregation deals with the former while the epistemic approach in judgment aggregation deals with the latter.

We take the epistemic approach. The epistemic approach aims to track the truth. When it comes to aggregating judgments rather than preferences, this approach seems very natural. In the court trial example, the jury's main target seems to be to find out two independent facts rather than to reach a conclusion that is fair to jurors. This approach has been much less explored than the more common procedural approach in judgment aggregation theory while it is well-established in a literature which deals with voting between two alternatives, or equivalently, with single-issue judgment aggregation problems (Austen-Smith and Banks, 1996 and Feddersen and Pesendorfer, 1997). In Part II, we apply their methods and results to multi-issue judgment aggregation problems and extend their work beyond single-issue agendas.

All three chapters in Part II assume that there is a group of voters having to accept or reject each of two propositions. Each proposition is factually true or false and voters hold private information about which propositions are true. Voters share a common preference for true collective judgments. Chapter 3 considers two independent propositions. Chapter 4 adds interconnections between the two propositions. Chapter 5 still assumes independence but the private information structure is different. While Chapter 3 and 4 consider discrete binary private information for each proposition, private information is continuous rather than binary in Chapter 5. Our aim is common among the three chapters: we want to analyse the resulting strategic incentives and determine which voting rules lead to collective judgments that efficiently use all private information under each setting. In the setting of Chapter 3, we find that in many, but not all cases a quota rule should be used, which decides on each issue according to whether the number of yes votes exceeds a particular quota. When interconnections are introduced, this result does not persist. Chapter 4 characterizes the rare situations in which efficient information aggregation is possible with a voting rule, and gives the nature of such rules. We only focus on quota rules in Chapter 5, and we find that efficient information aggregation by quota rules is not always possible.

Part I

Bargaining

Chapter 2

Bargaining with endogenous disagreement

2.1 Introduction

In the bargaining problem of Nash (1950) each player can unilaterally enforce the disagreement outcome if negotiations fail. In some cases, however, it may not be clear what the disagreement outcome is or whether the players can, or want to, enforce it if agreement is not reached. In the classical example of employer-union wage negotiations the union can call out a strike if it is not satisfied with the wage offered by the employer. But how long should the strike last? What will be its consequences? Will all workers join? Are there perhaps different and better ways to put pressure on management? Also, which outcome can the employer enforce, if any, in case no agreement is reached?

In this chapter, following Vartiainen (2007), we assume that the disagreement outcome is determined endogenously, namely by the bargaining solution. Specifically, the bargaining solution assigns a *pair* of outcomes, namely a compromise outcome and a disagreement outcome. The possible interpretations of such a bargaining solution are parallel to the usual interpretations of a classical bargaining solution in the situation where the disagreement outcome is exogenous. From a positive point of view, a *classical* bargaining solution predicts or describes the compromise outcome, i.e., it tells us what this outcome is *given* that the players reach agreement. From this point of view, a bargaining solution in the situation *without* exogenous disagreement outcome predicts both the compromise outcome for the case that the players reach an agreement, and the disagreement outcome in the opposite case. From a normative point of view, a *classical* bargaining solution functions like an outside arbitrator and proposes a compromise outcome, but this outcome should be 'reasonable' given the exogenous disagreement outcome. In the situation *without* exogenous disagreement outcome, a bargaining solution proposes a compromise as well as a disagreement outcome, such that the compromise is 'reasonable' when compared to the disagreement outcome. We will refine and detail these interpretations in Section 2.2 below.

Within this framework, Vartiainen (2007) proposes and axiomatically characterizes a bargaining solution which extends the classical Nash bargaining solution for bargaining problems with fixed, exogenous disagreement point. That solution maximizes the Nash product, i.e., the product of the gains of the players from the compromise outcome over the disagreement outcome.

By contrast, the solution proposed in our work depends explicitly on the utopia point and extends the solution of Raiffa (1953) and Kalai-Smorodinsky (1975) for classical bargaining problems. This extension works as follows. First, the assigned compromise point is indeed the classical Kalai-Smorodinsky (KS) outcome for the assigned disagreement outcome. That is, it is the Pareto optimal point on the straight line joining this disagreement outcome and its associated utopia point. Second, the assigned disagreement outcome is the point on the straight line joining the assigned compromise point and the associated 'anti-utopia point', obtained by taking the minimum utilities of the players below the compromise point; it is, thus, a 'converse' KS outcome. The main original condition justifying the classical Kalai-Smorodinsky solution is individual monotonicity: it implies that if the utopia point stays fixed, then the players should benefit from increased availability of favorable outcomes. In defining the KS solution for the case where the disagreement outcome is not exogenous, we thus apply the same logic also to the determination of the disagreement outcome: given that the anti-utopia point does not change, the players should suffer from the increased availability of unfavorable outcomes.

We present two axiomatic characterizations of this solution. In the first one, the crucial axiom is called Independence of Non-Utopia information (INU). This condition is relatively strong and, under an additional condition, says that the compromise and disagreement outcomes in two different problems should be the same if the associated utopia and anti-utopia points coincide. In the second characterization, INU is replaced by three much weaker axioms, including a monotonicity condition.

Another extension of the Kalai-Smorodinsky solution to bargaining problems without fixed disagreement point is proposed in Vartiainen (2002), but this solution is quite different from our extension.[1]

The framework in our work and in Vartiainen (2007) has resemblance to the one in Thomson (1981), who also considers bargaining problems without disagreement

[1] It assigns the points of intersection of the straight line connecting the *global* utopia and anti-utopia points with the boundary of the feasible set and, thus, extends the Kalai-Rosenthal (Kalai and Rosenthal, 1978) solution rather than the Kalai-Smorodinsky solution.

point. Thus, a bargaining problem is defined merely as a utility-possibility set. Thomson introduces the notion of *reference point* as a *function* of the bargaining problem.[2] The key difference to the classical disagreement point is that no player can unilaterally enforce the reference point. It may thus serve, rather, as a hypothetical outcome to which the players compare proposals made during negotiations. The key difference with our (and Vartiainen's) approach is that we assume that also the reference point (disagreement outcome) is determined by the solution.

In situations where an arbitrator, or a mediator, makes choices for the players (cf. Luce and Raiffa, 1957), the reference point may also result from a noncooperative, strategic game between the players, and the arbitrator (bargaining solution) assigns a compromise point based on the reference point. Effectively, this way a noncooperative game is turned into a strictly competitive game which may have a value, comparable to a zero-sum game. Such arbitration games have received renewed attention recently, see Kalai and Kalai (2010).

In Section 2.2 we present a more detailed discussion of bargaining with endogenous disagreement. In Section 2.3 we formally introduce the extended Kalai-Smorodinsky solution, show that it is non-empty valued and characterize the domain of bargaining problems for which it is single-valued. In Section 2.4 we present two axiomatic characterizations of the solution on domains where it is single-valued. We also show that the axioms in these characterizations are independent.

All proofs are collected in the Appendix.

Notation For $x, y \in \mathbb{R}^2$, $x > y$ means $x_i > y_i$ and $x \geqslant y$ means $x_i \geqslant y_i$ for $i = 1, 2$. Similarly for $<$ and $\leqslant$. By $[x, y]$ we denote the line segment with endpoints x and y. The cardinality of a set $X \subseteq \mathbb{R}^2$ is denoted by $|X|$. For $a, x \in \mathbb{R}^2$, $ax := (a_1x_1, a_2x_2)$, $aX := \{ax \mid x \in X\}$, and $a + X := \{a + x \mid x \in X\}$. The set $(-1, -1)X$ is also denoted by $-X$. By $\mathbb{R}^2_+$ we denote the (strictly) positive quadrant of $\mathbb{R}^2$. By $\text{conv}(X)$ we denote the convex hull of the set X.

2.2 Bargaining with endogenous disagreement

A *bargaining problem* U is a compact and convex subset of $\mathbb{R}^2$ such that $x > y$ for some $x, y \in U$. Elements of U are called *outcomes* and represent the utilities of two players. By $\mathcal{U}$ we denote the set of all bargaining problems.

A *classical bargaining problem* is a pair (U, d), where $U \in \mathcal{U}$ and $d \in U$; the outcome d is called the *disagreement outcome*, and it results if the players do not reach agreement. By $\mathcal{B}$ we denote the set of all classical bargaining problems. A *classical bargaining solution* is a map $F : \mathcal{B} \to \mathbb{R}^2$ with $F(U, d) \in U$ for all $(U, d) \in \mathcal{B}$.

[2]Herrero (1998) considers endogenous reference points in so-called bargaining problems with claims. Also these reference points are a function of the bargaining problem and, in this case, the claims point.

In contrast, a *bargaining solution* or, briefly, a *solution* is a correspondence $f : \mathcal{U} \to \mathbb{R}^2 \times \mathbb{R}^2$ such that $s, r \in U$ and $s \neq r$ for all $U \in \mathcal{U}$ and $(s, r) \in f(U)$. For a pair $(s, r) \in f(U)$, we call s the *compromise outcome* and r the *disagreement outcome*.

We now discuss how solutions with endogenous disagreement can be interpreted, also offering some perspectives that go beyond Vartiainen (2007).[3] Classical bargaining theory commonly distinguishes between positive interpretations, according to which bargaining solutions aim to predict the outcome of a bargaining process, and normative interpretations, according to which solutions express a judgment of what outcome would be normatively 'best' or 'fairest' and should therefore be proposed by an arbitrator or mediator if such a person is appointed. Since our extended bargaining solutions return two outcomes – a compromise outcome s and a disagreement outcome r – we may classify potential interpretations according to which of the two outcomes are interpreted positively ('players') and which normatively ('mediator'). This yields four possible interpretations overall, see Table 2.2.1.

Table 2.2.1: Four potential interpretations of bargaining solutions

		disagreement r	
		players	mediator
compromise s	players	*Case 1*	*Case 3*
	mediator	*Case 2*	*Case 4*

We discuss these cases in turn. In the 'doubly positive' Case 1, s is the predicted compromise outcome of bargaining (without arbitration or mediation), and r the predicted outcome failing agreement. Outcome r plays the role of players' mental reference point, representing their common beliefs of what would happen failing agreement. Both s and r are predicted to emerge as the result of the bargaining process, in the course of which various proposals and threats might have been on the table.[4]

Cases 2, 3 and 4 represent three variants of how a mediator could intervene in the bargaining process. These variants are not merely hypothetical but can be observed in practice. For instance, they correspond to different forms of so-called *Alternative Dispute Resolution* (ADR). Especially in Anglo-Saxon common law systems, ADR has become a wide-spread practice aimed to avoid costly and lengthy court trials

[3]This discussion has benefitted from helpful comments of an anonymous referee.

[4]Case 1 is perhaps closest to the idea of a reference point as in Thomson (1981).

through reaching a compromise beforehand.[5] While ADR always assigns a central role to a so-called mediator, the exact nature of this role differs across different forms of ADR. Bargaining theory with endogenous disagreement allows one to study some existing forms of ADR in virtue of the interpretations of Cases 2, 3 and 4. To understand why this is so, two pieces of background information are worth noting. Firstly, the role of the mediator in ADR does typically *not* consist in elaborating a binding compromise. Instead, any compromise needs both parties' approval. Should this compromise have been proposed by the mediator, this proposal was non-binding. This marks a key difference between ADR and orthodox forms of dispute resolution such as *court trials* and *arbitration*; there, the role of the judge resp. arbitrator is precisely to dictate a binding compromise.[6] Secondly, prior to entering ADR both parties have contractually agreed to the mediator's precise role, whatever this role consists in. So, parties cannot later withdraw from the ADR procedure, and any threats or incentives placed by the mediator are credible in the game-theoretic sense. Now we turn to the specific Cases 2, 3 and 4.

In Case 2, the mediator proposes a non-binding compromise s (after listening to both parties, i.e., 'learning' the bargaining problem U at hand). This makes s salient and externally approved. If both players accept s, it is implemented. If the parties do not both accept the proposal and do not reach an alternative compromise, the non-cooperative outcome r is predicted. So, r once again operates as a reference point or 'threat', creating an incentive to accept the proposal s (as long as $s > r$).

In Case 4, the mediator not just proposes a non-binding compromise, but also underpins this proposal with the threat of forcing a 'bad' binding outcome r on the parties (typically including sanctions or fines) which takes effect in the eventuality that the players neither agree to s nor manage to reach an alternative compromise. This of course presupposes that players have contractually authorized the mediator to dictate a binding disagreement outcome (which players may plausibly do to facilitate a compromise). Once the mediator has announced r, players effectively face an exogenous disagreement outcome. While classical bargaining theory can be used to model bargaining *given* the mediator's announced r, we also address how r is determined.

Case 3 gives more responsibility to the parties: the mediator does not propose a compromise to the parties but mediates between them to help them find a compromise by themselves. Just as in Case 4, to create an incentive to compromise, the mediator imposes a binding outcome r (typically including sanctions or fines) that

[5]The United Kingdom legislation strongly encourages, if not *de facto* forces, parties to engage in an ADR process prior to meeting before court (since the 2004 judgment in the *Halsey* landmark case). For general introductions to ADR, see for instance Lynch (2001) and Blake et al. (2011).

[6]The difference between a court trial and arbitration is that the former is instituted by the state, whereas the latter is based on a contractual agreement between both parties to submit to an arbitration procedure.

takes effect if no compromise is reached.

We ultimately leave it to the reader which interpretation to prefer and which applications to focus on. As with bargaining theory in general, the theory with endogenous disagreement captures its intended applications only in a stylized and simplified way. For instance, the model abstracts away certain goals of ADR, such as the goal of inducing a change and ideally a convergence of the parties' preferences. We hope that the connection to ADR will motivate future research and generalizations.

2.3 Bargaining problems and the extended Kalai-Smorodinsky solution

In this chapter we focus on a particular solution, which extends the classical Kalai-Smorodinsky bargaining solution (Raiffa, 1953; Kalai and Smorodinsky, 1975). For a bargaining problem $U \in \mathcal{U}$, the *Pareto optimal set* is the set

$$P(U) := \{x \in U \mid \text{for all } y \in U,\ y \geqslant x \text{ implies } y = x\}$$

and the *anti-Pareto optimal set* is the set

$$AP(U) := \{x \in U \mid \text{for all } y \in U,\ y \leqslant x \text{ implies } y = x\}.$$

The classical Kalai-Smorodinsky bargaining solution assigns to each classical bargaining problem (U, d) the unique point $KS(U, d) \in P(U)$ on the straight line through d and the *utopia point*

$$u(U, d) = \left(\max_{x \in U,\ x \geqslant d} x_1,\ \max_{x \in U,\ x \geqslant d} x_2 \right).$$

The *extended Kalai-Smorodinsky solution* is the correspondence $k : \mathcal{U} \to \mathbb{R}^2 \times \mathbb{R}^2$ defined by

$$(s, r) \in k(U) \Leftrightarrow s = KS(U, r),\ r = -KS(-U, -s) \text{ and } s \neq r$$

for all $U \in \mathcal{U}$ and $s, r \in U$. Thus, $(s, r) \in k(U)$ exactly if the following three conditions are satisfied: (i) $s \neq r$; (ii) s is the classical Kalai-Smorodinsky outcome when r is viewed as the disagreement outcome; and (iii) r results similarly from s when we reverse the problem or, equivalently, r is the unique point in $AP(U)$ on the straight line through s and the *anti-utopia point*

$$a(U, s) := \left(\min_{x \in U,\ x \leqslant s} x_1,\ \min_{x \in U,\ x \leqslant s} x_2 \right).$$

See Figure 2.3.1 for an illustration.

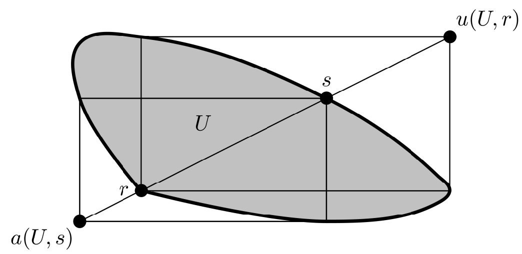


Figure 2.3.1: An illustration of the extended Kalai-Smorodinsky solution

Our first result is that the extended Kalai-Smorodinsky solution is non-empty valued. The proof is based on an elementary fixed point argument, slightly complicated by the fact that the Pareto and anti-Pareto optimal sets of a bargaining problem U may have one or both endpoints in common. Clearly, in that case, by definition of k – in particular the condition $s \neq r$ – such an endpoint cannot be the solution outcome.

Theorem 2.3.1 *$k(U) \neq \emptyset$ for all $U \in \mathcal{U}$.*

We note that k does not have to assign a *unique* pair of outcomes to a bargaining problem. For instance, let U be the convex hull of the points $(6, 0)$, $(8, 0)$, $(0, 6)$, and $(0, 8)$. Then it is not difficult to check that

$$k(U) = \{((s_1, s_2), (r_1, r_2)) \mid 2 \leqslant s_1 \leqslant 6,\ r_1 = s_1 - 1,\ s_1 + s_2 = 8,\ r_1 + r_2 = 6\}.$$

In this example the Pareto optimal and anti-Pareto optimal sets are parallel line segments. In fact, a sufficient but not necessary condition for k to assign a unique pair of outcomes to a problem U is that $P(U)$ and $AP(U)$ do *not* contain parallel line segments. Theorem 2.3.2 below provides an exact description of the class of all bargaining problems on which k is unique. We first introduce some additional terminology.

For $x \neq y$ and $x' \neq y'$ the line segments $[x, y]$ and $[x', y']$ are *parallel* if the straight lines ℓ and ℓ' containing these line segments are parallel. In that case, the *vertical distance* between $[x, y]$ and $[x', y']$ is the number $v = |z_2 - z'_2|$ for (any) $z \in \ell$ and $z' \in \ell'$ with $z_1 = z'_1$; v is infinite if ℓ and ℓ' are vertical. Similarly, the *horizontal distance* between $[x, y]$ and $[x', y']$ is the number $h = |z_1 - z'_1|$ for (any) $z \in \ell$ and $z' \in \ell'$ with $z_2 = z'_2$; h is infinite if ℓ and ℓ' are horizontal.

Now let $\mathcal{D}_k$ denote the set of bargaining problems U with $|k(U)| = 1$.

Theorem 2.3.2 *Let $U \in \mathcal{U}$. Then $U \in \mathcal{D}_k$ if and only if there are* no *parallel line segments $[\bar{x}, \underline{x}] \subseteq AP(U)]$ and $[\bar{y}, \underline{y}] \subseteq P(U)$ with $\bar{x}_1 < \underline{x}_1$ and $\bar{y}_1 < \underline{y}_1$ and such that the vertical distance v and horizontal distance h between these line segments satisfy the following conditions:*

(i) $\frac{1}{2}v = \bar{y}_2 - \bar{x}_2 = \underline{y}_2 - \underline{x}_2$,

(ii) *the lengths*[7] *of $[\bar{x}, \underline{x}]$ and $[\bar{y}, \underline{y}]$ both exceed $\sqrt{h^2 + v^2}$.*

See Figure 2.5.2 in the Appendix for an illustration. The theorem implies that we do not lose much generality if we restrict attention to domains of bargaining problems within $\mathcal{D}_k$. We conclude this section with a remark, listing some domains on which k is single-valued.

Remark 2.3.1 Theorem 2.3.2 implies that the extended Kalai-Smorodinsky solution k is single-valued on each of the following domains:

(a) $\{U \in \mathcal{U} \mid U \text{ is strictly convex}\}$.

(b) $\{U \in \mathcal{U} \mid AP(U) \text{ or } P(U) \text{ contains no line segment}\}$.

(c) $\{U \in \mathcal{U} \mid \text{no line segments } S \subseteq AP(U) \text{ and } S' \subseteq P(U) \text{ are parallel}\}$.

Clearly, the domain in (a) is a subset of the domain in (b), which in turn is a subset of the domain in (c).

2.4 Two axiomatic characterizations of the extended Kalai-Smorodinsky solution

We give two axiomatic characterizations of the extended Kalai-Smorodinsky solution k on domains on which k is single-valued in this section. In each characterization all axioms except one are basic and shared with the extended Nash solution. In the first characterization the additional axiom is an informational constraint (Independence of Non-Utopia Information), while in the second it is a monotonicity property analogous to such properties used in characterizations of the classical Kalai-Smorodinsky solution.

We formulate our conditions for a solution f defined on a domain $\mathcal{D} \subseteq \mathcal{U}$ with $|f(U)| = 1$ for all $U \in \mathcal{D}$. Instead of $f(U) = \{(s, r)\}$ we write $f(U) = (s, r)$ and regard f as a function rather than a correspondence.

[7] The length of a line segment is the Euclidean distance between its endpoints.

A bargaining problem $U' \in \mathcal{U}$ is a *positive affine transformation* of a bargaining problem $U \in \mathcal{U}$ if there are $a \in \mathbb{R}^2_+$ and $b \in \mathbb{R}^2$ such that $U' = aU + b$. A bargaining problem $U \in \mathcal{U}$ is *symmetric* if $(x_1, x_2) \in U \Leftrightarrow (x_2, x_1) \in U$ for all $x \in \mathbb{R}^2$.

The first condition is an extended version of the usual Pareto optimality condition.

Extended Pareto Optimality (EPO): For each $U \in \mathcal{D}$, $f(U) \in P(U) \times AP(U)$.

In particular from a normative view point it is natural to require Pareto optimality of the compromise outcome. Requiring anti-Pareto optimality of the disagreement outcome reflects that we wish this outcome to be as severe as possible in order to induce acceptance of the compromise outcome.[8]

The following two conditions are standard in classical axiomatic bargaining theory. They have similar justifications in the present model.

Symmetry (SYM): For each symmetric $U \in \mathcal{D}$, if $f(U) = (s, r)$ then $s_1 = s_2$ and $r_1 = r_2$.

Scale Invariance (SI): For all $U \in \mathcal{D}$ and $a \in \mathbb{R}^2_+$, $b \in \mathbb{R}^2$ with $aU + b \in \mathcal{D}$, if $f(U) = (s, r)$ then $f(aU + b) = (as + b, ar + b)$.

We now turn to axioms used in only one of our two characterizations. The first characterization is based on an informational restriction which extends and modifies similar conditions used in characterizations of the classical Kalai-Smorodinsky solution.

Independence of Non-Utopia Information (INU): For all $U, V \in \mathcal{D}$, if $f(V) = (s, r) \in P(U) \times AP(U)$, $u(U, r) = u(V, r)$ and $a(U, s) = a(V, s)$, then $f(U) = (s, r)$.

This condition says that if $f(V) = (s, r)$ and we consider a problem U such that s and r are Pareto and anti-Pareto optimal in U and also the associated utopia and anti-utopia points do not change, then the solution does not change: $f(U) = (s, r)$ as well.

Our second characterization replaces INU by three other axioms, each of which seems normatively defensible and extends classical axioms. The first of these axioms requires that the compromise outcome weakly Pareto dominates the disagreement outcome, i.e., that the disagreement outcome is a threat to both players.

Pareto Dominance (PD): For every $U \in \mathcal{D}$, if $f(U) = (s, r)$ then $s \geqslant r$.

The next condition requires that the outcome for any bargaining problem U be unchanged if one removes possible alternatives x from U that are extreme in the sense of giving some individual even less utility than under the original disagreement

[8] Disagreement poses a threat to the players only if $s > r$ (where $f(U) = (s, r)$), as a referee rightly noticed. In our two characterization results, EPO could be weakened by restricting the anti-Pareto optimality requirement on r to those cases in which $s > r$.

outcome while giving the other individual even more utility than under the original compromise outcome.

Independence of Extreme Alternatives (IEA): For all $U, U' \in \mathcal{D}$, writing $(s, r) = f(U)$, if $U' \subseteq U$ and for every $x \in U \backslash U'$ there is an agent i such that $x_i < r_i$ and $x_j > s_j$ for $j \neq i$, then $f(U') = f(U)$.

This condition is a weak version of the condition of IIA (Independence of Irrelevant Alternatives, extended to endogenous disagreement), which underlies the extended Nash bargaining solution. IEA relaxes IIA by restricting it to the case that two sets U and U' differ only in 'extreme' alternatives.

The final condition is a variant of classical monotonicity conditions. It is well-known from classical bargaining theory that plausible bargaining solutions usually satisfy *some but not any* kind of monotonicity property. Our monotonicity condition requires that if additional alternatives become available, then, at least under certain extra conditions, the compromise outcome improves weakly and the disagreement outcome worsens weakly for each player. Roughly speaking, the justification is that additional possibilities should give room for better compromise outcomes but also worse disagreement outcomes. In order to formulate the axiom we define, for a bargaining problem $U \in \mathcal{D}$, the *global utopia point* and the *global anti-utopia point* by

$$u(U) = \left(\max_{x \in U} x_1, \max_{x \in U} x_2\right), \quad a(U) = \left(\min_{x \in U} x_1, \min_{x \in U} x_2\right).$$

Restricted Monotonicity (RM): For all $U, U' \in \mathcal{D}$, writing $(s, r) = f(U)$ and $(s', r') = f(U')$, if $U \subseteq U'$, $u(U') = u(U, r)$, and $a(U') = a(U, s)$, then $s' \geq s$ and $r' \leq r$.

Clearly, the conditions on the utopia and global utopia points and the anti-utopia and global anti-utopia points considerably restrict this monotonicity condition.[9]

The domain $\mathcal{D}$ is *closed under truncation* if whenever it contains U then it also contains every bargaining problem of the form $\{x \in U \mid \alpha \leq x_i \leq \beta\}$ for some $i \in \{1, 2\}$ and $\alpha, \beta \in \mathbb{R}$ with $a_i(U) \leqslant \alpha < \beta \leqslant u_i(U)$. The domain $\mathcal{D}$ is *minimally rich* if it is closed under truncation and contains all polytopes in $\mathcal{D}_k$.[10] For instance, the whole domain $\mathcal{D}_k$ and the (small) domain of all polytopes in $\mathcal{D}_k$ are both minimally rich by Theorem 2.3.2.

Theorem 2.4.1 *Let $\mathcal{D} \subseteq \mathcal{D}_k$ be minimally rich and let $f : \mathcal{D} \to \mathbb{R}^2 \times \mathbb{R}^2$ be a solution satisfying $|f(U)| = 1$ for all $U \in \mathcal{D}$. Then the following statements are equivalent:*

(a) *f is the extended Kalai-Smorodinsky solution on $\mathcal{D}$.*

[9] Note that the antecedent in RM implies that $u(U) = u(U, r)$ and $a(U) = a(U, s)$.

[10] A polytope is the convex hull of finitely many points in $\mathbb{R}^2$.

(b) *f satisfies EPO, SYM, SI, and INU.*

(c) *f satisfies EPO, SYM, SI, PD, IEA, and RM.*

The characterizations in Theorem 2.4.1 are tight. We show this by means of examples of solutions defined on a minimally rich domain $\mathcal{D} \subseteq \mathcal{D}_k$. Proofs are left to the reader. We start with demonstrating tightness of the six axioms in characterization (c).

(1) For each $U \in \mathcal{D}$, write $(s, r) := k(U)$ and let $f^1(U) := (t, r)$, where t is the point in $[r, s]$ which is closest to s subject to U containing at least one of the points $(t_1, a_2(U, s))$ and $(a_1(U, s), t_2)$ (note that possibly $t = s$). Then f^1 satisfies SYM, SI, PD, IEA, and RM, but not EPO.

(2) Define the solution f^2 in the same way as k but now based on a non-symmetric version of the KS-solution (cf. Peters and Tijs, 1985). Such a solution satisfies EPO, SI, PD, IEA, and RM, but not SYM.

(3) Let T be the convex hull of $(0, 0)$, $(4, 0)$, and $(0, 2)$. Define the solution f^3 as follows. For all $U \in \mathcal{D}$ with $T \subseteq U$, $a(U) = (0, 0)$, and $u(U) = (4, 2)$, define $f^3(U)$ as $(s, (0, 0))$ where s is the point of intersection of $P(U)$ with the line segment $[(3, \frac{1}{2}), (4, 2)]$. Otherwise, define $f^3(U) = k(U)$. Then f^3 satisfies EPO, SYM, PD, IEA, and RM, but not SI.

(4) For each $U \in \mathcal{U}$, let $(s(U), r(U)) := k(U)$, let $\hat{s}(U)$ resp. $\hat{r}(U)$ be the element of U with first coordinate $r_1(U)$ resp. $s_1(U)$ and with maximal resp. minimal second coordinate, denote the set of non-extreme outcomes relative to k by $\overline{U} = \{x \in U \mid x_i \leqslant s_i(U) \text{ or } x_j \geqslant r_i(U) \text{ for all distinct } i, j\}$, and call $U \in \mathcal{U}$ *essentially symmetric* if some positive affine transformation of $\overline{U}$ is symmetric. For all $U \in \mathcal{D}$, define $f^4(U)$ as $(\hat{s}(U), \hat{r}(U))$ if [$\hat{s}(U) \in P(U)$, $\hat{r}(U) \in AP(U)$ and U is not essentially symmetric], and as $k(U)$ otherwise. Then f^4 satisfies EPO, SYM, SI, IEA, and RM, but not PD.

(5) For all $U \in \mathcal{D}$, define $f^5(U)$ as (s, r) where s [r] is the intersection of $P(U)$ [$AP(U)$] with the line segment joining the global utopia point and the global anti-utopia point of U. Then f^5 satisfies EPO, SYM, SI, PD, and RM, but not IEA.

(6) Let T be the convex hull of $(0, 0)$, $(4, 0)$, $(2, 1)$, and $(0, 1)$. For all $U \in \mathcal{D}$, define $f^6(U)$ as (s, b) if $U = aT + b$ for some $a \in \mathbb{R}^2_+$, $b \in \mathbb{R}^2$, where s is the Nash bargaining solution of (U, b), and as $k(U)$ otherwise. Then f^6 satisfies EPO, SYM, SI, PD, and IEA, but not RM.

Next, we show that the axioms in characterization (b) are tight.

(7) The solution f^1 satisfies SYM, SI, and INU, but not EPO.

(8) The solution f^2 satisfies EPO, SI, and INU, but not SYM.

(9) Let T be as in (3) and define the solution f^7 by $f^7(T) = ((3, \frac{1}{2}), (0,0))$, and by $f^7(U) = k(U)$ for all $U \in \mathcal{D}_k$ with $U \neq T$. Then f^7 satisfies EPO, SYM, INU, but not SI.

(10) The solutions f^4, f^5, and f^6 all satisfy EPO, SYM, and SI, but not INU.

We conclude with a few remarks.

Remark 2.4.1 A partial characterization of the extended Kalai-Smorodinsky solution on the whole domain $\mathcal{U}$ is provided in Valkengoed (2006)[11], at the expense of rather technical conditions.

Remark 2.4.2 Variations on the characterization of k can be obtained by imposing different conditions of 'minimal richness'. For instance, Theorem 2.4.1 would still hold – with some modifications of the proof – on some subdomains of $\mathcal{D}_k$ that contain all strictly convex bargaining problems.

2.5 Appendix: proofs

PROOF OF THEOREM 2.3.1. Let $U \in \mathcal{U}$. Then $AP(U)$ is the graph of a strictly decreasing convex function g on an interval $[\alpha, \beta]$ with $(\alpha, g(\alpha))$ and $(\beta, g(\beta))$ the points of $AP(U)$ with minimal and maximal first coordinates, respectively. If $\alpha = \beta$ (so that $AP(U)$ consists of a unique outcome) then $\{(KS(U, (\alpha, g(\alpha))), (\alpha, g(\alpha)))\} = k(U)$ and we are done. From now on we assume $\alpha < \beta$. Define the function $\varphi : [\alpha, \beta] \to [\alpha, \beta]$ by $\varphi(\gamma) = -KS_1(-U, -KS(U, (\gamma, g(\gamma))))$. Observe that if $\varphi(\gamma^*) = \gamma^*$ for some $\gamma^* \in [\alpha, \beta]$ and $KS(U, (\gamma^*, g(\gamma^*))) \neq (\gamma^*, g(\gamma^*))$ then $(KS(U, (\gamma^*, g(\gamma^*))), (\gamma^*, g(\gamma^*))) \in k(U)$.

Of course, $\varphi(\alpha) \geqslant \alpha$ and $\varphi(\beta) \leqslant \beta$. Suppose that $(\alpha, g(\alpha)) \in P(U)$. Then $\varphi(\alpha) = \alpha$, but $(KS(U, (\alpha, g(\alpha))), (\alpha, g(\alpha))) \notin k(U)$ since $KS(U, (\alpha, g(\alpha))) = (\alpha, g(\alpha))$. Below, however, we will prove:

$$\text{There is an } \varepsilon_1 > 0 \text{ with } \varphi(\gamma) > \gamma \text{ for all } \gamma \in (\alpha, \alpha + \varepsilon_1]. \tag{2.5.1}$$

Similarly, if $(\beta, g(\beta)) \in P(U)$ we have:

$$\text{There is an } \varepsilon_2 > 0 \text{ with } \varphi(\gamma) < \gamma \text{ for all } \gamma \in [\beta - \varepsilon_2, \beta). \tag{2.5.2}$$

[11] Master thesis, supervised by Prof. dr. Hans Peters.

Clearly, we can then take ε_1 and ε_2 in (2.5.1) and (2.5.2) such that $\alpha+\varepsilon_1 < \beta-\varepsilon_2$. Now define the interval $[\alpha', \beta']$ by $\alpha' = \alpha$ if $(\alpha, g(\alpha)) \notin P(U)$ and $\alpha' = \alpha + \varepsilon_1$ if $(\alpha, g(\alpha)) \in P(U)$, and $\beta' = \beta$ if $(\beta, g(\beta)) \notin P(U)$ and $\beta' = \beta - \varepsilon_2$ if $(\beta, g(\beta)) \in P(U)$. Then, since φ is continuous, the intermediate value theorem implies that in all cases there is a point $\gamma^* \in [\alpha', \beta']$ with $\varphi(\gamma^*) = \gamma^*$ and $KS(U, (\gamma^*, g(\gamma^*))) \neq (\gamma^*, g(\gamma^*))$ and, thus, $k(U) \neq \emptyset$.

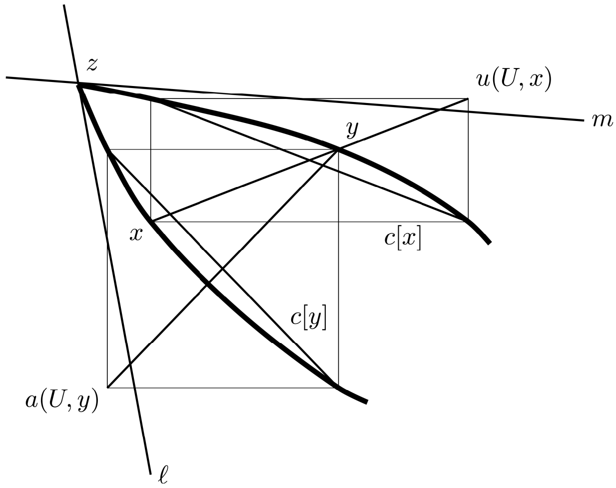


Figure 2.5.1: Illustrating the proof of (2.5.1)

We are left to prove (2.5.1) and (2.5.2). We only show (2.5.1), the proof of (2.5.2) is analogous. So suppose $z := (\alpha, g(\alpha)) \in P(U)$. See Figure 2.5.1 for an illustration of the remainder of the proof.

Let m and ℓ be the supporting lines of U at z as drawn in Figure 2.5.1. (That is, m is the limit of supporting lines at $P(U)$ and ℓ is the limit of supporting lines at $AP(U)$.) Let μ be the absolute value of the slope of m and let λ be the absolute value of the slope of ℓ. Then $\lambda > \mu$.

For $x \in AP(U) \setminus P(U)$ let $\sigma(x)$ denote the slope of the straight line through x and $u(U, x)$. Let $c[x]$ denote the line segment with endpoints $(x_1, u_2(U, x))$ and $(u_1(U, x), x_2)$. Then $\sigma(x)$ is equal to the absolute value of the slope of $c[x]$. Therefore, $\sigma(x)$ converges to μ if $x \in AP(U)$ converges to z.

For $y \in P(U) \setminus AP(U)$ let $\tau(y)$ denote the slope of the straight line through y and $a(U, y)$. Let $c[y]$ denote the line segment with endpoints $(y_1, a_2(U, y))$ and $(a_1(U, y), y_2)$. Then $\tau(y)$ is equal to the absolute value of the slope of $c[y]$. Therefore, $\tau(y)$ converges to λ if $y \in P(U)$ converges to z.

We conclude that $\tau(y) > \sigma(x)$ for $y \in P(U)$ and $x \in AP(U)$ close to z. This implies the existence of an ε_1 as in (2.5.1). $\square$

Proof of Theorem 2.3.2 For the only-if part, suppose that $[\bar{x}, \underline{x}] \subseteq AP(U)]$ and

$[\bar{y}, \underline{y}] \subseteq P(U)$ are as in the Theorem. Let $\bar{r} \in [\bar{x}, \underline{x}]$ with $\bar{r}_1 = \bar{y}_1$ and $\underline{r} \in [\bar{x}, \underline{x}]$ with $\underline{r}_2 = \underline{y}_2$. For each $r \in [\bar{r}, \underline{r}]$ let $s(r) \in [\bar{y}, \underline{y}]$ with $s(r)_2 - r_2 = v/2$. Then it is straightforward to check that $(s(r), r) \in k(U)$ for each $r \in [\bar{r}, \underline{r}]$. Thus, $U \notin \mathcal{D}_k$. See Figure 2.5.2 for an illustration.

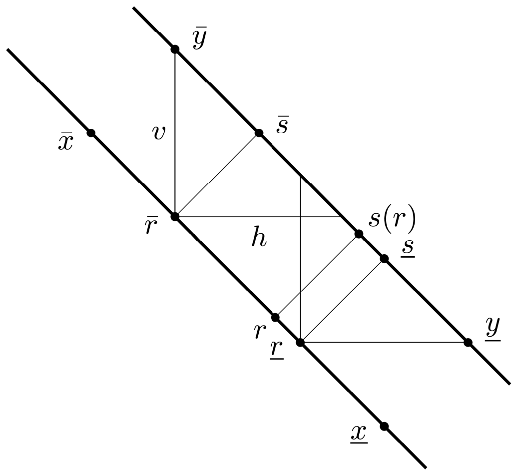


Figure 2.5.2: Illustrating the proof of Theorem 2.3.2

We now prove the if-part. Assume $U \notin \mathcal{D}_k$, i.e. $|k(U)| > 1$. We will construct $[\bar{x}, \underline{x}] \subseteq AP(U)]$ and $[\bar{y}, \underline{y}] \subseteq P(U)$ as in the theorem.

For any $x \in AP(U)$ let $\sigma(x)$ denote the slope of the straight line through x and $u(U, x)$ (as in the proof of Theorem 2.3.1). Since $\sigma(x)$ is equal to the absolute value of the slope of the line segment $c[x]$ connecting the points $(x_1, u_2(U, x))$ and $(u_1(U, x), x_2)$, and the absolute values of these slopes weakly increase if x_1 increases – the line segments $c[x]$ are chords of the weakly decreasing concave function the graph of which contains the Pareto optimal set of U – we have that $\sigma(x)$ weakly increases if x_1 increases. $(*)$

Similarly, for any $y \in P(U)$ let $\tau(y)$ denote the slope of the straight line through y and $a(U, y)$ (again as in the proof of Theorem 2.3.1). Then by an analogous argument $\tau(y)$ weakly increases if y_1 decreases. $(**)$

Let $(\bar{s}, \bar{r})$ and $(\underline{s}, \underline{r})$ be the elements of $k(U)$ with maximal and minimal second coordinates, respectively. By definition of k we have $\tau(s) = \sigma(r)$ for all $(s, r) \in k(U)$. Therefore, by $(*)$ and $(**)$ we must have $\sigma(x) = \tau(y)$ for all $x \in AP(U)$ with $\bar{r}_1 \leqslant x_1 \leqslant \underline{r}_1$ and all $y \in P(U)$ with $\bar{s}_1 \leqslant y_1 \leqslant \underline{s}_1$. In particular, $\sigma(x)$ is constant for $\bar{r}_1 \leqslant x_1 \leqslant \underline{r}_1$, which implies that the line segments $c[x]$ for $x \in [\bar{r}, \underline{r}]$ are parallel; but this means that they must be on the same straight line m through $\bar{s}$ and $\underline{s}$. Let $\bar{y}$ be the upper endpoint of $c[\bar{r}]$ and let $\underline{y}$ be the lower endpoint of $c[\underline{r}]$. Then $[\bar{y}, \underline{y}] \subseteq P(U)$, $\bar{y}_1 = \bar{r}_1$ and $\underline{y}_2 = \underline{r}_2$. See, again, Figure 2.5.2 for an illustration.

Similarly, let ℓ be the straight line through $\bar{r}$ and $\underline{r}$, then $[\bar{x}, \underline{x}] \subseteq AP(U)$, where $\bar{x}$ is the point of ℓ with $\bar{x}_2 = \bar{s}_2$ and $\underline{x}$ is the point of m with $\underline{x}_1 = \underline{s}_1$. Now it is straightforward to check that $[\bar{x}, \underline{x}]$ and $[\bar{y}, \underline{y}]$ satisfy the conditions in the theorem. $\square$

Proof of Theorem 2.4.1. (1) We first prove that (a) implies (b) and (c). We leave verification of EPO, SYM, SI, INU, PD, and IEA of k on $\mathcal{D}$ to the reader. To show RM, consider $U, U' \in \mathcal{D}$ satisfying the antecedent of RM, i.e., $U \subseteq U'$, $u(U') = u(U, r)$, and $a(U') = a(U, s)$, where $(s, r) := k(U)$. Since $r < s$ and $U \subseteq U'$, there are unique points $\bar{r} < \bar{s}$ in the intersection of the line through r and s with the boundary of U'. We show that $(\bar{s}, \bar{r}) = k(U')$, which completes the proof of RM since, clearly, $\bar{r} \leq r$ and $\bar{s} \geq s$. Note that $u(U', \bar{r}) \leq u(U') = u(U, r) \leq u(U', \bar{r})$, so that $u(U', \bar{r}) = u(U, r)$, whence $KS(U', \bar{r}) = \bar{s}$. Similarly, $a(U', \bar{s}) \geq a(U') = a(U, s) \geq a(U', \bar{s})$, so that $a(U', \bar{s}) = a(U, s)$, whence $KS(-U', -\bar{s}) = -\bar{r}$. It follows that $k(U') = (\bar{s}, \bar{r})$.

(2) We now prove that (b) implies (a). Suppose f satisfies the four conditions in (b) and let $U \in \mathcal{D}$. We have to prove that $f(U) = k(U)$. Let $k(U) = (s, r) \in P(U) \times AP(U)$. Then $s > r$ (this follows from the requirement that there must be $x, y \in U$ with $x > y$). Let V be the convex hull of the six points s, r, $(s_1, a_2(U, s))$, $(a_1(U, s), s_2)$, $(u_1(U, r), r_2)$, and $(r_1, u_2(U, r))$. We will prove that $V \in \mathcal{D}$ and $f(V) = (s, r)$. This will conclude the proof of (b) $\Rightarrow$ (a), since by INU, $f(V) = (s, r)$ implies $f(U) = (s, r)$ and, thus, $f(U) = k(U)$.

Consider the positive affine transformation

$$(x_1, x_2) \mapsto (\varphi_1(x_1), \varphi_2(x_2)) := \left(\frac{x_1 - r_1}{s_1 - r_1}, \frac{x_2 - r_2}{s_2 - r_2} \right)$$

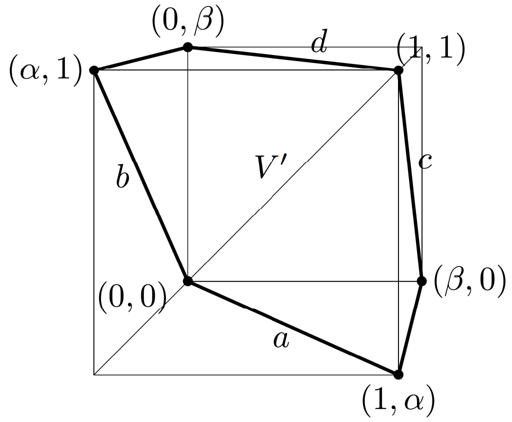
which maps r to $(0, 0)$, s to $(1, 1)$, and V to some set V'. Then V' is the convex hull of the set

$$\{(0,0), (1,1), (1, \frac{a_2(U,s) - r_2}{s_2 - r_2}), (\frac{a_1(U,s) - r_1}{s_1 - r_1}, 1), (\frac{u_1(U,r) - r_1}{s_1 - r_1}, 0), (0, \frac{u_2(U,r) - r_2}{s_2 - r_2})\} .$$

Note that

$$\frac{a_2(U,s) - r_2}{s_2 - r_2} = \frac{a_1(U,s) - r_1}{s_1 - r_1} \text{ and } \frac{u_2(U,r) - r_2}{s_2 - r_2} = \frac{u_1(U,r) - r_1}{s_1 - r_1} .$$

Thus, V' is a symmetric polytope, and it is sufficient to prove that $V' \in \mathcal{D}_k$: for this implies $V \in \mathcal{D}$ by minimal richness of $\mathcal{D}$; and by SYM and EPO, we have $f(V') = ((1, 1), (0, 0))$ and thus, by SI, $f(V) = (s, r)$.



The letters a, b, c, d denote line segments, $\alpha = \varphi_1(a_1(U, s)) = \varphi_2(a_2(U, s))$, and $\beta = \varphi_1(u_1(U, r)) = \varphi_2(u_2(U, r))$

Figure 2.5.3: Illustrating the proof of part (2) of Theorem 2.4.1.

We are left to prove that $V' \in \mathcal{D}_k$, i.e., that $|k(V')| = 1$. Consider Figure 2.5.3 with notations as there. For $k(V')$ to be non-unique there are, in view of Theorem 2.3.2, two possible cases to examine: (1) a is parallel to d and (2) a is parallel to c. (The cases involving b are analogous.) In case (1) we must have $\beta = 1 - \alpha > 1$. Denote the vertical and horizontal distances between a and d by v and h, then the length of a is equal to $\sqrt{1+\alpha^2}$ whereas $\sqrt{v^2+h^2} > \sqrt{\beta^2+\beta^2} > \sqrt{1+\alpha^2}$, so that a does not satisfy condition (ii) in Theorem 2.3.2. In case (2) we must have $\beta = 2$ and $\alpha = -1$. In particular, $AP(V')$ is the line segment $[(-1, 1), (1, -1)]$ and $P(V')$ is the line segment $[(0, 2), (2, 0)]$, so that again condition (ii) in Theorem 2.3.2 is violated.

(3) We finally prove that (c) implies (a). Suppose f satisfies the six conditions in (b) and let $U \in \mathcal{D}$. Let $(s, r) := f(U)$. We proceed in several steps.

Claim 1. $s > r$.

To prove this claim, assume the contrary. As $s \geqslant r$ by PD and $s \neq r$ by definition of a bargaining solution, we may assume $s_1 = r_1$ and $s_2 > r_2$ (the other case is analogous).

Consider first the truncated set $\widehat{U} = \{x \in U \mid x_2 \leqslant s_2\}$, which is in $\mathcal{D}$ by minimal richness. Note that for each $x \in U \setminus \widehat{U}$ we have $x_2 > s_2$ and hence $x_1 < s_1 = r_1$ as $s \in P(U)$; so by IEA, $f(\widehat{U}) = f(U)$. Next consider the set $T = \{x \in \widehat{U} \mid x_2 \geqslant r_2\}$, which is again in $\mathcal{D}$ by minimal richness. Next, note that for each $x \in \widehat{U} \setminus T$ we have $x_2 < r_2$ and hence $x_1 > r_1 = s_1$ as $r \in AP(\widehat{U})$; so again by IEA, $f(T) = f(\widehat{U})$. Altogether we have $f(T) = f(U)$ where $T = \{x \in U \mid r_2 \leqslant x_2 \leqslant s_2\}$. Moreover, by construction of T, we have $u(T, r) = u(T)$ and $a(T, s) = a(T)$.

As $T \supsetneq [s, r]$, we have $u(T) \neq s$ or $a(T) \neq r$. Suppose $u(T) \neq s$ (the proof is analogous if $a(T) \neq r$). Then, since $s \in P(T)$, $u(T) \notin T$. Choose $\alpha, \beta \in \mathbb{R}$ with $a_2(T) < \alpha < u_2(T)$ and $a_1(T) < \beta < u_1(T)$ such that the set

$$T' = \operatorname{conv}\{a(T), (a_1(T), u_2(T)), (u_1(T), a_2(T)), (u_1(T), \alpha), (\beta, u_2(T))\}$$

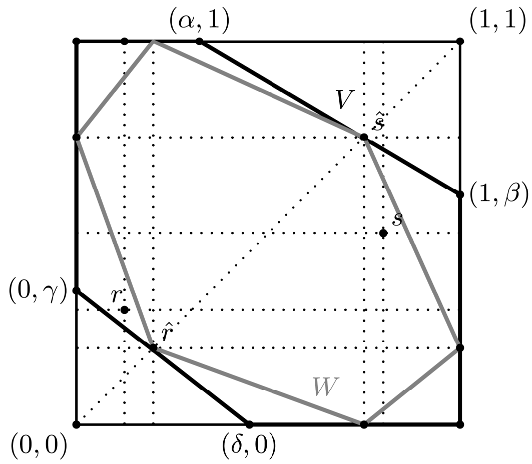
is a positive affine transformation of a symmetric polytope and such that $T \subseteq T'$. Then $T' \in \mathcal{D}$ as $\mathcal{D}$ is minimally rich and $T' \in \mathcal{D}_k$ by Theorem 2.3.2. Let $(s', r') = f(T')$. As T' is symmetric up to a positive affine transformation, SI and SYM imply that $r', s' \in T' \cap [a(T'), u(T')]$. So, as $u(T') \notin T'$, we have $s' < u(T') = u(T)$, whence in particular $s'_2 < u_2(T)$. On the other hand, since $T \subseteq T'$, $u(T') = u(T) = u(T, r)$ and $a(T') = a(T) = a(T, s)$, we have by RM that $s' \geqslant s$, so that $s'_2 = u_2(T)$. This contradiction completes the proof of Claim 1.

Claim 2. Let $U' = \{x \in U \mid a(U, s) \leqslant x \leqslant u(U, r)\}$. Then $U' \in \mathcal{D}$ and $f(U') = (s, r)$.

To prove this, first observe that, since $s > r$ by Claim 1, U' arises from U by a double truncation. Hence, $U' \in \mathcal{D}$. We next prove that all outcomes in $U \setminus U'$ are extreme alternatives in the sense of IEA. Suppose that $x \in U$ with $x_1 < r_1$. Suppose $x_2 \leqslant s_2$. Since $r < s$, we have $x < s$, hence $x \geqslant a(U, s)$. Also, $x < u(U, r)$ since $s \leqslant u(U, r)$. Thus, $x \in U'$. Hence, if $x \in U \setminus U'$ then $x_1 < r_1$ implies $x_2 > s_2$. Similarly, $x \in U \setminus U'$ and $x_2 < r_2$ imply $x_1 > s_1$. Suppose now $x \in U$ and $x \geqslant r$. Then $x \leqslant u(U, r)$ and since $r < s$, whence $r \geqslant a(U, s)$, we have $x \geqslant a(U, s)$, so that $x \in U'$. Altogether we have proved that the antecedent of IEA holds for $U' \subseteq U$, so that $f(U') = (s, r)$.

In view of Claim 2 and the definition of k it is sufficient to prove that $f(U') = k(U')$. In view of SI of f and k we may assume that $a(U', s) = (0, 0)$ and $u(U', r) = (1, 1)$. Denote $L = [(0, 0), (1, 1)]$ and $U^0 = \{x \in \mathbb{R}^2 \mid (0, 0) \leqslant x \leqslant (1, 1)\}$. If $r, s \in L$ then clearly $k(U') = (s, r) = f(U')$ and we are done. Otherwise, without loss of generality $s \notin L$. We proceed by choosing $\hat{s}, \hat{r} \in L$ as follows. If $r \notin L$ then choose $\hat{s}, \hat{r}$ such that: (i) $\hat{s} \not\geqslant s$, $\hat{r} \not\leqslant r$; (ii) there is a line ℓ through $\hat{s}$ intersecting the boundary of U^0 at points $(\alpha, 1)$ and $(1, \beta)$ such that $\hat{r}_1 < \alpha < 1$, $\hat{r}_2 < \beta < 1$ and such that the set U' is weakly below ℓ; (iii) there is a line m, not parallel to ℓ, through $\hat{r}$ intersecting the boundary of U^0 at points $(0, \gamma)$ and $(\delta, 0)$ such that $0 < \gamma < \hat{s}_2$, $0 < \delta < \hat{s}_1$ and such that the set U' is weakly above m.[12] See Figure 2.5.4 for an illustration. If $r \in L$ then we still choose $\hat{s}$ as above but set $\hat{r} = (0, 0)$, and $\gamma = \delta = 0$. Let V be the polytope with vertices $(0, 1)$, $(1, 0)$, $(\alpha, 1)$, $(1, \beta)$, $(0, \gamma)$, and $(\delta, 0)$. Then $V \in \mathcal{D}_k$ by Theorem 2.3.2 and therefore $V \in \mathcal{D}$ since $\mathcal{D}$ is minimally rich.

[12] The first candidates for $\hat{s}$ and $\hat{r}$ are the points s' and r', where s' is the point of intersection of L and $P(U')$ and r' is the point of intersection of L and $AP(U')$; and for ℓ and m the lines through s' and r' supporting U'. If those lines happen to be parallel, or if one or more of the numbers α, β, γ, and δ do not satisfy the desired constraints, one can take $\hat{s} = s' + (\varepsilon, \varepsilon)$ and/or $\hat{r} = r' - (\varepsilon', \varepsilon')$, where $0 < \varepsilon, \varepsilon'$ are sufficiently small, and shift up and if necessary slightly rotate ℓ and/or m.



The black curve is the boundary of the set V and the gray curve is the boundary of the set W.

Figure 2.5.4: Illustrating the proof of part (3) of Theorem 2.4.1.

Claim 3. $f(V) \in \{x \in V : x \geq s\} \times \{x \in V : x \leq r\}$.
The claim follows from RM and Claim 2, noting that $U' \subseteq V$ and that $u(V) = u(U', r)$ $(= (1, 1))$ and $a(V) = a(U', s)$ $(= (0, 0))$.

Let W be the convex hull of the points $\hat{s}$, $\hat{r}$, $(\hat{s}_1, 0)$, $(1, \hat{r}_2)$, $(0, \hat{s}_2)$, and $(\hat{r}_1, 1)$.

Claim 4. $W \in \mathcal{D}$ and $f(W) = (\hat{s}, \hat{r})$.
That $W \in \mathcal{D}$, in particular that $W \in \mathcal{D}_k$, follows by the same argument as used in the last part of (2) above. Since W is symmetric, EPO and SYM imply $f(W) = (\hat{s}, \hat{r})$.

Claim 5. $f(V) = (\hat{s}, \hat{r})$.
To prove this, we note that by construction of V and W we have $(1, 1) = u(V) = u(W, \hat{r})$ and $(0, 0) = a(V) = a(W, \hat{s})$. Since $W \subseteq V$, the claim follows from RM and Claim 4.

We can now complete the proof of part (3) and the theorem. Claim 5, Claim 3, and the definition of $\hat{s}$ and $\hat{r}$ imply that we must have $s, r \in L$ since otherwise we obtain a contradiction. But in that case we have $(s, r) = k(U') = f(U')$. □

References to Part I

Blake, S.H., Browne, J., Sime, S. (2011) A practical approach to alternative dispute resolution. *Oxford University Press*

Bozbay, I., Dietrich, F., Peters, H. (2012) Bargaining with endogenous disagreement: The extended KalaiSmorodinsky solution. *Games and Economic Behavior* 74, 407–417

Herrero, C. (1998) Endogenous reference points and the adjusted proportional solution for bargaining problems with claims. *Social Choice and Welfare* 15, 113–119

Kalai, A., Kalai, E. (2010) A cooperative value for Bayesian games. Discussion Paper #1512, CMS-EMS, Kellogg, Northwestern University

Kalai, E., Rosenthal, R.W. (1978) Arbitration of two-party disputes under ignorance. *International Journal of Game Theory* 7, 65–72

Kalai, E., Smorodinsky, M. (1975) Other solutions to Nash's bargaining problem. *Econometrica* 43, 513–518

Luce, D., Raiffa, H. (1957) Games and Decisions. Wiley, New York

Lynch, J. (2001) ADR and beyond: a systems approach to conflict management. *Negotiation Journal* 17, 207–216

Nash, J.F. (1950) The bargaining problem. *Econometrica* 28, 155–162

Peters, H., Tijs, S. (1985) Characterization of all individually monotonic bargaining solutions. *International Journal of Game Theory* 14, 219–228

Raiffa, H. (1953) Arbitration schemes for generalized two-person games. Contributions to the Theory of Games II. Ed. by Kuhn HW, Tucker AW. *Annals of Mathematics Studies* 28, 361–387. Princeton University Press, Princeton

Thomson, W. (1981) A class of solutions to the bargaining problem. *Journal of Economic Theory* 25, 431–442

Valkengoed, J. (2006) A monotone solution for bargaining problems with endogenous disagreement point. Master thesis, Maastricht University

Vartiainen, H. (2002) Bargaining without disagreement. Yrjö Jahnsson Foundation, Helsinki

Vartiainen, H. (2007) Collective choice with endogenous reference outcome. *Games and Economic Behavior* 58, 172–180

Part II

Judgment aggregation

Chapter 3

Judgment aggregation in search for the truth

3.1 Introduction

In the by now well-established theory of judgment aggregation, a group needs to form a 'yes' or 'no' judgment on different issues, based on the judgments of the group members on these issues. For instance, the jury in a court trial might need to form judgments on whether the defendant has broken the contract, and whether the contract is legally valid; the United Nations security council might need to form judgments on whether country X is threatened by a military coup, and whether the economy of country X is collapsing; and so on. Group judgments matter in practice. They may determine group action: in the court trial example, they may determine whether the defendant is convicted, and in the United Nations example they may determine whether a large-scale international intervention in country X will happen.

So far, nearly the entire judgment aggregation theory follows the classical social-choice theoretic approach of aiming to find out how – and whether – group judgments can reflect the individuals' judgments in a procedurally fair manner, where 'fair' is spelled out in terms of axiomatic conditions on the aggregation rule (such as the anonymity condition or the Pareto-type condition of respecting unanimous judgments). The recent *Symposium on Judgment Aggregation* in Journal of Economic Theory (C. List and B. Polak eds., 2010, vol. 145(2)) illustrates well this social-choice theoretic approach, as well as the state of the art of the theory, which we review below. This approach is certainly important in many contexts. It is nonetheless surprising that little attention is given to a different, 'epistemic' approach of aiming to *track the truth*, i.e., reach true group judgments. The theory does not model the private information underlying voters' judgments, thereby preventing itself from studying questions of efficient information aggregation. Yet such an epistemic perspective seems particularly natural in the context of aggregating judgments (rather than preferences[1]). In our court trial example, the ultimate goal seems indeed to be to find out independent facts (of whether the defendant has broken the contract and whether the contract is legally valid). So, the jury's voting rule should be optimised with respect to the goal that the resulting group judgments are *true*, not that they are fair to the jurors.

This alters the problem of designing a voting rule. Properties of voting rules standardly assumed in judgment aggregation theory, such as respecting unanimous judgments or anonymity, cannot be taken for granted anymore. If they turn out to be justified, they derive their justification from the truth-tracking goal rather than fairness considerations. To illustrate the contrast, suppose each juror expresses the judgment (opinion) that the contract was broken. A collective 'broken' judgment would then of course count as good from the classical social-choice theoretic perspective of procedural fairness. However, from a truth-tracking perspective, much depends on questions such as whether the jurors' judgments are sufficient evidence for breach of contract, and whether voters have expressed their judgments truthfully.

This chapter analyses judgment aggregation from the truth-tracking and strategic-voting perspective. We model voters' private information, allowing us to ask questions about efficient information aggregation and strategic voting in a Bayesian voting game setting. Though new within judgment aggregation theory, this approach is well-established in a different body of literature about voting between two alternatives, which started with seminal work by Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1997) and can be placed in the broader context of work on the Condorcet Jury Theorem (see the review below). In the base-line case, voters share a common interest of finding out the 'correct' alternative, but hold possibly conflicting private information about which of the alternatives might be 'correct'. The voting rule should be designed so as to help finding the 'correct' alternative by making optimal use of all the private information scattered across the voters. So, the goal is *efficient information aggregation.* Such an 'epistemic' binary collective choice problem can in fact be viewed as a special judgment aggregation problem, involving just one issue. Our court trial example involves two issues: firstly, whether the contract was broken, and secondly, whether it is legally valid. If instead only the first issue were on the jury's agenda, the jury would face a single-issue judgment aggregation problem, or equivalently, a binary collective choice problem. The entire machinery and results of the mentioned binary collective choice literature could then be applied in order to design the voting rule.

This chapter and the following two chapters therefore combines two bodies of work, namely the judgment aggregation literature and the mentioned binary col-

[1] In preference aggregation theory, the core of social choice theory, an epistemic perspective would be less natural since there is no 'true preference' to be found.

lective choice literature.[2] We believe that these two literatures can learn from each other, and that a fruitful combination can help fill gaps in each of them. Indeed, it seems important that the former benefits from methodologies developed by the latter, and that the latter is extended beyond single-issue agendas towards more complex agendas with multiple issues. Analysing this multi-issue case does not reduce to analysing each issue separately, since preferences establish links between different issues.

It is worth starting simple. We therefore assume that the group faces an agenda with just *two* issues, the simplest kind of multi-issue agenda; but many of our results generalize easily. Though simple, agendas with just two issues are important in practice. Our court trial example and United Nations example each involve two issues. To mention further two-issue agendas, a medical commission might need to issue joint judgments on whether a therapy is effective, and whether it is compatible with given ethical standards; members of a political party in charge of elaborating the party programme might seek joint judgments on whether a tax cut is affordable, and whether it is popular; a university hiring committee might seek joint judgments on whether a given candidate is good at research, and whether he or she is good at teaching; and finally, economic advisors to a government during the banking crisis in 2008 might need to issue collective judgments on whether a given bank has short-term liquidity problems, and whether it has long-term liquidity problems.

The issues of an agenda could in principle be mutually interconnected, so that the judgments taken on the issues logically constrain each other; for instance, a 'no' judgment on all issues might be inconsistent. Indeed, interconnections are what render judgment aggregation non-trivial if the usual social-choice theoretic approach of procedural fairness is taken.[3] However, within our truth-tracking approach, mechanism design is non-trivial even if the issues are mutually independent. We therefore assume independence between issues in this chapter, while we study the case of interconnections in the next chapter.

Section 3.2 introduces our model, in which voters vote on the basis of private information and are guided by 'truth-tracking preferences', i.e., aim for true collective judgments. Section 3.3 addresses the key question of how to design the voting rule such that it leads to efficient decisions as well as simple-minded, truthful voting behaviour in equilibrium. It will turn out that in many, but not all cases one should use a 'quota rule', which decides on each issue according to whether the number of 'yes' judgments on the issue exceeds a particular quota. The details depend on

[2] Recent works by Ahn and Oliveros (2011) and Eliaz and de Clippel (2011) follow a similar approach by combining the two bodies of work, however by asking different questions. See the literature review.

[3] In the absence of interconnections one can safely aggregate by taking a separate vote on each issue. This never generates inconsistent collective judgments and meets all standard social-choice theoretic requirements such as anonymity.

the exact kind of truth-tracking preferences, i.e., whether preferences are 'simple' or 'consequentialist' in a sense defined below. Under simple preferences, the only voting rule which induces an efficient and truthful Bayesian Nash equilibrium is the quota rule defined precisely through the model parameters. Under consequentialist preferences of type 1, an additional monotonicity requirement on the voting rules lead to the quota rule defined equivalently. Under consequentialist preferences of type 2, we characterize voting rules which lead to an efficient and truthful Bayesian Nash equilibrium by 'quota rules with exception' of which quota rules are a special case. Section 3.4 analyses the notion of truthful behaviour, by determining the conditions under which a 'sincere' voter directly reveals his information in his vote. Finally, the appendix contains all proofs.

3.1.1 Literature review

We now selectively review the two literatures to which Part II connects, beginning with judgment aggregation theory. As mentioned, this theory's primary objective has so far been to find out which voting rules can aggregate the judgments of group members over some issues in accordance with certain axiomatic requirements with a classic social-choice theoretic flavour, such as *unanimity preservation* (the counterpart of the *Pareto principle*) and *independence* (the counterpart of Arrow's *independence of irrelevant alternatives*). A series of possibility and impossibility results successfully address this query, by giving answers which depend, firstly, on the axiomatic requirements on the voting rule, and secondly, on the agenda of issues under consideration (e.g., List and Pettit 2002, Dietrich 2006, 2007, 2010, Nehring and Puppe 2008, 2010, Dietrich and List 2007a, 2008, Dokow and Holzman 2010a, 2010b, Dietrich and Mongin 2010; see also precursor results by Guilbaud 1952 and Wilson 1975; for an introductory overview see List and Polak 2010). By contrast, a small minority of papers about judgment aggregation take a truth-tracking perspective (e.g., Bovens and Rabinowicz 2006, List 2005 and Pivato 2011). Their innovation is to apply the classical Condorcet Jury Theorem to judgment aggregation. Despite taking a truth-tracking perspective, they have little in common with our work, since private information and strategic incentives are not being considered.[4] A recent work by Eliaz and de Clippel (2011) consider a judgment aggregation problem with common values and private information. They compare the asymptotic efficiency of premise-based and outcome-based methods. Ahn and Oliveros (2011) study elections with two issues in the context of Condorcet Jury Theorem and make

[4]Ahn and Oliveros (2012) study multi-issue elections where voters have private values over the issues. Dietrich and List (2007b) analyse strategic voting in judgment aggregation, but in a sense not relevant to us since strategic voting is not modelled as coming from private information and a voter is motivated by the somewhat different goal that the collective judgments match *his own* judgments. Such assumptions are more natural under common knowledge of each other's judgments than under informational asymmetry. See also related work by Nehring and Puppe (2002, 2007).

a similar comparison between joint trials and severe trials. Both works study a fixed mechanism to compare different voting games while we take a mechanism design approach. List and Pettit (2011) provide the most systematic philosophical analysis of the truth-tracking approach, already discussing strategic incentives and private information and drawing on the second body of literature to which we now turn.

As for this second body of literature, it is concerned with voting rules for binary choice problems in which disagreements are driven (partly or totally) by conflicting information rather than conflicting interests. Specifically, the utilities which voters derive from decisions are affected by the same unknown 'state of the world', about which voters have private information. Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1997) show that it typically cannot be rational for all voters to vote sincerely, and that the choice of voting rule matters considerably for sincere voting and efficient information aggregation. While the former authors consider the 'purely epistemic' case without conflict of interest, the latter authors introduce some preference heterogeneity (and focus primarily on large electorates). Austen-Smith and Feddersen (2005, 2006) add an extra dimension of pre-voting deliberation. Duggan and Martinelli (2001) and Meirowitz (2002) extend the approach to continuous rather than binary private information. Feddersen and Pesendorfer (1998), Coughlan (2000) and Gerardi (2000) examine the (in)effectiveness of unanimity rule in 'protecting the innocent' in jury trials. Goertz and Maniquet (2011) analyse efficient information aggregation in large electorates, showing that approval voting outperforms other voting rules in their setting.

3.2 The Model

3.2.1 A simple judgment aggregation problem

We consider a group of voters, labelled $i = 1, ..., n$, where $n \geq 2$. This group needs a collective judgment on whether some proposition p or its negation $\bar{p}$ is true, and whether some other proposition q or its negation $\bar{q}$ is true. In our court trial example, p states that the contract was broken, and q that it is legally valid; in our job candidate example, p states that the candidate is good at research, and q that he or she is good at teaching; and so on for our other examples. The four possible judgment sets are $\{p, q\}, \{p, \bar{q}\}, \{\bar{p}, q\}$ and $\{\bar{p}, \bar{q}\}$; we abbreviate them by pq, $p\bar{q}$, $\bar{p}q$ and $\bar{p}\bar{q}$, respectively. For instance, $p\bar{q}$ means accepting p but not q. Each voter votes for a judgment set in $\mathcal{J} = \{pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$. After all voters cast their votes, a collective decision in $\mathcal{J}$ is taken using a voting rule. Formally, a voting rule is a function $f : \mathcal{J}^n \to \mathcal{J}$, mapping each voting profile $\mathbf{v} = (v_1, ..., v_n)$ to a decision $d \equiv f(\mathbf{v})$. Among the various voting rules, *quota rules* stand out as particularly natural and common. A quota rule is given by two thresholds $m_p, m_q \in \{0, 1, ..., n+1\}$, and for each voting profile it accepts p $[q]$ if and only if at least m_p $[m_q]$ voters accept it in the profile. Quota rules have three salient properties:

- *Anonymity:* For all voting profiles $(v_1, ..., v_n) \in \mathcal{J}^n$ and all permutations $(i_1, ..., i_n)$ of the voters, $f(v_{i_1}, ..., v_{i_n}) = f(v_1, ..., v_n)$. Informally, the voters are treated equally.

- *Monotonicity:* For all voting profiles $\mathbf{v}, \mathbf{v}' \in \mathcal{J}^n$, if for each r in $f(\mathbf{v})$ the voters who accept r in $\mathbf{v}$ also accept r in $\mathbf{v}'$, then $f(\mathbf{v}') = f(\mathbf{v})$. Informally, additional support for the collectively accepted propositions never reverses the collective acceptance of these propositions.

- *Independence:* The decision on each proposition $r \in \{p, q\}$ only depends on the votes on r.[5] Informally, the group in effect takes two separate votes, one between p and $\bar{p}$ and one between q and $\bar{q}$.

Remark 3.2.1 *A voting rule $f : \mathcal{J}^n \to \mathcal{J}$ is a quota rule if and only if it is anonymous, monotonic and independent.*

We briefly sketch the proof of the non-trivial direction of implication. As can be shown, if a voting rule $f : \mathcal{J}^n \to \mathcal{J}$ is anonymous and independent, then it is given by two sets $M_p, M_q \subseteq \{0, 1, ..., n\}$, in the sense that for each voting profile $\mathbf{v} \in \mathcal{J}^n$ the decision $f(\mathbf{v})$ contains r ($\in \{p, q\}$) if and only if the number of votes in $\mathbf{v}$ containing r belongs to M_r. If f is moreover monotonic, each set M_r can be shown to take the form $\{m_r, m_r + 1, ..., n\}$ for some threshold $m_r \in \{0, 1, ..., n + 1\}$. Clearly, f is the quota rule with thresholds m_p and m_q.

3.2.2 A common preference for true collective judgments

Exactly one judgment set in $\mathcal{J}$ is 'correct', i.e., contains propositions which are factually true. It is called the *state (of the world)* and is generically denoted by s. For instance, the state might be $p\bar{q}$, so that p and $\bar{q}$ are true (and $\bar{p}$ and q are false). Voters have identical preferences, captured by a common utility function $u : \mathcal{J} \times \mathcal{J} \to \mathbb{R}$ which maps any decision-state pair (d, s) to its utility $u(d, s)$. Given voters' truth-tracking goal, one would expect $u(d, s)$ to be high if $d = s$, i.e., if the decision is correct. But how exactly should u be specified? We focus on two natural kinds of preferences:

Simple preferences. Here, the utility function is given by

$$u(d, s) = \begin{cases} 1 & \text{if } d = s \text{ (correct decision)} \\ 0 & \text{if } d \neq s \text{ (incorrect decision).} \end{cases} \tag{3.2.1}$$

[5] Given a voting profile $\mathbf{v}$, the subprofile with respect to r is denoted $\mathbf{v}_r$ ($\in \{r, \bar{r}\}^n$), and the collective decision with respect to r is denoted $f_r(\mathbf{v})$ ($\in \{r, \bar{r}\}$). Independence means that for all voting profiles $\mathbf{v}, \mathbf{v}' \in \mathcal{J}^n$, if $\mathbf{v}_r = \mathbf{v}'_r$, then $f_r(\mathbf{v}) = f_r(\mathbf{v}')$.

Such preferences are the simplest candidate for truth-tracking preferences. Correct decisions are preferred to incorrect ones, without further sophistication.

Consequentialist preferences. Here, we assume that the decision leads to one of two possible consequences, typically representing group actions. This is captured by a *consequence function* Co which maps the set of possible decisions $\mathcal{J}$ to a two-element set of possible consequences. The consequence function might look as follows in examples given earlier. In our court trial example, the court decision pq leads to conviction, since both premises of guilt are found to be satisfied ($\text{Co}(pq) =$ 'conviction'), while the other decisions all lead to acquittal ($\text{Co}(\bar{p}\bar{q}) = \text{Co}(p\bar{q}) = \text{Co}(\bar{p}q) =$ 'acquittal'). In our job candidate example, the decision pq leads to a hire since the candidate is seen as meeting both criteria ($\text{Co}(pq) =$ 'hire'), while the other decisions all lead to no hire ($\text{Co}(\bar{p}\bar{q}) = \text{Co}(p\bar{q}) = \text{Co}(\bar{p}q) =$ 'no hire'). In our United Nations example, the decisions $p\bar{q}$ and $\bar{p}q$ each lead to a large-scale international intervention in country X ($\text{Co}(p\bar{q}) = \text{Co}(\bar{p}q) =$ 'intervention'), whereas the decisions pq and $\bar{p}\bar{q}$ both lead to no intervention since the United Nations then consider an intervention as being too risky or unnecessary, respectively ($\text{Co}(pq) = \text{Co}(\bar{p}\bar{q}) =$ 'no intervention'). In our bank rescuing example, the decisions $p\bar{q}$ and $\bar{p}q$ each lead to a governmental rescue plan for the bank ($\text{Co}(p\bar{q}) = \text{Co}(\bar{p}q) =$ 'rescue'), whereas the decisions pq and $\bar{p}\bar{q}$ both lead to no rescue plan since a rescue is seen as infeasible or unnecessary, respectively ($\text{Co}(pq) = \text{Co}(\bar{p}\bar{q}) =$ 'no rescue'). The consequentialist utility function is given by

$$u(d,s) = \begin{cases} 1 & \text{if } \text{Co}(d) = \text{Co}(s) \text{ (correct consequence)} \\ 0 & \text{if } \text{Co}(d) \neq \text{Co}(s) \text{ (incorrect consequence).} \end{cases} \tag{3.2.2}$$

Incorrect decisions ($d \neq s$) can have correct consequences ($\text{Co}(d) = \text{Co}(s)$). The hiring committee might view the candidate as good at research and bad at teaching when in fact the opposite is true, so that the resulting consequence ('no hire') is correct for wrong reasons. This gives high utility under consequentialist preferences, but low utility under simple preferences.[6]

3.2.3 Private information and strategies

If voters had not just common preferences, but also common information about what the state might be, then no disagreement could arise. We however allow for informational asymmetry. Each voter has a *type*, representing private information

[6] In the judgment aggregation literature, the two possible consequences are usually represented by two conclusion propositions, c and $\bar{c}$. In our first two examples, the consequence function is encoded in the biconditional $c \leftrightarrow (p \wedge q)$, whereas in our last two examples it is encoded in the biconditional $c \leftrightarrow ((p \wedge q) \vee (\bar{p} \wedge \bar{q}))$.

or evidence.[7] A voter's type takes the form of an element of $\mathcal{J}$, generically denoted by t. For instance, a voter of type $t = p\bar{q}$ has evidence for p and for $\bar{q}$. We write $\mathbf{t} = (t_1, ..., t_n) \in \mathcal{J}^n$ for a profile of voters' types. Nature draws a state-types combination $(s, \mathbf{t}) \in \mathcal{J}^{n+1}$ according to a probability measure denoted Pr. When a proposition $r \in \{p, \bar{p}, q, \bar{q}\}$ is meant to represent part of voter i's type rather than part of the true state, we often write r_i for r. For instance, $\Pr(p_i|p)$ is the probability that voter i has evidence for p given that p is true. By assumption, the prior probability that r ($\in \{p, \bar{p}, q, \bar{q}\}$) is true is denoted

$$\pi_r = \Pr(r)$$

and belongs to $(0, 1)$, and the probability of getting evidence for r given that r is true is denoted

$$a_r = \Pr(r_i|r),$$

belongs to $(1/2, 1)$, and does not depend on the voter i. The parameters $a_p, a_{\bar{p}}, a_q, a_{\bar{q}}$ measure the reliability of private information, as they represent probabilities of receiving 'truth-telling' information. The lower bound of $1/2$ reflects the (standard) idea that information is more reliable than a fair coin.

By assumption, voters' types are independent conditional on the state, and in addition the state and the types w.r.t. p are independent of the state and the types w.r.t. q.[8] These independence assumptions allow one to express the joint distribution of the state and the types by just a few parameters, namely $\pi_p, \pi_q, a_p, a_{\bar{p}}, a_q, a_{\bar{q}}$. For instance, the probability that the state is pq and all voters receive the truth-telling evidence pq is

$$\Pr(pq, p_1q_1, p_2q_2, ..., p_nq_n) = \Pr(pq)\Pr(p_1q_1, p_2q_2, ..., p_nq_n|pq) = \pi_p\pi_q a_p^n a_q^n.$$

Each voter submits a vote in $\mathcal{J}$ based on his type. A *(voting) strategy* is a function $\sigma : \mathcal{J} \to \mathcal{J}$, mapping each type $t \in \mathcal{J}$ to the type's vote $v = \sigma(t)$. We write $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ for a profile of voters' strategies. Together with a voting rule f and a common utility function u, we now have a well-defined Bayesian game.

For a given type profile $\mathbf{t} \in \mathcal{J}^n$, we call a decision d *efficient* if it has maximal expected utility conditional on the full information $\mathbf{t}$.[9] Some common notions of voting behaviour can now be adapted to our framework:

[7] The type could represent information that is not shared with other voters because of a lack of deliberation or limits of deliberation. More generally, a voter i's type could represent uncertainty of other voters about i's beliefs.

[8] Recall that the state consists of a proposition in $\{p, \bar{p}\}$ and another in $\{q, \bar{q}\}$. The first [second] of these propositions is what we call the *state w.r.t.* p [q]. A voter's *type w.r.t.* p [q] is defined similarly.

[9] I.e., d maximizes $E(u(d, S)|\mathbf{t}) = \sum_{s \in \mathcal{J}} u(d, s)\Pr(s|\mathbf{t})$, where '$S$' denotes the random variable generating the state s in $\mathcal{J}$.

- A strategy σ of a voter is *informative* if $\sigma(t) = t$ for all types t. An informative voter directly reveals his information in his vote.

- A strategy σ of a voter is *sincere* if for every type t, the vote $\sigma(t)$ maximises the expected utility conditional on the information t. A sincere voter votes for the decision which maximises the expected utility conditional on his type; so, he acts as if his vote alone determined the decision, neglecting the other voters and their strategies. Technically, this amounts to optimal behaviour in a hypothetical single-player decision problem.

- A strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ is *rational* if each strategy is a best response to the other strategies, i.e., if the profile is a Nash equilibrium of the corresponding Bayesian game. Hence, each voter maximises the expected utility of the collective decision given the strategies of the other voters. (In this maximisation exercise, it turns out that a voter must only consider cases in which his vote is pivotal. Under a quota rule with majority thresholds, a voter is for instance pivotal if half of the other voters votes pq and the other half votes $\bar{p}\bar{q}$.)

- A strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ is *efficient* if for every type profile $\mathbf{t} = (t_1, ..., t_n)$ the resulting decision $d = f(\sigma_1(t_1), ..., \sigma_n(t_n))$ is efficient (i.e., has maximal expected utility conditional on full information $\mathbf{t}$). Hence, all the information spread across the group is used efficiently: the collective decision is no worse than a decision of a (virtual) social planner who has full information.

While informativeness and sincerity are properties of a single strategy (or voter), rationality and efficiency refer to an entire profile.

Finally, to avoid distraction by special cases, we make two assumptions. First, we exclude the degenerate case in which some decision in $\mathcal{J}$ is not efficient for any type profile whatsoever. Second, we exclude efficiency ties, i.e., we exclude those special parameter combinations such that some type profile $\mathbf{t}$ leads to different efficient decisions (with different consequences when we assume consequentialist preferences).

3.3 Which voting rules lead to efficient information aggregation?

3.3.1 Setting the stage

Our objective is to design the voting rule ('mechanism') in such a way as to yield efficient decisions on the basis of informative votes. In short, the voting rule should render informative voting efficient.[10] We begin by justifying this objective. *Prima facie*, two goals are of interest. The rule should, firstly, lead to efficient outcomes, and, secondly, encourage simple-minded, truthful behaviour. By such behaviour we mean informative voting.[11] To reach the second goal, informative voting should be rational, i.e., occur in equilibrium. If informative voting is not just rational, but also efficient, both goals are reached. So, the double-goal is that informative voting be efficient and rational. Following a well-known result by McLennan (1998), whenever informative voting is efficient, it is *a fortiori* also rational – which explains our primary objective that informative voting be efficient.

Theorem 3.3.1 *Consider an arbitrary common utility function $u : \mathcal{J}^2 \to \mathbb{R}$.*

(a) *For any voting rule, if a strategy profile is efficient, then it is rational (McLennan 1998).*

(b) *There is an anonymous voting rule for which informative voting is efficient (hence, rational).*

This theorem is general in that it applies to any kind of (common) preferences. The converse of part (a) does not hold: for instance, a constant voting rule makes all strategy profiles rational, but typically not efficient. The message of part (b) is positive but so far vague: it is always possible to make informative voting efficient (and rational), but apart from anonymity we do not know anything about the kind of voting rule we can use. And indeed, for some kinds of common preferences, it may not be possible to aggregate in an independent or monotonic way (as counterexamples show). But, once we narrow down to simple or consequentialist preferences, can – or even must – we aggregate in a monotonic resp. independent way? When can – or even must – we use a quota rule? Such questions are answered below.

3.3.2 Simple preferences

This section addresses the case of simple preferences, given by the common utility function (3.2.1). Which rules render informative voting efficient (hence, rational)? The answer is 'simple', as we will see. To state our result, we first define two

[10] By saying "informative voting" without referring to a particular voter, we mean "informative voting by all voters".

[11] One might alternatively mean sincere voting – but in practice there is little difference, since informative and sincere voting coincide under reasonable informational assumptions. As one can show, if informative voting is *not* sincere, then there exists a decision $d \in \mathcal{J}$ such that no voter ever finds himself in an informational position to consider d as best – a rather uninteresting, if not unnatural scenario.

coefficients:[12]

$$k_p := \min \left\{ k \in \{0, 1, ..., n+1\} : \frac{\pi_p}{1-\pi_p} > \left(\frac{1-a_{\bar{p}}}{a_p}\right)^k \left(\frac{a_{\bar{p}}}{1-a_p}\right)^{n-k} \right\}, \quad (3.3.1)$$

$$k_q := \min \left\{ k \in \{0, 1, ..., n+1\} : \frac{\pi_q}{1-\pi_q} > \left(\frac{1-a_{\bar{q}}}{a_q}\right)^k \left(\frac{a_{\bar{q}}}{1-a_q}\right)^{n-k} \right\}. \quad (3.3.2)$$

These coefficients have an interpretation: as can be proved, for p $[q]$ to be more probably true than false given all information, at least k_p $[k_q]$ individuals need to receive evidence for p $[q]$, i.e., need to have a type containing p $[q]$.

Theorem 3.3.2 *Assume simple preferences. Informative voting is efficient if and only if f is the quota rule with the thresholds k_p and k_q.*

This result shows that the quota rule with thresholds k_p and k_q is the *only* rule we may use in view of making informative voting efficient (hence, rational). This result is much more specific than the purely existential claim in part (b) of Theorem 3.3.1. This progress was possible by focusing on simple preferences.

3.3.3 Consequentialist preferences: first type

We now turn to consequentialist preferences. Much depends on the nature of the consequence function. In principle, there exist $2^4 = 16$ potential consequence functions from $\mathcal{J}$ to a binary set of consequences. But, as we shall see shortly, there are only two non-degenerate consequence functions up to isomorphism. We therefore define two types of consequentialist functions. Consequentialist preferences (or the consequence function) are said to be:

- of *type 1* if $\text{Co}(pq) = \text{Co}(\bar{p}\bar{q}) \neq \text{Co}(p\bar{q}) = \text{Co}(\bar{p}q)$;
- of *type 2* if $\text{Co}(pq) \neq \text{Co}(\bar{p}\bar{q}) = \text{Co}(p\bar{q}) = \text{Co}(\bar{p}q)$.

Our first two examples of consequentialist preferences in Section 3.2.2 are of type 1, while our last two examples are of type 2. But why are *all* non-degenerate consequences of one of these two types? Firstly, consequence functions for which each decision in $\mathcal{J}$ has the *same* consequence are of course degenerate and therefore uninteresting. Also consequence functions which depend only on the decision between p and $\bar{p}$, or only on the decision between q and $\bar{q}$, are degenerate, since in this case we

[12]The minimum defining k_p or k_q should be interpreted as $n+1$ if the set whose minimum is being taken is *empty*. In fact, emptiness is impossible under simple preferences. This follows from our non-degeneracy assumption on the model parameters (which also implies that $k_p, k_q \in \{1, ..., n\}$). Note that in (3.3.1) and (3.3.2) the right hand side of the inequality is strictly decreasing in k.

are essentially back to a decision problem with a single proposition-negation pair, which has already been studied in the literature.[13] The non-degenerate consequence functions are those which genuinely depend on both propositions. Among all of them, some assign each consequence to exactly two decisions in $\mathcal{J}$, while the others assign one consequence to three decisions and the other consequence to just one decision. As one can show, the former consequence functions are of type 1, while the latter are of type 2 up to isomorphism (i.e., up to exchanging p and $\bar{p}$ and/or exchanging q and $\bar{q}$). Thus, by studying our two types of consequence functions, we will have covered non-degenerate consequentialist preferences exhaustively.

We now address the first type, while the next subsection turns to the second type. One might at first expect there to be little resemblance between the current preferences and simple preferences in terms of the appropriate voting rule. For instance, even when all individuals have type pq, so that there is overwhelming evidence for state pq, the current preferences allow us to efficiently decide for $\bar{p}\bar{q}$, since this decision has the same consequence as pq. Surprisingly, despite the differences, consequentialist preferences of type 1 come much closer to simple preferences than to consequentialist preferences of type 2 in terms of mechanism design. The coefficients k_p and k_q, defined earlier for simple preferences, again play a key role.

Theorem 3.3.3 *Assume consequentialist preferences of type 1. A voting rule f makes informative voting efficient and is monotonic if and only if it is the quota rule with thresholds k_p and k_q.*

So, as for simple preferences, the social planner is led to impose a quota rule with the particular thresholds k_p and k_q. What distinguishes Theorem 3.3.3 from Theorem 3.3.2 is, for one, its somewhat different (and longer) proof, and secondly, the additional monotonicity requirement. Without this extra condition, a number of other voting rules become possible:

Corollary 3.3.1 *Assume consequentialist preferences of type 1. A voting rule f makes informative voting efficient if and only if for every voting profile $\mathbf{v} \in \mathcal{J}^n$ the decision $f(\mathbf{v})$ has the same consequence as the decision under the quota rule with thresholds k_p and k_q (i.e., $\mathrm{Co} \circ f = \mathrm{Co} \circ g$, where g is this quota rule).*

So, once we drop the monotonicity requirement, there is not just one possible voting rule, as for simple preferences, but 2^{4^n} possible rules (since there are 2 allowed decisions for each of the 4^n profiles in $\mathcal{J}^n$).

[13] For instance, our UN intervention example would be degenerate if the question of whether to intervene only depended on whether the country is considered as being threatened by a military coup (p or $\bar{p}$). The other pair of propositions (q or $\bar{q}$) could then be eliminated from the voting process.

3.3.4 Consequentialist preferences: second type

We now turn to consequentialist preferences of type 2. The space of aggregation possibilities is somewhat different here. As we shall show, quota rules are *not* always possible, and when they are, the two thresholds must be calculated differently.

For all $k, l \in \mathbb{R}$, we define the coefficient

$$\beta(k,l) = \frac{\pi_p a_p^k (1-a_p)^{n-k}}{\pi_p a_p^k (1-a_p)^{n-k} + \pi_{\bar{p}} a_{\bar{p}}^{n-k} (1-a_{\bar{p}})^k} \times \frac{\pi_q a_q^l (1-a_q)^{n-l}}{\pi_q a_q^l (1-a_q)^{n-l} + \pi_{\bar{q}} a_{\bar{q}}^{n-l} (1-a_{\bar{q}})^l}. \tag{3.3.3}$$

One can show that $\beta(k,l)$ has a natural interpretation if $k, l \in \{0, 1, ..., n\}$: it is the probability that the state is pq conditional on having k times evidence for (and $n-k$ times evidence against) p and l times evidence for (and $n-l$ times evidence against) q. So, $\beta(k,l) = \Pr(pq|\mathbf{t})$ for some (hence, any) type profile $\mathbf{t} \in \mathcal{J}^n$ containing p exactly k times and q exactly l times; or equivalently,

$$\beta(k,l) = \Pr(p|p_1, ...p_k, \bar{p}_{k+1}, ..., \bar{p}_n) \times \Pr(q|q_1, ..., q_l, \bar{q}_{l+1}, ..., \bar{q}_n).$$

As one can prove by drawing on the definition of the consequence function, given a type profile $\mathbf{t}$ containing p exactly k times and q exactly l times, if $\beta(k,l) > 1/2$ then only the decision pq is efficient, while otherwise the three other decisions are all efficient. This implies a first, simple characterization result. Henceforth, the number of votes for a proposition r in a voting profile $\mathbf{v}$ is written $n_r^{\mathbf{v}}$.

Proposition 3.3.1 *Assume consequentialist preferences of type 2. A voting rule f makes informative voting efficient if and only if for every voting profile $\mathbf{v} \in \mathcal{J}^n$ the decision $f(\mathbf{v})$ is pq if $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$ and in $\{p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$ otherwise.*

Which possibilities – if any – are left if we require the rule to be a quota rule? We begin by introducing two coefficients. Given that all voters hold evidence for q, how many voters with evidence for p does it minimally take for the decision pq to become efficient? Similarly, given that all voters hold evidence for p, how many voters with evidence for q does it take for the decision pq to become efficient? The answer to these questions is given by the following numbers, respectively:[14]

$$l_p := \min\{k \in \{0, ..., n\} : \beta(k,n) > 1/2\} \tag{3.3.4}$$
$$l_q := \min\{k \in \{0, ..., n\} : \beta(n,k) > 1/2\}. \tag{3.3.5}$$

Theorem 3.3.4 *Assume consequentialist preferences of type 2. There exists a quota rule making informative voting efficient if and only if $\beta(l_p, l_q) > 1/2$. In this case, that quota rule is unique and has the thresholds l_p and l_q.*

[14] These two minima are taken over *non-empty* sets of values of k (by the non-degeneracy assumption at the end of Section 3.2.3).

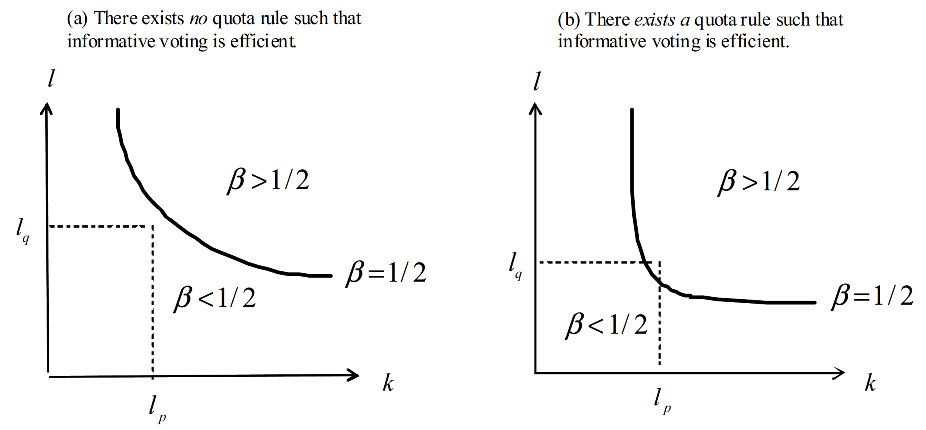


Figure 3.3.1: The function β

Figure 3.3.1b illustrates the region to which (l_p, l_q) must belong for a quota rule to be available. Unlike when preferences are simple or consequentialist of type 1, and unlike in the classic literature for a single pair of propositions $p, \bar{p}$, we have a partial impossibility:

Corollary 3.3.2 *Assume consequentialist preferences of type 2. For some, but not all combinations of values of the model parameters ($\pi_p, \pi_q, a_p, a_{\bar{p}}, a_q, a_{\bar{q}}$ and n), there exists a quota rule making informative voting efficient.*

For instance, if $\pi_p = \pi_q = 0.5$, $a_p = a_q = a_{\bar{p}} = a_{\bar{q}} = 0.7$ and $n = 3$, no quota rule makes informative voting efficient, whereas if instead $\pi_p = \pi_q = 0.6$, the quota rule with thresholds $l_p = l_q = 2$ makes informative voting efficient.

While by Corollary 3.3.2 it may be utopian to aim for a full-fledged quota rule, we now show that one can always achieve two characteristic properties of quota rules, namely anonymity and monotonicity, while often losing the third characteristic property, namely independence. Specifically, we characterize the class of all monotonic and anonymous (but not necessarily independent) aggregation possibilities. As we shall see, this class consists of so-called quota rules 'with exception'. Such rules behave like a quota rule as long as the profile does not fall into an 'exception domain', while generating the 'exception decision' pq on the exception domain. Formally, a *quota rule with exception* $f : \mathcal{J}^n \to \mathcal{J}$ is given by thresholds $m_p, m_q \in \{0, ..., n+1\}$ and an 'exception domain' $\mathcal{E} \subseteq \mathcal{J}^n$, and is defined as follows for all voting profiles $\mathbf{v} \in \mathcal{J}$: if $\mathbf{v} \notin \mathcal{E}$ then $f(\mathbf{v})$ contains any proposition r in $\{p, q\}$ if and only if $n_r^{\mathbf{v}} \geq m_r$, while if $\mathbf{v} \in \mathcal{E}$ then $f(\mathbf{v}) = pq$; or equivalently, $f(\mathbf{v})$ contains any r in $\{p, q\}$ if and only if $[n_r^{\mathbf{v}} \geq m_r$ or $\mathbf{v} \in \mathcal{E}]$.[15] Standard quota rules arise as special cases with an empty exception domain. In our characterization theorem, the exception domain is $\mathcal{E} = \{\mathbf{v} : \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2\}$, so that

$$f(\mathbf{v}) \text{ contains } r \Leftrightarrow [n_r^{\mathbf{v}} \geq m_r \text{ or } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2], \text{ for all } r \in \{p, q\} \text{ and } \mathbf{v} \in \mathcal{J}. \tag{3.3.6}$$

Theorem 3.3.5 *Assume consequentialist preferences of type 2. A voting rule f makes informative voting efficient and is monotonic and anonymous if and only if f is the quota rule with exception (3.3.6) for some thresholds m_p, m_q such that $\beta(m_p, l_q), \beta(l_p, m_q) > 1/2$.*

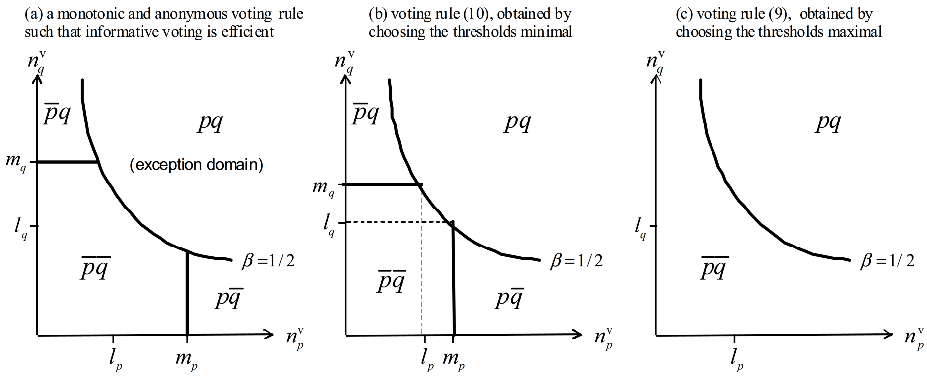


Figure 3.3.2: Illustration of Theorem 3.3.5: the decision as a function of the number of votes for p and q

Figure 3.3.2 shows three voting rules of the kind given in Theorem 3.3.5, which differ in the choice of the thresholds m_p and m_q. In Figure 3.3.2a, the thresholds

[15] The notion of a quota rules with exception could be generalized by allowing the exception decision to differ from pq. The exception decision is pq for us due to the privileged status of pq under consequentialist preferences of type 2.

are chosen in a 'non-extreme' way. In Figure 3.3.2c, the thresholds are maximal, i.e., $m_p = m_q = n + 1$, so that the voting rule takes a particularly simple form:

$$f(\mathbf{v}) = \begin{cases} pq & \text{if } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2 \\ \bar{p}\bar{q} & \text{if } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2 \end{cases} \tag{3.3.7}$$

for all voting profiles $\mathbf{v} \in \mathcal{J}^n$. In Figure 3.3.2b, the thresholds are minimal, so that the voting rule is given as follows:

$$f(\mathbf{v}) = \begin{cases} pq & \text{if } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2 \\ p\bar{q} & \text{if } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2 \text{ and } \beta(n_p^{\mathbf{v}}, l_q) > 1/2 \\ \bar{p}q & \text{if } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2 \text{ and } \beta(l_p, n_q^{\mathbf{v}}) > 1/2 \\ \bar{p}\bar{q} & \text{otherwise} \end{cases} \tag{3.3.8}$$

The latter rule is special in that it reduces to the quota rule making informative voting efficient (defined in Theorem 3.3.4) whenever such a quota rule exists.

3.4 When is informative voting sincere?

While the previous section focuses on mechanism design, the present section does not depend on the voting rule (mechanism). We focus on a single voter and answer the question of when informative voting is sincere, that is, when the naive strategy of 'following the evidence' is worthwhile for a sincere voter. For each type of preference, we fully characterize the parameter combinations for which this is so. We begin with simple preferences.

Theorem 3.4.1 *Under simple preferences, the informative voting strategy is sincere if and only if* $\frac{a_{\bar{r}}}{1-a_r} \geq \frac{\pi_r}{1-\pi_r} \geq \frac{1-a_{\bar{r}}}{a_r}$ *for each* $r \in \{p, q\}$.

This result has an intuitive interpretation. We know that necessarily the upper bound $\frac{a_{\bar{r}}}{1-a_r}$ for $\frac{\pi_r}{1-\pi_r}$ exceeds 1 and the lower bound $\frac{1-a_{\bar{r}}}{a_r}$ is below 1, since $a_r, a_{\bar{r}} > 1/2$. For very high or very low values of the prior probabilities π_r, the ratio $\frac{\pi_r}{1-\pi_r}$ is far from 1, so that one of the bounds is violated and informative voting is *not* sincere. This makes sense since if voters have 'strong' prior beliefs, then the evidence collected cannot overrule the prior beliefs: sincere votes cease to be sensitive to evidence, i.e., depart from informative votes. By contrast, for less strong prior beliefs, the inequalities are satisfied, so that informative voting is sincere, i.e., it is worth following the evidence as a sincere voter.

Another useful perspective on the result is obtained by focusing not on the parameters π_r representing prior beliefs, but on the parameters a_r and $a_{\bar{r}}$ representing 'strength of evidence'. The larger a_r and $a_{\bar{r}}$ are (i.e., the 'stronger' private evidence for r and $\bar{r}$ is), the greater the upper bound for $\frac{\pi_r}{1-\pi_r}$ is and the smaller the lower bound is, which makes it easier to meet both inequalities. In summary, sufficiently strong evidence and/or sufficiently weak prior beliefs imply that it is worth voting informatively ('following the evidence') as a sincere voter.

Surprisingly, the characterization remains the same as we move from simple preferences to consequentialist preferences of type 1 (though the proof is quite different):

Theorem 3.4.2 *Under consequentialist preferences of type 1, the informative voting strategy is sincere if and only if* $\frac{a_{\bar{r}}}{1-a_r} \geq \frac{\pi_r}{1-\pi_r} \geq \frac{1-a_{\bar{r}}}{a_r}$ *for each* $r \in \{p, q\}$.

One can interpret this result in a similar way as done for simple preferences.

Finally, we turn to consequentialist preferences of type 2. Here, the characterization is based on the following three coefficients:

$$A := \frac{\pi_p}{1-\pi_p} \times \frac{a_{\bar{q}}}{1-a_q} + \frac{\pi_q}{1-\pi_q} \times \frac{1-a_{\bar{p}}}{a_p} + \frac{1-a_{\bar{p}}}{a_p} \times \frac{a_{\bar{q}}}{1-a_q}$$
$$B := \frac{\pi_p}{1-\pi_p} \times \frac{1-a_{\bar{q}}}{a_q} + \frac{\pi_q}{1-\pi_q} \times \frac{a_{\bar{p}}}{1-a_p} + \frac{a_{\bar{p}}}{1-a_p} \times \frac{1-a_{\bar{q}}}{a_q}$$
$$C := \frac{\pi_p}{1-\pi_p} \times \frac{1-a_{\bar{q}}}{a_q} + \frac{\pi_q}{1-\pi_q} \times \frac{1-a_{\bar{p}}}{a_p} + \frac{1-a_{\bar{p}}}{a_p} \times \frac{1-a_{\bar{q}}}{a_q}.$$

Theorem 3.4.3 *Under consequentialist preferences of type 2, the informative voting strategy is sincere if and only if* $A, B \geq \frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} \geq C$.

Although the characterizing inequalities are more complicated than for the previous two kinds of preference, an interpretation in terms of strength of evidence is again possible. If the voter's evidence is sufficiently strong (i.e., if $a_p, a_{\bar{p}}, a_q, a_{\bar{q}}$ are sufficiently close 1), then C is well below 1 and A and B are well above 1, so that the inequalities are likely to hold; as a result, informative voting is sincere, i.e., it is worth following the evidence as a sincere voter.

3.5 Appendix: proofs

We begin by some preliminary derivations, and then prove our results in a new order obtained by clustering the results according to the kind of preference.

Conventions. Recall the notation 'f_r' introduced in fn. 5 and the notation 'S' for the random variable generating the state s in $\mathcal{J}$ introduced in fn. 9. Double-negations cancel each other out, i.e., $\bar{\bar{p}}$ stands for p, and $\bar{\bar{q}}$ for q. We refer to the two technical assumptions made at the end of Section 3.2.3 as 'non-degeneracy' and 'no efficiency ties', respectively.

3.5.1 Preliminary derivations

The joint probability of a state-types vector $(s, \mathbf{t}) = (s_p s_q, t_{1p} t_{1q}, ..., t_{np} t_{nq}) \in \mathcal{J}^{n+1}$ is

$$\Pr(s, \mathbf{t}) = \Pr(s) \Pr(\mathbf{t}|s) = \Pr(s) \prod_i \Pr(t_i|s) = \Pr(s_p) \Pr(s_q) \prod_i \Pr(t_{ip}|s_p) \Pr(t_{iq}|s_q),$$

where the last two equations follow from our independence assumptions. A voter's probability of a state $s = p_s q_s \in \mathcal{J}$ given his type $t = p_t q_t \in \mathcal{J}$ is given by $\Pr(s|t) = \Pr(p_s|p_t) \Pr(q_s|q_t)$, which reduces to

$$\Pr(s|t) = \frac{\pi_{p_s} a_{p_s}}{\pi_{p_s} a_{p_s} + \pi_{\overline{p_s}}(1 - a_{\overline{p_s}})} \times \frac{\pi_{q_s} a_{q_s}}{\pi_{q_s} a_{q_s} + \pi_{\overline{q_s}}(1 - a_{\overline{q_s}})} \text{ if } p_s = p_t, q_s = q_t \quad (3.5.1)$$

$$\Pr(s|t) = \frac{\pi_{p_s} a_{p_s}}{\pi_{p_s} a_{p_s} + \pi_{\overline{p_s}}(1 - a_{\overline{p_s}})} \times \frac{\pi_{q_s}(1 - a_{q_s})}{\pi_{q_s}(1 - a_{q_s}) + \pi_{\overline{q_s}} a_{\overline{q_s}}} \text{ if } p_s = p_t, q_s \neq q_t \quad (3.5.2)$$

$$\Pr(s|t) = \frac{\pi_{p_s}(1 - a_{p_s})}{\pi_{p_s}(1 - a_{p_s}) + \pi_{\overline{p_s}} a_{\overline{p_s}}} \times \frac{\pi_{q_s} a_{q_s}}{\pi_{q_s} a_{q_s} + \pi_{\overline{q_s}}(1 - a_{\overline{q_s}})} \text{ if } p_s \neq p_t, q_s = q_t \quad (3.5.3)$$

$$\Pr(s|t) = \frac{\pi_{p_s}(1 - a_{p_s})}{\pi_{p_s}(1 - a_{p_s}) + \pi_{\overline{p_s}} a_{\overline{p_s}}} \times \frac{\pi_{q_s}(1 - a_{q_s})}{\pi_{q_s}(1 - a_{q_s}) + \pi_{\overline{q_s}} a_{\overline{q_s}}} \text{ if } p_s \neq p_t, q_s \neq q_t \quad (3.5.4)$$

The probability of the four states in $\mathcal{J}$ conditional on the *full* information $\mathbf{t} \in \mathcal{J}^n$ is given as follows, where $k := n_p^{\mathbf{t}}$ and $l := n_q^{\mathbf{t}}$:

$$\Pr(pq|\mathbf{t}) = \frac{\pi_p a_p^k (1 - a_p)^{n-k} \pi_q a_q^l (1 - a_q)^{n-l}}{\Pr(\mathbf{t})} \quad (3.5.5)$$

$$\Pr(p\bar{q}|\mathbf{t}) = \frac{\pi_p a_p^k (1 - a_p)^{n-k} \pi_{\bar{q}} (1 - a_{\bar{q}})^l a_{\bar{q}}^{n-l}}{\Pr(\mathbf{t})} \quad (3.5.6)$$

$$\Pr(\bar{p}q|\mathbf{t}) = \frac{\pi_{\bar{p}} (1 - a_{\bar{p}})^k a_{\bar{p}}^{n-k} \pi_q a_q^l (1 - a_q)^{n-l}}{\Pr(\mathbf{t})} \quad (3.5.7)$$

$$\Pr(\bar{p}\bar{q}|\mathbf{t}) = \frac{\pi_{\bar{p}} (1 - a_{\bar{p}})^k a_{\bar{p}}^{n-k} \pi_{\bar{q}} (1 - a_{\bar{q}})^l a_{\bar{q}}^{n-l}}{\Pr(\mathbf{t})}. \quad (3.5.8)$$

3.5.2 General preferences

Proof of Theorem 3.3.1. (a) We write T_i $(= T_{ip} T_{iq})$ for the random variable generating voter i's type in $\mathcal{J}$, and $\mathbf{T} = (T_1, ..., T_n)$ for the random type profile. Consider any voting rule $f : \mathcal{J}^n \to \mathcal{J}$ and any efficient strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$. To show that $\boldsymbol{\sigma}$ is rational, consider any voter i and type $t_i \in \mathcal{J}$. We have to show that i's vote $\sigma_i(t_i)$ maximizes the expected utility conditional on i's type, i.e., that

$$E(u(f(\sigma_i(t_i), \boldsymbol{\sigma}_{-i}(\mathbf{T}_{-i})), S)|t_i) \geq E(u(f(v_i, \boldsymbol{\sigma}_{-i}(\mathbf{T}_{-i})), S)|t_i) \text{ for all } v_i \in \mathcal{J},$$

where $(\sigma_i(t_i), \boldsymbol{\sigma}_{-i}(\mathbf{T}_{-i}))$ and $(v_i, \boldsymbol{\sigma}_{-i}(\mathbf{T}_{-i}))$ of course denote the voting profiles in which i votes v_i resp. $\sigma_i(t_i)$ and each $j \neq i$ votes $\sigma_j(T_j)$. To show this, note that for all $v_i \in \mathcal{J}$,

$$\begin{aligned} E(u(f(v_i, \boldsymbol{\sigma}_{-i}(\mathbf{T}_{-i})), S)|t_i) &= \sum_{\mathbf{t}_{-i} \in \mathcal{J}^{n-1}} \Pr(\mathbf{t}_{-i}|t_i) E(u(f(v_i, \boldsymbol{\sigma}_{-i}(\mathbf{t}_{-i})), S)|t_i, \mathbf{t}_{-i}) \\ &\leq \sum_{\mathbf{t}_{-i} \in \mathcal{J}^{n-1}} \Pr(\mathbf{t}_{-i}|t_i) E(u(f(\sigma_i(t_i), \boldsymbol{\sigma}_{-i}(\mathbf{t}_{-i})), S)|t_i, \mathbf{t}_{-i}) \\ &= E(u(f(\sigma_i(t_i), \boldsymbol{\sigma}_{-i}(\mathbf{T}_{-i})), S)|t_i), \end{aligned}$$

where the inequality holds because the strategy profile $(\sigma_i, \boldsymbol{\sigma}_{-i}) = \boldsymbol{\sigma}$ is efficient for the type profile $(t_i, \mathbf{t}_{-i}) = \mathbf{t}$.

(b) Since by (3.5.5)-(3.5.8) the conditional distribution of the state given full information $\mathbf{t} \in \mathcal{J}^n$ depends on $\mathbf{t}$ only via the numbers $n_p^{\mathbf{t}}$ and $n_q^{\mathbf{t}}$, so does the conditional expected utility of each decision, and hence, the set of efficient decisions. For each $(k, l) \in \{0, 1, ..., n\}^2$, let $F(k, l) \in \mathcal{J}$ be a decision that is efficient for some (hence, every) $\mathbf{t} \in \mathcal{J}^n$ for which $n_p^{\mathbf{t}} = k$ and $n_q^{\mathbf{t}} = l$. The voting rule f defined by $\mathbf{v} \mapsto f(\mathbf{v}) = F(n_p^{\mathbf{v}}, n_q^{\mathbf{v}})$ is clearly anonymous and renders informative voting efficient. □

3.5.3 Simple preferences

We begin by two lemmas.

Lemma 3.5.1 *Assume simple preferences. The expected utility of a decision $d \in \mathcal{J}$ is*

$$E(u(d, S)) = \Pr(S = d),$$

and the conditional expected utility of d given a type or a type profile is given by the analogous expression with a conditional probability instead of an unconditional one.

Proof. The claim follows immediately from the definition of the utility function. □

The next lemma invokes the coefficients k_p and k_q defined in (3.3.1) and (3.3.2).

Lemma 3.5.2 *Assume simple preferences. For all type profiles $\mathbf{t} \in \mathcal{J}^n$, all $r \in \{p, q\}$, and all decisions $d, d' \in \mathcal{J}$ such that d but not d' contains r, and d and d' share the other proposition,*

$$E(u(d, S)|\mathbf{t}) > E(u(d', S)|\mathbf{t}) \Leftrightarrow n_r^{\mathbf{t}} \geq k_r.$$

PROOF. Let $\mathbf{t} \in \mathcal{J}^n$. We first prove the equivalence for $r = p$, $d = pq$ and $d' = \bar{p}q$. By the definition of k_p, the inequality $n_p^{\mathbf{t}} \geq k_p$ is equivalent to

$$\frac{\pi_p}{1-\pi_p} > \left(\frac{1-a_{\bar{p}}}{a_p}\right)^{n_p^{\mathbf{t}}} \left(\frac{a_{\bar{p}}}{1-a_p}\right)^{n-n_p^{\mathbf{t}}}, \tag{3.5.9}$$

which by (3.5.5) and (3.5.7) is equivalent to $\Pr(pq|\mathbf{t}) > \Pr(\bar{p}q|\mathbf{t})$, and hence by Lemma 3.5.1 to $E(u(pq,S)|\mathbf{t}) > E(u(\bar{p}q,S)|\mathbf{t})$. Next, suppose $r = p$, $d = p\bar{q}$ and $d' = \bar{p}\bar{q}$. Using (3.5.6) and (3.5.8), the inequality (3.5.9) is equivalent to $\Pr(p\bar{q}|\mathbf{t}) > \Pr(\bar{p}\bar{q}|\mathbf{t})$, and hence, to $E(u(p\bar{q},S)|\mathbf{t}) > E(u(\bar{p}\bar{q},S)|\mathbf{t})$. The proof for the remaining cases is analogous. $\square$

We are now in a position to prove the two theorems about simple preferences.

PROOF OF THEOREM 3.3.2. Consider a rule $f : \mathcal{J}^n \to \mathcal{J}$.

A. First, assume f is the quota rule with thresholds k_p and k_q. Consider a given type profile $\mathbf{t} \in \mathcal{J}^n$. Supposing that voters vote informatively, the resulting voting profile is $\mathbf{v} = \mathbf{t}$. We have to show that the decision $d := f(\mathbf{v})$ is efficient for $\mathbf{t}$, i.e., that (*) $E(u(d,S)|\mathbf{t}) > E(u(d',S)|\mathbf{t})$ for all $d' \in \mathcal{J}\backslash\{d\}$. (We use '$>$' rather than '$\geq$' in (*) because of our 'no efficiency ties' assumption.) The property (*) follows from Lemma 3.5.2. For instance, if $d = pq$, then by definition of f we have $n_p^{\mathbf{t}} \geq k_p$ and $n_q^{\mathbf{t}} \geq k_q$, so that Lemma 3.5.2 implies the inequality in (*) for $d' = \bar{p}q$ and $d' = p\bar{q}$

For instance, if $d = pq$, then by definition of f we have $n_p^{\mathbf{t}} \geq k_p$ and $n_q^{\mathbf{t}} \geq k_q$, so that Lemma 3.5.2 implies that

$$E(u(pq,S)|\mathbf{t}) > E(u(\bar{p}q,S)|\mathbf{t}), E(u(p\bar{q},S)|\mathbf{t}) > E(u(\bar{p}\bar{q},S)|\mathbf{t}),$$

which in turn implies (*); and if $d = \bar{p}\bar{q}$, then $n_p^{\mathbf{t}} < k_p$ and $n_q^{\mathbf{t}} < k_q$, so that Lemma 3.5.2 implies that

$$E(u(\bar{p}\bar{q},S)|\mathbf{t}) > E(u(p\bar{q},S)|\mathbf{t}), E(u(\bar{p}q,S)|\mathbf{t}) > E(u(pq,S)|\mathbf{t}),$$

which again implies (*).

B. Conversely, suppose informative voting is efficient under f. We consider any $\mathbf{v} \in \mathcal{J}^n$ and $r \in \{p,q\}$, and must show that (**) $f_r(\mathbf{v}) = r \Leftrightarrow n_r^{\mathbf{v}} \geq k_r$. Consider the type profile $\mathbf{t} = \mathbf{v}$. Since informative voting is efficient, the decision $d = f(\mathbf{v})$ is efficient for $\mathbf{t}$ $(= \mathbf{v})$, i.e., satisfies condition (*) above. Lemma 3.5.2 and (*) together imply (**). For instance, if $f(\mathbf{v}) = pq$, then (**) holds because, firstly, $f_r(\mathbf{v}) = r$, and secondly, $n_r^{\mathbf{v}} \geq k_r$ by (*) and Lemma 3.5.2. $\square$

PROOF OF THEOREM 3.4.1. A. First, assume informative voting is sincere. Equivalently, for any given type $t \in \mathcal{J}$, $E(u(d,S)|t)$ is maximal at $d = t$, i.e., by Lemma

3.5.1 (*) $\Pr(d|t)$ is maximal at $d = t$. Applying (*) to type $t = pq$, we have $\Pr(pq|t) \geq \Pr(\bar{p}q|t)$, which implies $\frac{\pi_p}{1-\pi_p} \geq \frac{1-a_{\bar{p}}}{a_p}$ by (3.5.1) and (3.5.3). Now applying (*) to type $t = \bar{p}\bar{q}$, we obtain $\Pr(\bar{p}\bar{q}|t) \geq \Pr(p\bar{q}|t)$, which by (3.5.1) and (3.5.3) implies $\frac{a_{\bar{p}}}{1-a_p} \geq \frac{\pi_p}{1-\pi_p}$. We have shown both inequalities relating to p. The two inequalities relating to q can be proved analogously.

B. Now suppose $\frac{a_{\bar{r}}}{1-a_r} \geq \frac{\pi_r}{1-\pi_r} \geq \frac{1-a_{\bar{r}}}{a_r}$ for each $r \in \{p, q\}$. We consider any type $t \in \mathcal{J}$ and have to show that the decision $d = t$ has maximal expected utility given t, or equivalently, that (*) holds.

We show (*) first in the case $t = pq$. Here, the inequality $\frac{\pi_p}{1-\pi_p} \geq \frac{1-a_{\bar{p}}}{a_p}$ implies $\Pr(pq|t) \geq \Pr(\bar{p}q|t)$ by (3.5.1) and (3.5.3), and it implies $\Pr(p\bar{q}|t) \geq \Pr(\bar{p}\bar{q}|t)$ by (3.5.2) and (3.5.4). Further, the inequality $\frac{\pi_q}{1-\pi_q} \geq \frac{1-a_{\bar{q}}}{a_q}$ implies $\Pr(pq|t) \geq \Pr(p\bar{q}|t)$ by (3.5.1) and (3.5.2). This shows (*) for $t = pq$.

Now we show (*) for the case $t = p\bar{q}$. As $\frac{\pi_p}{1-\pi_p} \geq \frac{1-a_{\bar{p}}}{a_p}$, we here have $\Pr(p\bar{q}|t) \geq \Pr(\bar{p}\bar{q}|t)$ by (3.5.1) and (3.5.3), and we have $\Pr(pq|t) \geq \Pr(\bar{p}q|t)$ by (3.5.2) and (3.5.4). As $\frac{a_{\bar{q}}}{1-a_q} \geq \frac{\pi_q}{1-\pi_q}$, we also have $\Pr(p\bar{q}|t) \geq \Pr(pq|t)$ by (3.5.1) and (3.5.2). This proves (*) for $t = p\bar{q}$.

By similar arguments, one shows (*) for $t = \bar{p}q$ and for $t = \bar{p}\bar{q}$. □

3.5.4 Consequentialist preferences: type 1

We begin by two lemmas, which are the counterparts of Lemmas 3.5.1 and 3.5.2 for the current preferences.

Lemma 3.5.3 *Assume consequentialist preferences of type 1. The expected utility of a decision $d \in \mathcal{J}$ is*

$$E(u(d, S)) = \begin{cases} \Pr(pq) + \Pr(\bar{p}\bar{q}) & \text{if } d \in \{pq, \bar{p}\bar{q}\} \\ \Pr(p\bar{q}) + \Pr(\bar{p}q) & \text{if } d \in \{p\bar{q}, \bar{p}q\}, \end{cases}$$

and the conditional expected utility of d given a type or a type profile is given by the analogous expression with conditional probabilities instead of unconditional ones.

Proof. The claim follows easily from the definition of the utility function. □

Lemma 3.5.4 *Assume consequentialist preferences of type 1. For each type profile $\mathbf{t} \in \mathcal{J}^n$ and decisions $d \in \{pq, \bar{p}\bar{q}\}$ and $d' \in \{p\bar{q}, \bar{p}q\}$*

$$E(u(d, S)|\mathbf{t}) > E(u(d', S)|\mathbf{t}) \Leftrightarrow [n_r^{\mathbf{t}} \geq k_r \text{ for both or no } r \in \{p, q\}].$$

Proof. Consider any $\mathbf{t} \in \mathcal{J}^n$, $d \in \{pq, \bar{p}\bar{q}\}$ and $d' \in \{p\bar{q}, \bar{p}q\}$. Define $g_r(k) := \pi_r a_r^k (1 - a_r)^{n-k}$ and $g_{\bar{r}}(k) := (1 - \pi_r)(1 - a_{\bar{r}})^k a_{\bar{r}}^{n-k}$ for all $r \in \{p, q\}$ and $k \in \mathbb{R}$. For each $r \in \{p, q\}$, the definition of k_r can now be rewritten as $k_r = \min\{k \in \{0, 1, ..., n + 1\} : g_r(k) > g_{\bar{r}}(k)\}$. So, (*) for each $k \in \{0, 1, ..., n + 1\}$, $k \geq k_r \Leftrightarrow$

$g_r(k) > g_{\bar{r}}(k)$. (Here, the implication '$\Rightarrow$' uses that $g_r(k)$ $[g_{\bar{r}}(k)]$ is strictly increasing [decreasing] in $k \in \mathbb{R}$.) Now,

$$\begin{array}{ll} & E(u(d,S)|\mathbf{t}) > E(u(d',S)|\mathbf{t}) \\ \Leftrightarrow & \Pr(pq|\mathbf{t}) + \Pr(\bar{p}\bar{q}|\mathbf{t}) > \Pr(p\bar{q}|\mathbf{t}) + \Pr(\bar{p}q|\mathbf{t}) \text{ by Lemma 3.5.3} \\ \Leftrightarrow & g_p(n_p^{\mathbf{t}})g_q(n_q^{\mathbf{t}}) + g_{\bar{p}}(n_p^{\mathbf{t}})g_{\bar{q}}(n_q^{\mathbf{t}}) > g_p(n_p^{\mathbf{t}})g_{\bar{q}}(n_q^{\mathbf{t}}) + g_{\bar{p}}(n_p^{\mathbf{t}})g_q(n_q^{\mathbf{t}}) \text{ by (3.5.5)-(3.5.8)} \\ \Leftrightarrow & \left[g_p(n_p^{\mathbf{t}}) - g_{\bar{p}}(n_p^{\mathbf{t}})\right]\left[g_q(n_q^{\mathbf{t}}) - g_{\bar{q}}(n_q^{\mathbf{t}})\right] > 0 \\ \Leftrightarrow & [n_r^{\mathbf{t}} \geq k_r \text{ for both or no } r \in \{p,q\}] \text{ by (*)}. \end{array}$$

□

We can now prove our two theorems about the present preferences.

Proof of Theorem 3.3.3. Consider a rule $f : \mathcal{J}^n \to \mathcal{J}$.

A. Assume f is the quota rule with thresholds k_p and k_q. Firstly, f is monotonic. Secondly, to show that informative voting is efficient, consider a given type profile $\mathbf{t} \in \mathcal{J}^n$. Supposing informative voting, the resulting voting profile is then $\mathbf{v} := \mathbf{t}$. We have to show that $d := f(\mathbf{v})$ is efficient for $\mathbf{t}$, i.e., that for each $d' \in \mathcal{J}$ with $\mathrm{Co}(d') \neq \mathrm{Co}(d)$ we have (*) $E(u(d,S)|\mathbf{t}) \geq E(u(d',S)|\mathbf{t})$. Consider any $d' \in \mathcal{J}$ with $\mathrm{Co}(d') \neq \mathrm{Co}(d)$. If $d = pq$, then $n_r^{\mathbf{t}} \geq k_r$ for both $r \in \{p,q\}$, implying (*) by Lemma 3.5.4. If $d = \bar{p}\bar{q}$, then $n_r^{\mathbf{t}} \geq k_r$ for *no* $r \in \{p,q\}$, again implying (*) by Lemma 3.5.4. Finally, if d is $\bar{p}q$ or $p\bar{q}$, then $n_r^{\mathbf{t}} \geq k_r$ for *exactly one* $r \in \{p,q\}$, so that (*) holds once again by Lemma 3.5.4.

B. Conversely, assume f is monotonic and makes informative voting efficient. We consider any $\mathbf{v} \in \mathcal{J}^n$ and must show that (**) $f_r(\mathbf{v}) = r \Leftrightarrow n_r^{\mathbf{v}} \geq k_r$ for each $r \in \{p,q\}$. As one can show using our non-degeneracy assumption,

$$k_r \notin \{0, n+1\} \text{ for some } r \in \{p,q\}; \tag{3.5.10}$$

for instance, if k_r were zero for each $r \in \{p,q\}$, then by Lemma 3.5.4 the decisions $\bar{p}q$ and $p\bar{q}$ would be inefficient for each type profile, violating non-degeneracy. We now prove (**) by distinguishing four cases.

Case 1: $n_r^{\mathbf{v}} \geq k_r$ for each $r \in \{p,q\}$. We must show that $f(\mathbf{v}) = pq$. Since the decision $f(\mathbf{v})$ is efficient for the type profile $\mathbf{t} = \mathbf{v}$, by Lemma 3.5.4, $f(\mathbf{v}) \in \{pq, \bar{p}\bar{q}\}$. Suppose for a contradiction $f(\mathbf{v}) = \bar{p}\bar{q}$. By (3.5.10), $k_r \geq 1$ for some $r \in \{p,q\}$. Suppose $k_p > 0$ (the case that $k_q > 0$ being analogous). Let $\mathbf{v}'$ be the voting profile obtained from $\mathbf{v}$ by replacing each occurring p by $\bar{p}$. By monotonicity, the decision is $f(\mathbf{v}') = \bar{p}\bar{q}$. By Lemma 3.5.4, for the type profile $\mathbf{t}' = \mathbf{v}'$ only $\bar{p}q$ and $p\bar{q}$ are efficient since $n_p^{\mathbf{t}'} = 0 < k_p$ and $n_q^{\mathbf{t}'} = n_q^{\mathbf{v}} \geq k_q$. So, the decision $f(\mathbf{v}')$ $(= \bar{p}\bar{q})$ is inefficient, a contradiction since f makes informative voting efficient.

Case 2: $n_p^{\mathbf{v}} \geq k_p$ and $n_q^{\mathbf{v}} < k_q$. We must show that $f(\mathbf{v}) = p\bar{q}$. By Lemma 3.5.4, $p\bar{q}$ and $\bar{p}q$ are both efficient for the type profile $\mathbf{t} = \mathbf{v}$. So, as informative voting is efficient, $f(\mathbf{v}) \in \{p\bar{q}, \bar{p}q\}$. Suppose for a contradiction $f(\mathbf{v}) = \bar{p}q$. By (3.5.10),

$k_p > 0$ or $k_q \leq n$. First, if $k_p > 0$, define $\mathbf{v}'$ as in Case 1. By monotonicity, the decision is $f(\mathbf{v}') = \bar{p}q$, which is inefficient for the type profile $\mathbf{t}' = \mathbf{v}'$ by Lemma 3.5.4 as $n_p^{\mathbf{t}'} = 0 < k_p$ and $n_q^{\mathbf{t}'} = n_q^{\mathbf{v}} < k_q$, a contradiction. Second, if $k_q \leq n$, define $\mathbf{v}'$ as the voting profile obtained from $\mathbf{v}$ by replacing each occurring $\bar{q}$ by q. By monotonicity, the decision is $f(\mathbf{v}') = \bar{p}q$, which is again inefficient for the type profile $\mathbf{t}' = \mathbf{v}'$ by Lemma 3.5.4 as $n_p^{\mathbf{t}'} = n_p^{\mathbf{v}} \geq k_p$ and $n_q^{\mathbf{t}'} = n \geq k_q$, a contradiction.

Case 3: $n_p^{\mathbf{v}} < k_p$ and $n_q^{\mathbf{v}} \geq k_q$. One can show that $f(\mathbf{v}) = \bar{p}q$ like in Case 2.

Case 4: $n_r^{\mathbf{v}} < k_r$ for each $r \in \{p, q\}$. We must show that $f(\mathbf{v}) = \bar{p}\bar{q}$. By informative voting being efficient and by Lemma 3.5.4 applied to $\mathbf{t} = \mathbf{v}$, $f(\mathbf{v}) \in \{pq, \bar{p}\bar{q}\}$. Suppose for a contradiction that $f(\mathbf{v}) = pq$. By (3.5.10), $k_r \leq n$ for some $r \in \{p, q\}$. We assume that $k_p \leq n$ (the proof being analogous if $k_q \leq n$). Let the voting profile $\mathbf{v}' \in \mathcal{J}^n$ arise from $\mathbf{v}$ by replacing each occurring $\bar{p}$ by p. By monotonicity, $f(\mathbf{v}') = pq$. This outcome is inefficient for the type profile $\mathbf{t}' = \mathbf{v}'$ by Lemma 3.5.4 and $n_p^{\mathbf{t}'} = n \geq k_p$ and $n_q^{\mathbf{t}'} = n_q^{\mathbf{v}} < k_q$. □

Proof of Theorem 3.4.2. A. First, let informative voting be sincere. Equivalently, for any type $t \in \mathcal{J}$, (*) $E(u(d, S)|t)$ is maximal at $d = t$. Using (*) with $t = pq$, we have $E(u(pq, S)|t) \geq E(u(\bar{p}q, S)|t)$, which by Lemma 3.5.3 is equivalent to $\Pr(pq|t) + \Pr(\bar{p}\bar{q}|t) \geq \Pr(p\bar{q}|t) + \Pr(\bar{p}q|t)$. Using (3.5.1)-(3.5.4), the latter is equivalent to

$$\frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} + \frac{1-a_{\bar{p}}}{a_p} \times \frac{1-a_{\bar{q}}}{a_q} \geq \frac{\pi_p}{1-\pi_p} \times \frac{1-a_{\bar{q}}}{a_q} + \frac{\pi_q}{1-\pi_q} \times \frac{1-a_{\bar{p}}}{a_p},$$

which can be rearranged as

$$\left(\frac{\pi_p}{1-\pi_p} - \frac{1-a_{\bar{p}}}{a_p}\right)\left(\frac{\pi_q}{1-\pi_q} - \frac{1-a_{\bar{q}}}{a_q}\right) \geq 0. \tag{3.5.11}$$

Analogously, using (*) three more times, with $t = p\bar{q}$, then $t = \bar{p}q$ and finally $t = \bar{p}\bar{q}$, we obtain

$$\begin{aligned}
\left(\frac{\pi_p}{1-\pi_p} - \frac{1-a_{\bar{p}}}{a_p}\right)\left(\frac{1-\pi_q}{\pi_q} - \frac{1-a_q}{a_{\bar{q}}}\right) &\geq 0 \quad (3.5.12)\\
\left(\frac{\pi_p}{1-\pi_p} - \frac{a_{\bar{p}}}{1-a_p}\right)\left(\frac{1-\pi_q}{\pi_q} - \frac{a_q}{1-a_{\bar{q}}}\right) &\geq 0 \quad (3.5.13)\\
\left(\frac{\pi_p}{1-\pi_p} - \frac{a_{\bar{p}}}{1-a_p}\right)\left(\frac{\pi_q}{1-\pi_q} - \frac{a_{\bar{q}}}{1-a_q}\right) &\geq 0. \quad (3.5.14)
\end{aligned}$$

Firstly, (i) $\frac{\pi_q}{1-\pi_q} \geq \frac{1-a_{\bar{q}}}{a_q}$, since otherwise by (3.5.11) we would get $\frac{\pi_p}{1-\pi_p} \leq \frac{1-a_{\bar{p}}}{a_p}$ (< 1), whereas by (3.5.13) we get $\frac{\pi_p}{1-\pi_p} \geq \frac{a_{\bar{p}}}{1-a_p}$ (> 1), a contradiction. Secondly, (ii) $\frac{\pi_p}{1-\pi_p} \geq \frac{1-a_{\bar{p}}}{a_p}$, because if (i) holds with a strict inequality, then (ii) follows from (3.5.11), whereas if (i) holds with equality, then $\frac{\pi_q}{1-\pi_q} < 1 < \frac{a_{\bar{q}}}{1-a_q}$, which together with (3.5.12) implies (ii). We finally show that (iii) $\frac{\pi_p}{1-\pi_p} \leq \frac{a_{\bar{p}}}{1-a_p}$ and (iv) $\frac{\pi_q}{1-\pi_q} \leq \frac{a_{\bar{q}}}{1-a_q}$. First, suppose (ii) holds with equality. Then $\frac{\pi_p}{1-\pi_p} < 1 < \frac{a_{\bar{p}}}{1-a_p}$, which implies (iii), and with (3.5.14) also implies (iv). Second, suppose (ii) holds with a strict inequality. Then with (3.5.12) we get (iv). If (iv) holds with a strict inequality, then we get (iii) by (3.5.14), while if (iv) holds with equality, then $\frac{1-\pi_q}{\pi_q} = \frac{1-a_q}{a_{\bar{q}}} < 1 < \frac{a_q}{1-a_{\bar{q}}}$, which by (3.5.13) implies (iii).

B. Conversely, assume $\frac{a_{\bar{r}}}{1-a_r} \geq \frac{\pi_r}{1-\pi_r} \geq \frac{1-a_{\bar{r}}}{a_r}$ for each $r \in \{p, q\}$. We have to show that informative voting is sincere, i.e., that (*) holds for each type $t \in \mathcal{J}$. As one can check, the inequalities (3.5.11)-(3.5.14) all hold. These inequalities imply that (*) holds for each type $t \in \mathcal{J}$. For instance, as shown in part A, (3.5.11) reduces to $E(u(pq, S)|t) \geq E(u(\bar{p}q, S)|t)$ for $t = pq$. □

3.5.5 Consequentialist preferences: type 2

We begin by a simple lemma, the counterpart of Lemmas 3.5.1 and 3.5.3.

Lemma 3.5.5 *Assume consequentialist preferences of type 2. The expected utility of a decision $d \in \mathcal{J}$ is*

$$E(u(d,S)) = \begin{cases} \Pr(pq) & \text{if } d = pq \\ 1 - \Pr(pq) & \text{if } d \neq pq, \end{cases}$$

and the conditional expected utility of d given a type or a type profile is given by the analogous expression with conditional probabilities instead of unconditional ones.

Proof. The claim follows from the specification of the utility function. □

We now prove our results about the current preferences. Some proofs implicitly extend $\beta(k, l)$ to values of $k.l$ not in $\{0, ..., n\}$, using the expression (3.3.3).

Proof of proposition 3.3.1. The claim can easily be shown by elaborating the informal argument given in the text. □

Proof of theorem 3.3.4. A. First, suppose $f : \mathcal{J}^n \to \mathcal{J}$ is a quota rule with thresholds m_p and m_q making informative voting efficient. The following claims must be shown.

Claim 1: $m_p = l_p$ and $m_q = l_q$.

Consider a type profile $\mathbf{t} \in \mathcal{J}^n$ for which $n_p^{\mathbf{t}} = n$ and $n_q^{\mathbf{t}} = l_q$. Assuming informative voting, the resulting voting profile is $\mathbf{v} = \mathbf{t}$. By definition of l_q, $\beta(n, l_q) > 1/2$. So $f(\mathbf{v}) = pq$ by Proposition 3.3.1. Thus, $l_q \geq m_q$ by definition of f. One analogously shows that $l_p \geq m_p$. To show the converse inequalities, consider a voting profile $\mathbf{v} \in \mathcal{J}^n$ for which $n_p^{\mathbf{v}} = m_p$ and $n_q^{\mathbf{v}} = n$ $(\geq m_q)$. The resulting decision is $f(\mathbf{v}) = pq$ by definition of f. So, by Proposition 3.3.1, $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) = \beta(m_p, n) > 1/2$. Hence, $m_p \geq l_p$ by definition of l_p. Analogously, one shows that $m_q \geq l_q$.

Claim 2: $\beta(l_p, l_q) > 1/2$.

For any voting profile $\mathbf{v} \in \mathcal{J}^n$ for which $n_p^{\mathbf{v}} = l_p$ $(= m_p)$ and $n_q^{\mathbf{v}} = l_q$ $(= m_q)$, we have $f(\mathbf{v}) = pq$ by definition of f, so that by Proposition 3.3.1 $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$, i.e., $\beta(l_p, l_q) > 1/2$.

B. Conversely, assume $\beta(l_p, l_q) > 1/2$. We show that the quota rule f with thresholds l_p and l_q makes informative voting efficient. We first prove that for all $k, l \in \{0, ..., n\}$,

$$\beta(k, l) > 1/2 \Leftrightarrow [k \geq l_p \text{ and } l \geq l_q]. \tag{3.5.15}$$

Let $k, l \in \{0, ..., n\}$. If $k \geq l_p$ and $l \geq l_q$, then $\beta(k, l) \geq \beta(l_p, l_q) > 1/2$, where the first inequality holds because β is increasing in each argument. If $k < l_p$, then $\beta(k, l) \leq \beta(k, n) \leq 1/2$, where the last inequality holds by definition of l_p $(> k)$. Analogously, if $l \leq l_q$, then $\beta(k, l) \leq 1/2$.

Now consider any type profile $\mathbf{t} \in \mathcal{J}^n$. Assuming informative voting, the resulting voting profile is $\mathbf{v} = \mathbf{t}$. We have to show that the decision $f(\mathbf{v})$ is efficient for $\mathbf{t}$ $(= \mathbf{v})$. First, if $n_p^{\mathbf{t}} \geq l_p$ and $n_q^{\mathbf{t}} \geq l_q$, the decision is $f(\mathbf{v}) = pq$, which is efficient by Proposition 3.3.1 since $\beta(n_p^{\mathbf{t}}, n_q^{\mathbf{t}}) > 1/2$ by (3.5.15). Second, if $n_p^{\mathbf{t}} < l_p$ or $n_q^{\mathbf{t}} < l_q$, the resulting decision $f(\mathbf{v})$ is in $\{\bar{p}q, p\bar{q}, \bar{p}\bar{q}\}$, which is efficient by Proposition 3.3.1 since $\beta(n_p^{\mathbf{t}}, n_q^{\mathbf{t}}) \leq 1/2$ by (3.5.15). □

Proof of theorem 3.3.5. Consider a rule $f : \mathcal{J}^n \to \mathcal{J}$. We repeatedly draw on the fact that (*) $\beta(k, l)$ is strictly increasing in each argument.

A. First, assume f is defined by (3.3.6) for thresholds m_p and m_q satisfying $\beta(m_p, l_q)$, $\beta(l_p, m_q) > 1/2$. Clearly, f is anonymous. To show that informative voting is efficient, it suffices by Proposition 3.3.1 to prove that for all $\mathbf{v} \in \mathcal{J}^n$,

$$f(\mathbf{v}) = pq \Leftrightarrow \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2. \tag{3.5.16}$$

If $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$, then clearly $f(\mathbf{v}) = pq$ by (3.3.6). Conversely, assume $f(\mathbf{v}) = pq$. Then, by definition of f, either $n_r^{\mathbf{v}} \geq m_r$ for each $r \in \{p, q\}$, or $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$. In the second case, we are done. Now assume the first case. Since $\beta(m_p, l_q) > 1/2$, we have $\beta(m_p, n) > 1/2$ by (*), whence $m_p \geq l_p$ by definition of l_p. Using (*) and that $n_p^{\mathbf{v}} \geq m_p \geq l_p$ and $n_q^{\mathbf{v}} \geq m_q$, we have $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \geq \beta(l_p, m_q)$. Moreover, $\beta(l_p, m_q) > 1/2$ by definition of m_q. So, $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$, which completes the proof of (3.5.16).

It remains to show monotonicity of f. Take two voting profiles $\mathbf{v}, \mathbf{v}' \in \mathcal{J}^n$ such that for all $r \in f(\mathbf{v})$, the voters who vote for r in $\mathbf{v}$ also vote for r in $\mathbf{v}'$.

Case 1: $f(\mathbf{v}) = pq$. Then, $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$ by (3.5.16). Also, $n_p^{\mathbf{v}'} \geq n_p^{\mathbf{v}}$ and $n_q^{\mathbf{v}'} \geq n_q^{\mathbf{v}}$, so that $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) \geq \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}})$ by (*). It follows that $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) > 1/2$, so that $f(\mathbf{v}') = pq$ by (3.5.16).

Case 2: $f(\mathbf{v}) = p\bar{q}$. We have to show that $f(\mathbf{v}') = p\bar{q}$, i.e., that

$$n_p^{\mathbf{v}'} \geq m_p,\ n_q^{\mathbf{v}'} < m_q, \text{ and } \beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) \leq 1/2. \tag{3.5.17}$$

Since $f(\mathbf{v}) = p\bar{q}$, the definition of f implies $n_p^{\mathbf{v}} \geq m_p$ and $n_q^{\mathbf{v}} < m_q$, and the definition of $\mathbf{v}'$ implies $n_p^{\mathbf{v}'} \geq n_p^{\mathbf{v}}$ and $n_q^{\mathbf{v}'} \leq n_q^{\mathbf{v}}$; hence, the first two inequalities in (3.5.17) hold. As $\beta(m_p, l_q) > 1/2$ and $n_p^{\mathbf{v}} \geq m_p$, we have $\beta(n_p^{\mathbf{v}}, l_q) > 1/2$ by (*). Also, since $f(\mathbf{v}) = p\bar{q}$, we have $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$ by (3.5.16). Hence, $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) < \beta(n_p^{\mathbf{v}}, l_q)$. So, $n_q^{\mathbf{v}} < l_q$ by (*), whence $n_q^{\mathbf{v}'} < l_q$ as $n_q^{\mathbf{v}'} \leq n_q^{\mathbf{v}}$. Thus, by definition of l_q, $\beta(n, n_q^{\mathbf{v}'}) \leq 1/2$, so that $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) \leq 1/2$ by (*), proving (3.5.17).

Case 3: $f(\mathbf{v}) = \bar{p}q$. One can show that $f(\mathbf{v}') = \bar{p}q$ analogously to Case 2.

Case 4: $f(\mathbf{v}) = \bar{p}\bar{q}$. Then, $n_p^{\mathbf{v}} < m_p$, $n_q^{\mathbf{v}} < m_q$, and $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$. We have to show that $f(\mathbf{v}') = \bar{p}\bar{q}$, i.e., that these three inequalities still hold if $\mathbf{v}$ is replaced by $\mathbf{v}'$. This follows from the fact that $n_p^{\mathbf{v}'} \leq n_p^{\mathbf{v}}$ and $n_q^{\mathbf{v}'} \leq n_q^{\mathbf{v}}$ (by definition of $\mathbf{v}'$) and from (*).

B. Conversely, let f be monotonic and anonymous, and make informative voting efficient. For each $r \in \{p, q\}$, define

$$m_r := \min\{n_r^{\mathbf{v}} : \mathbf{v} \in \mathcal{J}^n \text{ such that } f_r(\mathbf{v}) = r \text{ and } \beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2\},$$

where this minimum is interpreted as $n + 1$ if it is taken over an empty set. We prove that f has the required form with respect to the so-defined thresholds m_p and m_q. The proof proceeds in several steps and is completed by Claims 5 and 6 below.

Claim 1: For all $\mathbf{v} \in \mathcal{J}^n$, if $n_p^{\mathbf{v}} \geq l_p$, $n_q^{\mathbf{v}} \geq l_q$ and $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$, then $f(\mathbf{v}) = \bar{p}\bar{q}$.

Let $\mathbf{v} \in \mathcal{J}^n$ satisfy the antecedent condition. First assume $f(\mathbf{v}) = p\bar{q}$ for a contradiction. Let $\mathbf{v}'$ be the voting profile obtained from $\mathbf{v}$ by replacing each $\bar{p}$ by p. By monotonicity, $f(\mathbf{v}') = p\bar{q}$. However, Proposition 3.3.1 implies that $f(\mathbf{v}') = pq$, since $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) = \beta(n, n_q^{\mathbf{v}}) \geq \beta(n, l_q) > 1/2$ (where the first inequality holds by $n_q^{\mathbf{v}} \geq l_q$, and the second by definition of l_q). This contradiction proves that $f(\mathbf{v}) \neq p\bar{q}$. One similarly proves that $f(\mathbf{v}) \neq \bar{p}q$. So, as $f(\mathbf{v}) \in \{p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$ by Proposition 3.3.1, we have $f(\mathbf{v}) = \bar{p}\bar{q}$, proving the claim.

Claim 2: For all $\mathbf{v} \in \mathcal{J}^n$, if $n_p^{\mathbf{v}} \leq l_p$, $n_q^{\mathbf{v}} \leq l_q$ and $\beta(l_p, l_q) \leq 1/2$, then $f(\mathbf{v}) = \bar{p}\bar{q}$.

Consider any $\mathbf{v} \in \mathcal{J}^n$ satisfying the antecedent condition. Let $\mathbf{w} \in \mathcal{J}^n$ arise from $\mathbf{v}$ by replacing $l_p - n_p^{\mathbf{v}}$ occurrences of $\bar{p}$ by p, $l_q - n_q^{\mathbf{v}}$ occurrences of $\bar{q}$ by q. Note that $n_p^{\mathbf{w}} = l_p$ and $n_q^{\mathbf{w}} = l_q$, whence by Claim 1 $f(\mathbf{w}) = \bar{p}\bar{q}$. By monotonicity, it follows that $f(\mathbf{v}) = \bar{p}\bar{q}$.

Claim 3: $m_p \geq l_p$ and $m_q \geq l_q$.

Suppose for a contradiction $m_p < l_p$. By definition of m_p, there is a $\mathbf{v} \in \mathcal{J}^n$ such that $m_p = n_p^{\mathbf{v}}$, $f_p(\mathbf{v}) = p$ and $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$. As by Proposition 3.3.1,

$f(\mathbf{v}) \in \{\bar{p}q, p\bar{q}, \bar{p}\bar{q}\}$, it follows that $f(\mathbf{v}) = p\bar{q}$. We consider two cases.

Case 1: $n_q^{\mathbf{v}} \geq l_q$. Let $\mathbf{v}' \in \mathcal{J}^n$ be the voting profile arising from $\mathbf{v}$ by replacing each $\bar{p}$ by p. By monotonicity, the resulting decision is $f(\mathbf{v}') = p\bar{q}$. But $f(\mathbf{v}') = pq$ by Proposition 3.3.1 as $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) = \beta(n, n_q^{\mathbf{v}}) \geq \beta(n, l_q) > 1/2$ (where the first inequality holds by $n_q^{\mathbf{v}} > l_q$ and the second by definition of l_q).

Case 2: $n_q^{\mathbf{v}} < l_q$. Then by Claim 2 $f(\mathbf{v}) = \bar{p}\bar{q}$, a contradiction since $f(\mathbf{v}) = p\bar{q}$.

We have shown one inequality of Claim 3; the other one has an analogous proof.

Claim 4: For all $\mathbf{v} \in \mathcal{J}^n$ with $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$, if $n_p^{\mathbf{v}} \geq m_p$ then $f(\mathbf{v}) = p\bar{q}$, and if $n_q^{\mathbf{v}} \geq m_q$ then $f(\mathbf{v}) = \bar{p}q$.

Consider any $\mathbf{v} \in \mathcal{J}^n$ with $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$. Suppose for a contradiction that $n_p^{\mathbf{v}} \geq m_p$ but $f(\mathbf{v}) \neq p\bar{q}$. Then, as by Proposition 3.3.1 $f(\mathbf{v}) \in \{\bar{p}q, p\bar{q}, \bar{p}\bar{q}\}$, either $f(\mathbf{v}) = \bar{p}q$ or $f(\mathbf{v}) = \bar{p}\bar{q}$.

Case 1: $f(\mathbf{v}) = \bar{p}q$. Let $\mathbf{v}' \in \mathcal{J}^n$ be the voting profile arising from $\mathbf{v}$ by replacing each $\bar{q}$ by q. By monotonicity, the resulting decision is $f(\mathbf{v}') = \bar{p}q$, whereas by Proposition 3.3.1 $f(\mathbf{v}') = pq$ because $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) = \beta(n_p^{\mathbf{v}}, n) \geq \beta(l_p, n) > 1/2$, where the first inequality holds because $n_p^{\mathbf{v}} \geq l_p$ (by Claim 3) and the second inequality holds by definition of l_p.

Case 2: $f(\mathbf{v}) = \bar{p}\bar{q}$. By definition of m_p there is a $\mathbf{w} \in \mathcal{J}^n$ such that $n_p^{\mathbf{w}} = m_p$, $f_p(\mathbf{w}) = p$ and $\beta(n_p^{\mathbf{w}}, n_q^{\mathbf{w}}) \leq 1/2$. As by Proposition 3.3.1 $f(\mathbf{w}) \in \{p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$, it follows that $f(\mathbf{w}) = p\bar{q}$. Let $\mathbf{v}'$ $[\mathbf{w}']$ be the voting profile arising from $\mathbf{v}$ $[\mathbf{w}]$ by replacing each q by $\bar{q}$. By monotonicity, $f(\mathbf{v}') = \bar{p}\bar{q}$ and $f(\mathbf{w}') = p\bar{q}$. Now let $\mathbf{w}''$ be a voting profile arising from $\mathbf{w}'$ by replacing $n_p^{\mathbf{v}'} - n_p^{\mathbf{w}'}$ $(= n_p^{\mathbf{v}} - m_p \geq 0)$ occurrences of $\bar{p}$ by p. By monotonicity, $f(\mathbf{w}'') = p\bar{q}$. So, $f(\mathbf{w}'') \neq f(\mathbf{v}')$, a contradiction by anonymity since $\mathbf{w}''$ is a permutation of $\mathbf{v}'$.

This shows the first implication in Claim 4. The second one can be shown similarly.

Claim 5: $\beta(m_p, l_q), \beta(l_p, m_q) > 1/2$.

We only show that $\beta(m_p, l_q) > 1/2$; the other inequality is analogous. Suppose for a contradiction that $\beta(m_p, l_q) \leq 1/2$. So, since $\beta(n+1, l_q) > \beta(n, l_q) > 1/2$ (by definition of l_q), we have $m_p \neq n+1$. Hence, there is a $\mathbf{v} \in \mathcal{J}^n$ such that $n_p^{\mathbf{v}} = m_p$ and $n_q^{\mathbf{v}} = l_q$. By Claim 4, $f(\mathbf{v}) = p\bar{q}$. Let $\mathbf{v}'$ be the voting profile arising from $\mathbf{v}$ by replacing each $\bar{p}$ by p. By monotonicity, $f(\mathbf{v}') = p\bar{q}$, a contradiction since by Proposition 3.3.1 $f(\mathbf{v}') = pq$ since $\beta(n_p^{\mathbf{v}'}, n_q^{\mathbf{v}'}) = \beta(n, l_q) > 1/2$.

Claim 6: f is given by (3.3.6).

Consider any $\mathbf{v} \in \mathcal{J}^n$ and $r \in \{p, q\}$. We show the equivalence (3.3.6) by distinguishing different cases. If $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) > 1/2$, then $f(\mathbf{v}) = pq$ by Proposition 3.3.1, implying (3.3.6). If $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$ and $n_r^{\mathbf{v}} \geq m_r$, then (3.3.6) holds by Claim 4. Finally, if $\beta(n_p^{\mathbf{v}}, n_q^{\mathbf{v}}) \leq 1/2$ and $n_r^{\mathbf{v}} < m_r$, then $f_r(\mathbf{v}) \neq r$ by definition of m_r, whence (3.3.6) again holds. $\square$

Proof of Theorem 3.4.3. A. First, suppose informative voting is sincere. Equivalently, for any given type $t \in \mathcal{J}$, the decision $d = t$ has maximal conditional expected utility, i.e., (*) $E(u(d,S)|t)$ is maximal at $d = t$. Applying (*) with $t = pq$, we have $E(u(pq,S)|t) \geq E(u(\bar{p}\bar{q},S)|t)$, which by Lemma 3.5.5 reduces to $\Pr(pq|t) \geq 1 - \Pr(pq|t)$, i.e., to $\Pr(pq|t) \geq 1/2$. Using (3.5.1), one derives that $\frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} \geq C$. Now applying (*) with $t = p\bar{q}$, we have $E(u(p\bar{q},S)|t) \geq E(u(pq,S)|t)$, which by Lemma 3.5.5 reduces to $1 - \Pr(pq|t) \geq \Pr(pq|t)$, so that $\Pr(pq|t) \leq 1/2$. Using (3.5.2), one obtains $\frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} \leq A$. Finally, applying (*) with $t = \bar{p}q$, we have $E(u(\bar{p}q,S)|t) \geq E(u(pq,S)|t)$, which by Lemma 3.5.5 reduces to $1 - \Pr(pq|t) \geq \Pr(pq|t)$, whence $\Pr(pq|t) \leq 1/2$. Using (3.5.3), one derives $\frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} \leq B$. This proves all inequalities.

B. Conversely, suppose $A, B \geq \frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} \geq C$. For each given type $t \in \mathcal{J}$, one has to show (*). As the reader can verify using Lemma 3.5.5 and (3.5.1)-(3.5.4), if $t = pq$ then (*) follows from $\frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q} \geq C$; if $t = p\bar{q}$ then (*) follows from $A \geq \frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q}$; if $t = \bar{p}q$ then (*) follows from $B \geq \frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q}$; and if $t = \bar{p}\bar{q}$ then (*) can be derived from $A \geq \frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q}$ or from $B \geq \frac{\pi_p}{1-\pi_p} \times \frac{\pi_q}{1-\pi_q}$. □

Chapter 4

Judgment aggregation in search for the truth: the case of interconnections

4.1 Introduction

The theory of judgment aggregation deals with situations where a group needs to make a collective 'yes' or 'no' judgment on several issues on the basis of group members' judgments on these issues. Many decision making problems in real life involve multiple issues. The jury in a court trial might need to form judgments on whether the defendant has broken the contract, and whether the contract is legally valid. The city council might need to reach judgments on whether the CO_2 level in the city is above the critical threshold and whether the chemical plant in the city should be closed down. In such problems, the issues on group's agenda might be mutually interconnected, in that the judgment made on one issue might constrain the judgment on another issue. In the city commission example, the CO_2 level being judged to be above the critical threshold might restrict the judgment on the second issue to 'yes'; i.e., lead to the closing of the chemical factory. Judgment aggregation models allow for the study of a wide range of realistic collective decision making problems.

When it comes to aggregate judgments in decision making bodies like juries or city councils, it seems natural to have epistemic concerns. Such problems are different than problems where individuals have conflicting aims. In the court trial example, the jury's problem is to find out two independent facts, whether the defendant has broken the contract and whether the contract is legally valid. It seems that the primary goal is reaching the truth in such problems. The epistemic approach in judgment aggregation aims to reach true group judgments. According to the classical social-choice theoretic approach in judgment aggregation – where voters are taken to have conflicting aims – a good voting rule should be fair to jurors while according to the epistemic approach, a good voting rule should track the truth. Whether a voting rule is good in tracking the truth or not depends on questions such as whether the individuals judgments are sufficient evidence for the truth value of each issue, and whether individuals have expressed their judgments truthfully.

This chapter assumes that the group faces *two* issues, and a 'no' answer to both issues is *inconsistent*. Consistency is a property that requires the collective decision to be free from any logical contradictions. In the city council example, the inconsistency arises in case of a 'yes' judgment on the first issue (CO_2 level is above the threshold) and a 'no' judgment on the second issue (the chemical plant should not be closed down). Note that by exchanging the roles of issues with their opposites, we can obtain every kind of interconnection between issues. We assume that voters share the common goal of tracking the truth but each has private information regarding the truth value of each issue. In this setting we want to answer the following question: Which voting rules lead to efficient and truthful Bayesian Nash equilibrium of the corresponding game? So, we want to design voting rules which first lead to truthful revelation of private information, and second lead to the efficiency in equilibrium. Note that individual reporting of private information need not always be truthful, even when voters have no conflicting aims. As Austen-Smith and Banks (1996) show, if a voter conditions her beliefs on being pivotal – on being able to change the outcome – she may not always find it best to report truthfully. The question of consistency here arises when one wants to use quota rules (where separate votes are taken on each proposition using acceptance thresholds) which are practical and common. This chapter also examines the possibility of truth-tracking in particular with quota rules.

The epistemic perspective with strategic concerns is well-established in a different body of the literature, which studies single issue problems (Austen-Smith and Banks 1996, Feddersen and Pesendorfer 1997, 1998). There have been few works taking the epistemic approach in judgment aggregation. The work by Bozbay, Dietrich and Peters (2011) studies judgment aggregation from the epistemic and strategic voting perspective and to our knowledge is the only work that deals with the mechanism design problem in the multi-issue case. Their model is similar in that they model voters' common interests and private information, however with an agenda with independent issues. For an extended survey of the two bodies of literature to which this chapter connects to, please see Section 3.1.1.

The chapter proceeds as follows. Two questions are answered in Section 4.2 and Section 4.3; respectively, when informative voting is efficient, and when informative voting is efficient by quota rules in particular. By informative voting, we generally mean 'following the evidence' whenever the evidence is consistent with the true state.[1] Section 4.2.1 presents the model, in which a group of voters is to decide whether to accept or reject each of two propositions while they can not reject both at the same time. The model presented in Section 4.3.1 differs from the model in Section 4.2.1 in that the collective decision might be inconsistent; hence, both propositions can be collectively rejected. In both models, voters hold truth-tracking preferences and they vote on the basis of private information which may possibly be inconsistent. Section 4.2.2, 4.2.3 and 4.2.4 address the key question of how to design the voting rule such that it leads to efficient decisions as well as simple-minded, truthful voting behaviour in equilibrium. The answer depends on both the kind of utility function in use and the definition of simple-minded behaviour. It turns out that in many situations such a voting rule does not exist. The necessary and sufficient conditions for the existence of such rules are also given and these rules are characterized by some properties. Section 4.3.2 addresses the possibility of efficient information aggregation with quota rules and provides an impossibility result. All proofs are in appendix.

4.2 Efficient information aggregation

4.2.1 The Model

A simple judgment aggregation problem

We consider a group of voters, labeled $i = 1, ..., n$, where $n \geq 2$. There are two propositions p and q, and their negations $\bar{p}$ and $\bar{q}$. The group of voters wants to obtain a collective judgment on whether p or $\bar{p}$ is true, and whether q or $\bar{q}$ is true. While doing so, voters know that the combination $\{\bar{p}, \bar{q}\}$ is not possible. The three possible judgment sets are $\{p, q\}, \{p, \bar{q}\}, \{\bar{p}, q\}$, abbreviated by pq, $p\bar{q}$ and $\bar{p}q$, respectively[2]. Each voter votes for a judgment set in $\mathcal{J} = \{pq, p\bar{q}, \bar{p}q\}$. A collective decision is taken using a voting rule. A voting rule is defined as a function $f : \mathcal{J}^n \to \mathcal{J}$, which maps each voting profile $\mathbf{v} = (v_1, ..., v_n)$ to a decision $d \equiv f(\mathbf{v})$. Some salient properties of voting rules are defined below:

- *Anonymity:* For all voting profiles $(v_1, ..., v_n) \in \mathcal{J}^n$ and all permutations $(i_1, ..., i_n)$ of the voters, $f(v_{i_1}, ..., v_{i_n}) = f(v_1, ..., v_n)$. Informally, the voters are treated equally.

- *Monotonicity:* For all voting profiles $\mathbf{v}, \mathbf{v}' \in \mathcal{J}^n$, if for each r in $f(\mathbf{v})$ the voters who accept r in $\mathbf{v}$ also accept r in $\mathbf{v}'$, then $f(\mathbf{v}') = f(\mathbf{v})$. Informally,

[1] An informative voter reveals her private information in her vote whenever the private information is non-conflicting with the true state. Two kinds of informative behaviour is analysed in this work.

[2] Similarly, $\{\bar{p}, \bar{q}\}$ is abbreviated by $\bar{p}\bar{q}$.

additional support for the collectively accepted propositions never reverses the collective acceptance of these propositions.

- *Independence:* The decision on each proposition $r \in \{p, q\}$ only depends on the votes on r. Informally, the group in effect takes two separate votes, one between p and $\bar{p}$ and one between q and $\bar{q}$.

To define the next property, we introduce some notation. Given a voting profile $\mathbf{v} = (v_1, ..., v_n)$, for each $r \in \{p, q\}$ let $\mathbf{v}_r := (v_{1r}, ..., v_{nr})$ be the vector with entities defined as follows: for $i = 1, ..., n$, $v_{ir} = 1$ if v_i contains r and $v_{ir} = 0$ if v_i contains $\bar{r}$.

- *Neutrality:* For every voting profile $\mathbf{v}$ and every voting profile $\mathbf{v}'$ for which there is no permutation $(i_1, ..., i_n)$ of the voters with $(v_{i_1}, ..., v_{i_n}) = (v'_1, ..., v'_n)$, if $\mathbf{v}_r = \mathbf{v}'_{r'}$ for each $r, r' \in \{p, q\}$ with $r \neq r'$, then f accepts r in $\mathbf{v}$ if and only if f accepts r' in $\mathbf{v}'$. Informally, if two voting profiles have the exact same acceptance regime between different propositions, so do the decisions.

A common preference for true collective judgments

There is one 'correct' judgment set in $\mathcal{J}$, which we call the *state (of the world)*, denoted by s. The state is unobservable by voters. Voters have identical preferences, captured by a common utility function $u : \mathcal{J} \times \mathcal{J} \to \mathbb{R}$ which maps any decision-state pair (d, s) to its utility $u(d, s)$. The notion of truth-tracking requires the utility to be high if the decision is correct, but details matter. We focus on two natural kinds of preferences:

Simple preferences. The utility function is given by

$$u(d, s) = \begin{cases} 1 & \text{if } d = s \text{ (correct decision)} \\ 0 & \text{if } d \neq s \text{ (incorrect decision).} \end{cases} \tag{4.2.1}$$

Simple preferences are the simplest candidate for truth-tracking preferences.[3]

Consequentialist preferences. Here, we assume that the decision leads to one of two possible consequences which represents group actions.[4] This is captured by a *consequence function* Co which maps the set of possible decisions $\mathcal{J}$ to a two-element set of possible consequences. Consider a market with only one firm, Firm A. Firms B and C are interested in entering the market and Firm B has higher capacity than Firm C. The executive board of firm C is to make judgments on whether or

[3]A voter tracks the truth on a proposition p if the following is true: if p were true, the voter would accept p and if p were false, the agent would accept $\bar{p}$ (Nozick, 1981).

[4]This two-consequence situation corresponds to problems where the group action is represented by a third proposition – conclusion proposition – which might be true or false. Judging it to be true leads to one of the actions/consequences while judging it to be false leads to the other.

not firm A will expand its capacity (p) and firm B will enter the market (q). While doing so, the board members know that if Firm A does not increase capacity, then Firm B will enter the market; hence, $\bar{p} \to q$. If both issues are judged to be true, then the consequence is 'no market entry' (Co(pq) ='no market entry'), while if only one of the issues is judged to be true, the consequence is 'market entry' (Co($p\bar{q}$) = Co($\bar{p}q$) ='market entry').[5] It turns out that this consequence function with the property Co(pq) $\neq$ Co($p\bar{q}$) = Co($\bar{p}q$) is the only interesting consequence function up to isomorphism. (See Section 2.4 for further discussion.) The consequentialist utility function is given by

$$u(d,s) = \begin{cases} 1 & \text{if } \mathrm{Co}(d) = \mathrm{Co}(s) \text{ (correct consequence)} \\ 0 & \text{if } \mathrm{Co}(d) \neq \mathrm{Co}(s) \text{ (incorrect consequence).} \end{cases} \tag{4.2.2}$$

Private information and strategies

Each voter has a *type*, which represents private information or evidence about whether p is true and information about whether q is true. A voter's type takes the form of an element of $\mathcal{T} := \{pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$, generically denoted by t. For instance, the type $t = p\bar{q}$ represents evidence for p and for $\bar{q}$, and the type $t = \bar{p}\bar{q}$ represents evidence for $\bar{p}$ and for $\bar{q}$, which is conflicting information since $\bar{p}\bar{q} \notin \mathcal{J}$. We write $\mathbf{t} = (t_1, ..., t_n) \in \mathcal{T}^n$ for a profile of voters' types.

Nature draws a state-types combination $(s, \mathbf{t})$ in $\mathcal{J} \times \mathcal{T}^n$ according to a probability measure denoted Pr. When a proposition r in $\{p, \bar{p}, q, \bar{q}\}$ represents (part of) voter i's type rather than (part of) the true state, we often write r_i for r. For instance, $\Pr(p_i|p)$ is the probability that voter i has evidence for p given that p is true. By convention, the prior probability of state $s \in \mathcal{J}$ is denoted

$$\pi_s = \Pr(s)$$

and is assumed to be in the interval $(0, 1)$. The probability of getting evidence for r given that r is true is denoted

$$a_r = \Pr(r_i|r)$$

and by assumption belongs to $(1/2, 1)$ and does not depend on the voter i.

By assumption, voters' types are independent given the state. Moreover, given the truth about p (i.e., either p or $\bar{p}$), a voter's evidence about p (i.e., either p_i or $\bar{p}_i$) is independent of the truth and the evidence about q; and similarly, given the truth about q, a voter's evidence about q is independent of the truth and the evidence about p. These independence assumptions allow one to express the joint distribution

[5] There is still demand left for Firm C if only one of the companies is in the market.

of the state and the types:

$$\Pr(s, \mathbf{t}) = \Pr(s) \times \prod_{i=1}^{n} \Pr(t_i|s).$$

Here, $\Pr(s) = \pi_s$, and the term $\Pr(t_i|s)$ is also expressible in terms of our parameters; for instance,

$$\begin{aligned}
\Pr(p_i q_i|pq) &= \Pr(p_i|pq)\Pr(q_i|pq, p_i) = \Pr(p_i|p)\Pr(q_i|q) = a_p a_q. \\
\Pr(p_i \bar{q}_i|pq) &= \Pr(p_i|pq)\Pr(\bar{q}_i|pq, p_i) = \Pr(p_i|p)\Pr(\bar{q}_i|q) = a_p(1 - a_q).
\end{aligned}$$

Each voter submits a vote in $\mathcal{J}$ based on his type. A *(voting) strategy* is a function $\sigma : \mathcal{T} \to \mathcal{J}$, mapping each type $t \in \mathcal{T}$ to the type's vote $v = \sigma(t)$. We write $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ for a profile of voters' strategies. Together with a voting rule f and a common utility function u, we now have a well-defined Bayesian game.

For a given type profile $\mathbf{t} \in \mathcal{T}^n$, we call a decision $d \in \mathcal{J}$ *efficient* if it has maximal expected utility conditional on the full information $\mathbf{t}$. We adapt some common notions of voting behaviour to this framework.

- A strategy σ of a voter is *mostly informative* if $\sigma(t) = t$ for all $t \in \mathcal{T} \setminus \{\bar{p}\bar{q}\}$.
- A strategy σ of a voter is *informative* if $\sigma(t) = t$ for all $t \in \mathcal{T} \setminus \{\bar{p}\bar{q}\}$ and $\sigma(\bar{p}\bar{q}) \in \{p\bar{q}, \bar{p}q\}$.
- A strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ is *rational* if each strategy is a best response to the other strategies, i.e., if the profile is a Nash equilibrium of the corresponding Bayesian game. Hence, each voter maximises the expected utility of the collective decision given the strategies of the other voters.
- A strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ is *efficient* if for every type profile $\mathbf{t} = (t_1, ..., t_n)$ the resulting decision $d = f(\sigma_1(t_1), ..., \sigma_n(t_n))$ is efficient (i.e., has maximal expected utility conditional on full information $\mathbf{t}$). Hence, all the information spread across the group is used efficiently: the collective decision is no worse than a decision of a (virtual) social planner who has full information.

A voter with mostly informative strategy votes for her type if her type is *non-conflicting*, i.e., not $\bar{p}\bar{q}$; while she completely ignores the *conflicting* evidence ($t = \bar{p}\bar{q}$). In the case of informative strategy, conflicting evidence is followed *partly.* Unless we particularly mean one of these strategies, we say informative behaviour to refer to them. Note that rationality and efficiency refer to a whole profile of strategies.

We make two assumptions to avoid trivialities. First, we exclude the degenerate case where some decision in $\mathcal{J}$ is not efficient for any type profile. Hence, each decision is efficient for at least one type profile. Second, we exclude efficiency ties, i.e., those special parameter combinations such that some type profile leads to different efficient decisions (with different consequences when we assume consequentialist preferences). Hence, we exclude those instances where a voter is indifferent between two decisions except in the case that these decisions lead to the same consequence.

4.2.2 A general (im)possibility

How should the voting rule be designed so that it leads to efficient decisions as well as simple-minded, truthful voting behaviour in equilibrium? The objective of the chapter is to answer this question. By simple-minded, truthful behaviour, we mean informative behaviour. A voting rule encourages simple-minded behaviour if it makes informative voting rational.[6] If informative voting is both rational and efficient, the objective is reached. Note that neither an informative strategy nor a mostly informative strategy is unique. Informative voting being efficient means that for any given type profile $\mathbf{t}$, *every* profile of corresponding informative strategies is efficient. By the following theorem, our objective is reduced to finding out when informative voting is efficient.

Theorem 4.2.1 *For any common utility function $u : \mathcal{J}^2 \to \mathbb{R}$, and for any voting rule $f : \mathcal{J}^n \to \mathcal{J}$, if a strategy profile is efficient, then it is rational.*

This result applies to any kind of common preferences. Is it always possible to find a voting rule which makes informative voting efficient, hence, rational? The next theorem answers this question.

Theorem 4.2.2 *Consider an arbitrary common utility function $u : \mathcal{J}^2 \to \mathbb{R}$. There exists no voting rule for which mostly informative voting is efficient.*

This theorem states that there is no voting rule which achieves efficient information aggregation for every possible mostly informative strategy profile. This result comes as a surprise when one considers the single-issue setting and multi-issue setting with no interconnections, where there is always a voting rule for which informative voting is efficient. This contrast comes from the fact that the notion of informative voting is not very clear in the current setting, since there is no straightforward way of adapting informativeness. It is clear what a simple-minded voter should do when she receives non-conflicting evidence about the state of the world, but what about the conflicting evidence, $\bar{p}\bar{q}$? Here, voters with the mostly informative strategy have no restriction upon receiving type $\bar{p}\bar{q}$. This leads to the question of whether the impossibility persists when one considers informativeness differently. Let us now

[6] Here, by informative voting, we mean informative behaviour – with mostly informative strategy or informative strategy – in general. This chapter analyses the case with informative strategy and the case with mostly informative strategy separately.

consider informative voting. Voters holding informative strategy follow the conflicting evidence *partly*. Does this additional requirement lead to any possibility for efficient information aggregation? The result is yes, and it is formalized by the coming theorem. To state the theorem, we first introduce some notation and a condition.

Given a type profile $\mathbf{t} = (t_1, ..., t_n)$, let $\mathbf{t}_{pq} = (t_{pq_1}, ..., t_{pq_n})$ be the vector with entities defined for $i = 1, ..., n$ as $t_{pq_i} = 1$ if $t_i = pq$ and $t_{pq_i} = 0$ otherwise.[7] Given a voting profile $\mathbf{v}$, $\mathbf{v}_{pq}$ is defined similarly. At this point, it is useful to remark that for any type profile, there is an efficient decision.

Condition 1: For any $\mathbf{t}, \mathbf{t}' \in \mathcal{T}^n$ with $\mathbf{t}_{pq} = \mathbf{t}'_{pq}$, there is a decision $d \in \mathcal{J}$ which is efficient for both $\mathbf{t}, \mathbf{t}'$.

Theorem 4.2.3 *Consider an arbitrary common utility function $u : \mathcal{J}^2 \to \mathbb{R}$. There exists a voting rule for which informative voting is efficient if and only if Condition 1 holds.*

When we require an informative voter to follow the conflicting evidence partly instead of completely ignoring it, efficient information aggregation is possible when some condition on the model parameters is satisfied. How does the voting rule making efficient information aggregation possible look like? The answer depends on how exactly the utility function is specified. We can say more about this rule only when we focus on specific kind of preferences.

Having a general impossibility for informative voting with mostly informative strategy, we focus on informative voting for the rest of the chapter. To see how strong condition 1 is, one has to narrow focus on specific preferences. We study the two natural kinds of preferences – simple and consequentialist preferences – in the following subsections.

4.2.3 Simple preferences

We start by addressing simple preferences. Under simple preferences, correct decisions are preferred to incorrect ones without further sophistication. By narrowing down the focus on simple preferences, can we say more about the voting rule which makes informative voting efficient under Condition 1 and obtain a more specific result than the existential claim in Theorem 4.2.3? For simple preferences, we obtain the following impossibility.

Theorem 4.2.4 *Under simple preferences, there exists no voting rule for which informative voting is efficient.*

[7] For instance, for the type profiles $\mathbf{t} = (p\bar{q}, pq, \bar{p}\bar{q})$ and $\mathbf{t}' = (\bar{p}q, pq, \bar{p}q)$, $\mathbf{t}_{pq} = \mathbf{t}'_{pq} = (0, 1, 0)$.

It turns out that Condition 1 is never satisfied under simple preferences. In addition to the impossibility of efficient information aggregation with informative strategy stated in Theorem 4.2.2, efficient information aggregation with informative strategy is impossible under simple preferences. Does the impossibility persist under consequentialist preferences? The next subsection addresses this question.

4.2.4 Consequentialist preferences

We now turn to consequentialist preferences. We consider situations where the decision leads to one of *two* possible consequences. Such problems are very common in practice and widely studied in the judgment aggregation literature, where the two possible consequences are represented by conclusion propositions, c and $\bar{c}$. The decision leads to either acceptance of the conclusion proposition or rejection of it[8]. Consequence functions which lead all decisions to the same consequence are degenerate and uninteresting. If the consequence function depends only on the decision between p and $\bar{p}$, or only on the decision between q and $\bar{q}$, then the decision problem reduces to a problem with a single proposition-negation pair which has already been studied in the literature. Therefore, there is only one interesting consequence function up to isomorphism, and this function has the property $\mathrm{Co}(pq) \neq \mathrm{Co}(p\bar{q}) = \mathrm{Co}(\bar{p}q)$.

To state our result, we first define two coefficients:

$$A := \pi_{p\bar{q}} \left(\frac{1 - a_{\bar{q}}}{a_q} \right)^n + \pi_{\bar{p}q} \left(\frac{1 - a_{\bar{p}}}{a_p} \right)^{n-1} \frac{a_{\bar{p}}}{1 - a_p}$$

$$B := \pi_{p\bar{q}} \left(\frac{1 - a_{\bar{q}}}{a_q} \right)^{n-1} \frac{a_{\bar{q}}}{1 - a_q} + \pi_{\bar{p}q} \left(\frac{1 - a_{\bar{p}}}{a_p} \right)^n$$

Theorem 4.2.5 *Under consequentialist preferences, the following statements are equivalent:*

(a) *There exists a voting rule for which informative voting is efficient.*

(b) $A, B > \pi_{pq}$.

(c) *pq is the efficient decision only for the type profile $\mathbf{t} = (pq, ..., pq)$.*

Unlike under simple preferences, efficient information aggregation is possible under consequentialist preferences, if pq is the efficient decision *only* when there is overwhelming evidence for pq. This is satisfied when the prior probability of pq

[8] Consider the lead example of judgment aggregation: a jury is to decide whether the defendant has broken the contract (p) or not ($\bar{p}$) and whether the contract is legally valid (q) or not ($\bar{q}$). The defendant is convicted if and only if both propositions are collectively accepted. The consequence function here is encoded by $c \leftrightarrow (p \wedge q)$.

is sufficiently low compared to prior probabilities of $p\bar{q}$ and $\bar{p}q$. For instance, if $\pi_{pq} = \pi_{p\bar{q}} = \pi_{\bar{p}q} = 0.7$, $a_p = a_q = a_{\bar{p}} = a_{\bar{q}} = 0.6$ and $n = 3$, no voting rule makes informative voting efficient, whereas if instead $\pi_{pq} = 0.6$, such a voting rule exists. Now comes the main question: how do such rules look like? Let us call the condition stated at Theorem 4.2.5(b) *Condition 2*. We start by a simple characterization of voting rules which make informative voting efficient.

Proposition 4.2.1 *Assume consequentialist preferences and Condition 2. A voting rule f makes informative voting efficient if and only if for every voting profile $\mathbf{v} \in \mathcal{J}^n$, the decision $f(\mathbf{v})$ is pq if $\mathbf{v} = (pq, ..., pq)$ and in $\{p\bar{q}, \bar{p}q\}$ otherwise.*

While some of these voting rules making informative voting efficient are anonymous, monotonic and neutral, some of them fail to satisfy any of these properties. The number of votes for a proposition r in a voting profile $\mathbf{v}$ is written $n_r^{\mathbf{v}}$. For $n = 5$, two examples of anonymous and monotonic rules are given in the figure below:

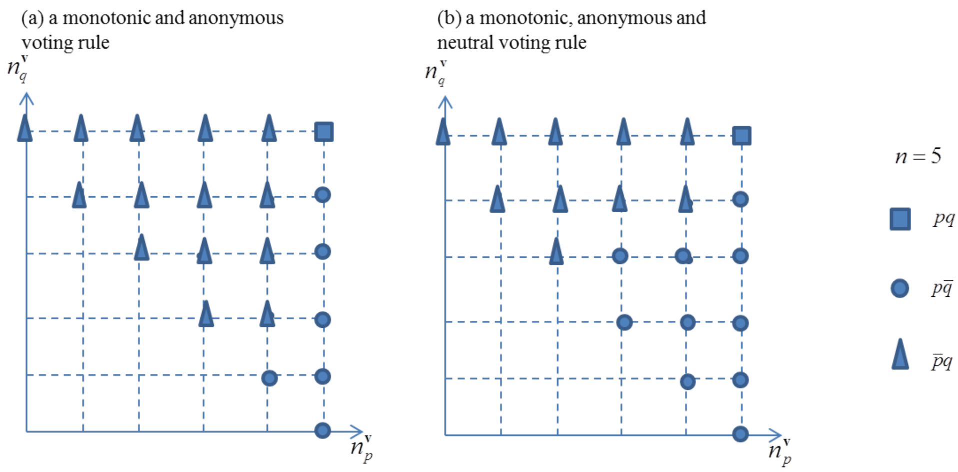


Figure 4.2.1: Two examples of voting rules given by Proposition 4.2.1 for $n = 5$

Figure 4.2.1(b) shows a neutral voting rule in addition to being anonymous and monotonic. This voting rule belongs to a class of voting rules defined by the following conditions. For each $\mathbf{v} \in \mathcal{J}^n$,

$$f(\mathbf{v}) = pq \iff n_p^{\mathbf{v}} = n_q^{\mathbf{v}} = n \tag{4.2.3}$$
$$f(\mathbf{v}) = p\bar{q} \text{ if } n_p^{\mathbf{v}} > n_q^{\mathbf{v}} \tag{4.2.4}$$
$$f(\mathbf{v}) = \bar{p}q \text{ if } n_p^{\mathbf{v}} < n_q^{\mathbf{v}} \tag{4.2.5}$$
$$f(\mathbf{v}) \in \{p\bar{q}, \bar{p}q\} \text{ if } n_p^{\mathbf{v}} = n_q^{\mathbf{v}} < n \tag{4.2.6}$$

By this class of voting rules defined by (4.2.3-4.2.6), we characterize the class of anonymous, monotonic and neutral voting rules making informative voting efficient under Condition 2.

Theorem 4.2.6 *Assume consequentialist preferences and Condition 2. A voting rule f makes informative voting efficient and is anonymous, monotonic and neutral if and only if f is defined by (4.2.3)-(4.2.6).*

Besides satisfying nice properties, these rules are reasonably practical and natural. We have now shown the necessary and sufficient conditions for the existence of a mechanism for efficient information aggregation, and characterized the mechanism with natural properties. Among the aggregation possibilities, anonymity, monotonicity and neutrality can be attained if required. What about independence? The next section answers this question.

4.3 Consistency and quota rules under consequentialist preferences

Quota rules are very natural and common among various voting rules. Under quota rules, separate votes are taken on each proposition using acceptance thresholds. In the previous section, we have seen that under consequentialist preferences, efficient information aggregation is possible in an anonymous, monotonic and neutral way when some condition on the model parameters is satisfied. Quota rules are monotonic, anonymous and *independent*, but not necessarily neutral.[9] This section examines the possibility of efficient information aggregation with quota rules under consequentialist preferences. To do so, one has to re-define a voting rule and utility function. The model described in the previous section applies to this section with exceptions described below.

Let $\mathcal{J}^* := \{pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$. For this section, a voting rule is a function $f : \mathcal{J}^n \to \mathcal{J}^*$, mapping each voting profile $\mathbf{v} = (v_1, ..., v_n)$ to a decision $d \equiv f(\mathbf{v})$. A voting rule f is called 'consistent' if it *never* returns $\bar{p}\bar{q}$.

The consequentialist utility function is now given as $u : \mathcal{J}^* \times \mathcal{J} \to \mathbb{R}$, mapping any decision-state pair $(d, s) \in \mathcal{J}^* \times \mathcal{J}$ to its utility $u(d, s)$. By assumption, the decision $\bar{p}\bar{q}$ never has the correct consequence, hence, $\mathrm{Co}(\bar{p}\bar{q}) \neq \mathrm{Co}(s)$. The consequentialist utility function is given by (4.2.2).

A quota rule is given by two thresholds $m_p, m_q \in \{0, 1, ..., n\}$, and for each voting profile it accepts p [q] if and only if at least m_p [m_q] voters accept it in the profile. Quota rules are characterized by anonymity, monotonicity and independence. The remark below follows from Theorem 2(c) in Dietrich and List (2007) and gives the necessary and sufficient conditions for a quota rule to be consistent under the given agenda.

[9] Whenever the acceptance thresholds for propositions are equal, they turn out to be neutral.

Remark 4.3.1 *A quota rule is consistent if and only if* $m_p + m_q \leq n + 1$.

The proof of this remark follows from the fact that there can be at most n votes in total for $\bar{p}$ and $\bar{q}$ (since $\bar{p}\bar{q} \notin \mathcal{J}$). It is easy to see that consistency is attained whenever $m_p + m_q \leq n$. Moreover, if $m_p + m_q = n + 1$, the number of p votes or the number of q votes exceeds the acceptance threshold, so, the resulting decision is never $\bar{p}\bar{q}$. Note that if a voting rule leads to efficient decisions, then it is consistent. So, as long as efficiency is guaranteed, consistency follows. Among all the rules making informative voting efficient, is there a quota rule? The theorem below answers this question.

Theorem 4.3.1 *Under consequentialist preferences, there exists no quota rule making informative voting efficient.*

A consistent quota rule is always available by Remark 4.3.1 regardless of the model parameters or the utility function chosen. However, there is no possibility for efficient information aggregation using quota rules.

4.4 Conclusion

We consider a model where a group of voters with common interests wants to form collective judgments over two propositions which are mutually interconnected. Each of these propositions is factually true or false, but the truth value is unknown to voters. Each voter has a type representing evidence about what the true state might be and this is private information. We study the problem of efficient information aggregation when propositions are mutually interconnected. The results depend particularly on how the utility function is specified. It turns out that a voting rule which makes informative voting efficient does not exist under simple preferences while such a rule exists under consequentialist preferences if some condition relating the model parameters and the utility function is satisfied. We want to design a voting rule which make every possible informative strategy profile efficient. Under simple preferences, it is of course possible to find a voting rule which makes *some* informative strategy profile efficient. However, we believe a voting rule which *sometimes* makes informative voting efficient is not really interesting.

We leave unanswered whether these results persist when conflicting private information – $\bar{p}\bar{q}$ in this case – is not allowed. In that case, we can no longer assume that a voter's evidence about one proposition is independent of the evidence about the other proposition conditional on the truth. For instance, if a voter has $\bar{p}$ in her type, she must have q. Informative strategy is then defined as direct revelation of types, and there is a unique informative strategy profile given a type profile.

4.5 Appendix: proofs

We begin by some preliminary derivations and then prove the results

4.5.1 Preliminary derivations

The joint probability of a state-types vector $(s, \mathbf{t}) = (s_p s_q, t_{1p} t_{1q}, ..., t_{np} t_{nq}) \in \mathcal{J}^{n+1}$ is

$$\Pr(s, \mathbf{t}) = \Pr(s) \Pr(\mathbf{t}|s) = \Pr(s) \prod_i \Pr(t_i|s) = \Pr(s_p) \Pr(s_q) \prod_i \Pr(t_{ip}|s_p) \Pr(t_{iq}|s_q),$$

where the last two equations follow from independence assumptions.

The probability of the three states in $\mathcal{J}$ conditional on the *full* information $\mathbf{t} \in \mathcal{J}^n$ is given as follows, where $k := n_p^{\mathbf{t}}$ and $l := n_q^{\mathbf{t}}$:

$$\Pr(pq|\mathbf{t}) = \frac{\pi_{pq} a_p^k (1 - a_p)^{n-k} a_q^l (1 - a_q)^{n-l}}{\Pr(\mathbf{t})} \tag{4.5.1}$$

$$\Pr(p\bar{q}|\mathbf{t}) = \frac{\pi_{p\bar{q}} a_p^k (1 - a_p)^{n-k} (1 - a_{\bar{q}})^l a_{\bar{q}}^{n-l}}{\Pr(\mathbf{t})} \tag{4.5.2}$$

$$\Pr(\bar{p}q|\mathbf{t}) = \frac{\pi_{\bar{p}q} (1 - a_{\bar{p}})^k a_{\bar{p}}^{n-k} a_q^l (1 - a_q)^{n-l}}{\Pr(\mathbf{t})}. \tag{4.5.3}$$

4.5.2 Proofs.

Proof of Theorem 4.2.1. Consider any voting rule $f : \mathcal{J}^n \to \mathcal{J}$ and any efficient strategy profile $\boldsymbol{\sigma}$. Consider any voter i and type $t_i \in \mathcal{T}$. To show that $\boldsymbol{\sigma}$ is rational, one has to show that i's vote $\sigma_i(t_i)$ maximizes her expected utility conditional on t_i. This follows from the fact that voters share common preferences. Since the resulting decision is efficient, it maximizes the expected utility of each voter. Thus, $\boldsymbol{\sigma}$ is rational. □

Proof of Theorem 4.2.2. Consider a voting rule $f : \mathcal{J}^n \to \mathcal{J}$. Suppose for a contradiction f makes informative voting efficient. Consider the type profile $\mathbf{t} = (\bar{p}\bar{q}, ..., \bar{p}\bar{q})$, where all voters have the type $\bar{p}\bar{q}$. Then, the set of all voting profiles which may result from informative voting is $\mathcal{J}^n$. Since informative voting is efficient, for each $\mathbf{v} \in \mathcal{J}^n$, $f(\mathbf{v})$ is efficient given $\mathbf{t}$. Then, it follows that for all type profiles in $\mathcal{T}^n$, $f(\mathbf{v})$ is efficient. This either means that some decision in $\mathcal{J}$ is not efficient for any type profile which contradicts to non-degeneracy assumption, or all decisions in $\mathcal{J}$ are always efficient for any type profile which contradicts to no efficiency ties assumption.[10] □

Proof of Theorem 4.2.3. To start with, we introduce some notation. Given a voting profile $\mathbf{v}$, let $\Theta(\mathbf{v})$ denote the set of all type profiles which possibly lead to $\mathbf{v}$ under informative voting. Given a type profile $\mathbf{t}$, let $\Omega(\mathbf{t})$ denote the set of all voting profiles which possibly result from $\mathbf{t}$ under informative voting. Consider a voting rule $f : \mathcal{J}^n \to \mathcal{J}$.

First, let Condition 1 hold. Suppose there is an exogenously given ordering of judgment sets, and let f be the following voting rule: for all $\mathbf{v} \in \mathcal{J}^n$, $f(\mathbf{v}) = d \iff d$ is the highest ordered decision among all decisions which are efficient for some $\mathbf{t} \in \Theta(\mathbf{v})$. Consider any type profile $\hat{\mathbf{t}} \in \mathcal{T}^n$ and suppose informative voting. We want to show that (*) for each $\mathbf{v} \in \Omega(\hat{\mathbf{t}})$, $f(\mathbf{v})$ is efficient for $\hat{\mathbf{t}}$. Let $\mathbf{v} \in \Omega(\hat{\mathbf{t}})$. One can show that all type profiles in $\Theta(\mathbf{v})$ share the same subvector restricted to pq. Since Condition 1 holds, there is some decision d which is efficient for all $\mathbf{t} \in \Theta(\mathbf{v})$, including $\hat{\mathbf{t}}$. It follows from Condition 1 that if any other decision $d' \neq d$ is efficient for *some* $\mathbf{t} \in \Theta(\mathbf{v})$, it is efficient for all $\mathbf{t} \in \Theta(\mathbf{v})$. Then, (*) holds.

Conversely, let f make informative voting efficient. Let $\mathbf{t}, \mathbf{t}'$ be two type profiles in $\mathcal{T}^n$ with $\mathbf{t}_{pq} = \mathbf{t}'_{pq}$. One has to show that (*) there is $d \in \mathcal{J}$ which is efficient for both $\mathbf{t}, \mathbf{t}'$. By construction, for each $\mathbf{v} \in \Omega(\mathbf{t})$, $\mathbf{t}' \in \Theta(\mathbf{v})$; and similarly, for each $\mathbf{v}' \in \Omega(\mathbf{t}')$, $\mathbf{t} \in \Theta(\mathbf{v}')$. Then, $f(\mathbf{v})$ must be efficient for $\mathbf{t}'$ (as well as $\mathbf{t}$) and $f(\mathbf{v}')$ must be efficient for $\mathbf{t}$ (as well as $\mathbf{t}'$) since informative voting is efficient. So, (*) holds. □

Proof of Theorem 4.2.4. By Theorem 4.2.2, it is sufficient to show that Condition 1 never holds under simple preferences. Suppose for a contradiction, it holds. Consider the two type profiles $\mathbf{t} = (p\bar{q}, ..., p\bar{q})$ and $\mathbf{t}' = (\bar{p}q, ..., \bar{p}q)$. Since $\mathbf{t}_{pq} = \mathbf{t}'_{pq}$ and Condition 1 holds, there is a decision which is efficient for both profiles. By non-degeneracy assumption, $p\bar{q}$ must be efficient for $\mathbf{t}$ since otherwise $p\bar{q}$ wouldn't be efficient for any type profile which contradicts to non-degeneracy assumption. Similarly, $\bar{p}q$ must be efficient for $\mathbf{t}'$. Hence, $p\bar{q}$ and $\bar{p}q$ are both efficient given $\mathbf{t}$ or $\mathbf{t}'$, which contradicts to no-efficiency ties assumption. □

Proof of Theorem 4.2.5. Let the statement in (b) be called Condition 2.

(1) We first prove that (c) implies (a) and (b). Assume Condition 2 holds. This implies that Condition 1 holds. By Theorem 4.2.3, there is a voting rule which makes informative voting efficient. Let $\mathbf{t}, \mathbf{t}'$ be type profiles with one $p\bar{q}$ and one $\bar{p}q$ respectively while each of the rest of the types is pq. Without loss of generality,

[10] For the case where the consequence function is defined in such a way that all three judgment sets in $\mathcal{J}$ lead to the same consequence, the second assumption is not violated. In such situations, there is no decision making problem since all decisions are equally good for each voter. Such utility functions are excluded.

let $\mathbf{t} = (pq, ..., \bar{p}q)$ and $\mathbf{t}' = (pq, ..., p\bar{q})$. By Condition 2, $p\bar{q}, \bar{p}q$ are both efficient for each of the type profiles. Using (4.5.1) and (4.5.3), we can write the following:

$$E(u(p\bar{q}, S)|\mathbf{t}) > E(u(pq, S)|\mathbf{t}) \tag{4.5.4}$$

$$\Leftrightarrow \pi_{p\bar{q}} a_p^{n-1}(1-a_p)(1-a_{\bar{q}})^n + \pi_{\bar{p}q}(1-a_{\bar{p}})^{n-1} a_{\bar{p}} a_q^n > \pi_{pq} a_p^{n-1}(1-a_p) a_q^n \tag{4.5.5}$$

$$\Leftrightarrow \pi_{p\bar{q}} \left(\frac{1-a_{\bar{q}}}{a_q}\right)^n + \pi_{\bar{p}q} \left(\frac{1-a_{\bar{p}}}{a_p}\right)^{n-1} \left(\frac{a_{\bar{p}}}{1-a_p}\right) > \pi_{pq}. \tag{4.5.6}$$

Similarly,

$$E(u(p\bar{q}, S)|\mathbf{t}') > E(u(pq, S)|\mathbf{t}') \tag{4.5.7}$$

$$\Leftrightarrow \pi_{p\bar{q}} a_p^n (1-a_{\bar{q}})^{n-1} a_{\bar{q}} + \pi_{\bar{p}q}(1-a_{\bar{p}})^n a_q^{n-1}(1-a_q) > \pi_{pq} a_p^n a_q^{n-1}(1-a_q) \tag{4.5.8}$$

$$\Leftrightarrow \pi_{p\bar{q}} \left(\frac{1-a_{\bar{q}}}{a_q}\right)^{n-1} \left(\frac{a_{\bar{q}}}{1-a_q}\right) + \pi_{\bar{p}q} \left(\frac{1-a_{\bar{p}}}{a_p}\right)^n > \pi_{pq}. \tag{4.5.9}$$

So, $A, B > \pi_{pq}$.

(2) We now prove that (a) implies (c). Consider a voting rule $f : \mathcal{J}^n \to \mathcal{J}$ and suppose f makes informative voting efficient. By Theorem 4.2.3, Condition 1 holds. Given a type profile $\mathbf{t} \in \mathcal{T}^n$, let $\Gamma(\mathbf{t})$ denote the set of type profiles which have the same subvector on pq as in $\mathbf{t}$. Recall that the number of occurrences for a proposition r in a type profile $\mathbf{t}$ is written $n_r^{\mathbf{t}}$. Now, take a type profile $\hat{\mathbf{t}} \in \mathcal{T}^n$ with k times pq where $1 \leq k < n$. The proof proceeds in several steps.

Claim 1: There is a type profile $\mathbf{t} \in \Gamma(\hat{\mathbf{t}})$ with $n_p^{\mathbf{t}} = k$ and $n_q^{\mathbf{t}} = k$.

Any type profile with k times pq and $n-k$ times $\bar{p}\bar{q}$ satisfies this condition and one of these type profiles is obviously in $\Gamma(\hat{\mathbf{t}})$. Now, take $\tilde{\mathbf{t}} \in \mathcal{T}^n$ with $k-1$ times pq.

Claim 2: There is a type profile $\mathbf{t} \in \Gamma(\tilde{\mathbf{t}})$ with $n_p^{\mathbf{t}} = k$ and $n_q^{\mathbf{t}} = k$.

One can easily see there is always a type profile with the exact same pq structure as $\tilde{\mathbf{t}}$ and with only one occurrence of $p\bar{q}$ and only one occurrence of $\bar{p}q$.

Claim 3: Under consequentialist preferences, for all $\mathbf{t}, \mathbf{t}' \in \mathcal{T}^n$ with $n_p^{\mathbf{t}} = n_p^{\mathbf{t}'}$ and $n_q^{\mathbf{t}} = n_q^{\mathbf{t}'}$, $E(u(d, S)|\mathbf{t}) = E(u(d, S)|\mathbf{t}')$ for each $d \in \mathcal{J}$.

The claim follows from the expressions (4.5.1)-(4.5.3). By Condition 1, there is a decision $d \in \mathcal{J}$ which is efficient for all $\mathbf{t} \in \Gamma(\hat{\mathbf{t}})$. Similarly, there is a decision $d \in \mathcal{J}$ which is efficient for all $\mathbf{t} \in \Gamma(\tilde{\mathbf{t}})$. Combining Claim 1, 2 and 3, one obtains that the same decision $d \in \mathcal{J}$ is efficient for all $\mathbf{t} \in \Gamma(\hat{\mathbf{t}})$ and all $\mathbf{t} \in \Gamma(\tilde{\mathbf{t}})$. Since this is true for all k with $1 \leq k < n$, there is a decision d which is efficient for all $\mathbf{t} \in \mathcal{T}^n \setminus \{(pq, ..., pq)\}$. By non-degeneracy assumption, pq is efficient for $\mathbf{t} = (pq, ..., pq)$. Hence, this decision must be in $\{p\bar{q}, \bar{p}q\}$ since otherwise pq would be efficient for all type profiles which contradicts to non-degeneracy assumption. Since $E(u(p\bar{q}, S)|\mathbf{t}) = E(u(\bar{p}q, S)|\mathbf{t})$ for all $\mathbf{t}$, both $p\bar{q}$ and $\bar{p}q$ are efficient for all $\mathbf{t} \in \mathcal{T}^n \setminus \{(pq, ..., pq)\}$.

Hence, Condition 2 holds.

(3) We finally prove that (b) implies (c). Let $A, B > \pi_{pq}$. To show that Condition 2 holds, we first show the following claim.

Claim 4: The expected utility of pq given a type profile $\mathbf{t}$ is an increasing function of $n_p^{\mathbf{t}}$ and $n_q^{\mathbf{t}}$.

The claim follows from the definition of the utility function and from $\Pr(S = pq|\mathbf{t})$ being an increasing function of $n_p^{\mathbf{t}}$ and $n_q^{\mathbf{t}}$. Let $\mathbf{t}, \mathbf{t}' \in \mathcal{T}^n$ be type profiles with one $p\bar{q}$ and one $\bar{p}q$ respectively while each of the rest of the types is pq. Without loss of generality, let $\mathbf{t} = (pq, ..., \bar{p}q)$ and $\mathbf{t}' = (pq, ..., p\bar{q})$. By (4.5.1) and (4.5.3), one has $E(u(p\bar{q}, S)|\mathbf{t}) > E(u(pq, S)|\mathbf{t})$ and $E(u(p\bar{q}, S)|\mathbf{t}') > E(u(pq, S)|\mathbf{t}')$. By the claim, it follows that $E(u(p\bar{q}, S)|\mathbf{t}) = E(u(\bar{p}q, S)|\mathbf{t}) > E(u(pq, S)|\mathbf{t})$ for all $\mathbf{t} \in \mathcal{T}^n \setminus \{(pq, ..., pq)\}$ which means $p\bar{q}, \bar{p}q$ are efficient for each $\mathbf{t} \in \mathcal{T}^n \setminus \{(pq, ..., pq)\}$. Thus, Condition 2 holds. □

Proof of Proposition 4.2.1. Consider a voting rule $f : \mathcal{J}^n \to \mathcal{J}$. Proof if the 'if' part is obvious and left to the reader. To show converse, let f make informative voting efficient. Since Condition 2 holds, for all voting profiles obtained by informative voting from any $\mathbf{t} \in \mathcal{T}^n \setminus \{(pq, ..., pq)\}$, $f(\mathbf{v}) \in \{p\bar{q}, \bar{p}q\}$. By non-degeneracy assumption, pq is efficient for $\mathbf{t} = (pq, ..., pq)$. By f making informative voting efficient, $f(\mathbf{v}) = pq$ if $\mathbf{v} = (pq, ..., pq)$. □

Proof of Theorem 4.2.6. Consider a voting rule $f : \mathcal{J}^n \to \mathcal{J}$. First, assume f is defined by (4.2.3)-(4.2.6). Clearly, f is anonymous. It follows from Proposition 4.2.1 that informative voting is efficient with f since for all $\mathbf{v} \in \mathcal{J}^n$, $f(\mathbf{v}) = pq$ if and only if $n_p^{\mathbf{v}} = n_q^{\mathbf{v}} = n$; so, if and only if $\mathbf{v} = (pq, ..., pq)$. To show monotonicity of f, take two voting profiles $\mathbf{v}, \mathbf{v}' \in \mathcal{J}^n$ such that for all $r \in f(\mathbf{v})$, the voters who vote for r in $\mathbf{v}$ also vote for r in $\mathbf{v}'$.

Case 1: $f(\mathbf{v}) = pq$. Then $\mathbf{v} = (pq, ..., pq)$. By definition, $\mathbf{v}' = \mathbf{v}$ and $f(\mathbf{v}') = pq$.

Case 2: $f(\mathbf{v}) = p\bar{q}$. The definition of f implies either $n_p^{\mathbf{v}} > n_q^{\mathbf{v}}$ or $n_p^{\mathbf{v}} = n_q^{\mathbf{v}} < n$; and the definition of $\mathbf{v}'$ implies $n_p^{\mathbf{v}'} \geq n_p^{\mathbf{v}}$ and $n_q^{\mathbf{v}'} \leq n_q^{\mathbf{v}}$. Suppose the former is true. Then, $n_p^{\mathbf{v}'} > n_q^{\mathbf{v}'}$ and $f(\mathbf{v}') = p\bar{q}$. Next, suppose $n_p^{\mathbf{v}} = n_q^{\mathbf{v}} < n$. If $\mathbf{v}' \neq \mathbf{v}$, one has $n_p^{\mathbf{v}'} > n_p^{\mathbf{v}}$ or $n_q^{\mathbf{v}'} < n_q^{\mathbf{v}}$ which means $n_p^{\mathbf{v}'} > n_q^{\mathbf{v}'}$ and $f(\mathbf{v}') = p\bar{q}$. It is obvious that if $\mathbf{v}' = \mathbf{v}$, we are done.

Case 3: $f(\mathbf{v}) = \bar{p}q$. One can show that $f(\mathbf{v}') = \bar{p}q$ analogously to Case 2.

It remains to show neutrality of f. Take two voting profiles $\mathbf{v}, \mathbf{v}' \in \mathcal{J}^n$ such that $\mathbf{v}_r = \mathbf{v}'_{r'}$ for every distinct $r, r' \in \{p, q\}$ and there is no permutation of voters $(i_1, ..., i_n)$ with $(v_{i_1}, ..., v_{i_n}) = (v'_1, ..., v'_n)$. We have to show that (*) f accepts r in $\mathbf{v}$ if and only if f accepts r' in $\mathbf{v}'$. We distinguish 3 cases:

Case 1: $f(\mathbf{v}) = pq$. It is clear that $\mathbf{v}' = \mathbf{v}$, and $f(\mathbf{v}') = pq$.

Case 2: $f(\mathbf{v}) = p\bar{q}$. By definition of f, either $n_p^{\mathbf{v}} > n_q^{\mathbf{v}}$ or $n_p^{\mathbf{v}} = n_q^{\mathbf{v}} < n$. One can see that the latter is not possible since then one could find a permutation of voters

$(i_1, ..., i_n)$ with $(v_{i_1}, ..., v_{i_n}) = (v'_1, ..., v'_n)$. Suppose the former is true. By definition of $\mathbf{v}'$, whenever p (q) is accepted in $\mathbf{v}$, q (p) is accepted in $\mathbf{v}'$. This means $n_p^{\mathbf{v}'} < n_q^{\mathbf{v}'}$ and $f(\mathbf{v}') = \bar{p}q$. So, f accepts p in $\mathbf{v}$ and q in $\mathbf{v}'$, and it accepts $\bar{q}$ in $\mathbf{v}$ and $\bar{p}$ in $\mathbf{v}'$. Hence, (*) holds.

Case 3: $f(\mathbf{v}) = \bar{p}q$. One can show that $f(\mathbf{v}') = \bar{p}q$ analogously to Case 2.

Conversely, let f be anonymous, monotonic and neutral, and make informative voting efficient. We have to show that (*) f is defined by (4.2.3)-(4.2.6). By Proposition 4.2.1 and informative voting being efficient, $f(\mathbf{v}) = pq$ if and only if $\mathbf{v} = (pq, ..., pq)$, equivalently $n_p^{\mathbf{v}} = n_q^{\mathbf{v}} = n$. Now, take a voting profile $\mathbf{v} \in \mathcal{J}^n \setminus \{(pq, ..., pq)\}$.

Case 1: $n_p^{\mathbf{v}} > n_q^{\mathbf{v}}$. Suppose for a contradiction, $f(\mathbf{v}) = \bar{p}q$. Let $\mathbf{v}'$ be a voting profile with $n_p^{\mathbf{v}'} = n_q^{\mathbf{v}}$ and $n_q^{\mathbf{v}'} = n_p^{\mathbf{v}}$. We start by proving the following claim.

Claim: For each combination of $k, l \in \{0, ..., n\}$, there is only one voting profile $\mathbf{v} \in \mathcal{J}^n$ with $n_p^{\mathbf{v}} = k$ and $n_p^{\mathbf{v}} = l$ up to the permutations of votes.

The claim follows from the fact that all votes containing $\bar{p}$ are $\bar{p}q$, and similarly, all votes containing $\bar{q}$ are $p\bar{q}$. Hence, subtracting number of p (q) occurrences in a profile from n gives the exact number of $\bar{p}q$ $(p\bar{q})$ votes. Then, there is only one voting profile with $n_p^{\mathbf{v}}$ times q and $n_q^{\mathbf{v}}$ times p up to permutations of votes by the claim. Hence, by neutrality and anonymity, $f(\mathbf{v}') = p\bar{q}$. However, by monotonicity of f, $f(\mathbf{v}') = \bar{p}q$ since $n_p^{\mathbf{v}'} \leq n_p^{\mathbf{v}}$ and $n_q^{\mathbf{v}'} \geq n_q^{\mathbf{v}}$, a contradiction. Then, $f(\mathbf{v}) = p\bar{q}$ if $n_p^{\mathbf{v}} > n_q^{\mathbf{v}}$.

Case 2: $n_p^{\mathbf{v}} < n_q^{\mathbf{v}}$. One can show that $f(\mathbf{v}) = \bar{p}q$ analogously to Case 1.

Case 3: $n_p^{\mathbf{v}} = n_q^{\mathbf{v}} < n$. By Proposition 4.2.1 and informative voting being efficient, $f(\mathbf{v}) \in \{p\bar{q}, \bar{p}q\}$.

So, (*) is true. □

Proof of Theorem 4.3.1. Consider a quota rule $f : \mathcal{J}^n \to \mathcal{J}^*$ with thresholds m_p and m_q. Suppose for a contradiction, f makes informative voting efficient. By Theorem 4.2.5, this means Condition 2 holds. Moreover, $f(\mathbf{v}) = pq$ if and only if $\mathbf{v} = (pq, ..., pq)$ by Condition 2 and informative voting being efficient. So, $m_p = m_q = n$. Now, consider any voting profile $\mathbf{v}$ which has the following property: if n is even, there are $\frac{n}{2}$ times $p\bar{q}$ and $\frac{n}{2}$ times $\bar{p}q$ in $\mathbf{v}$, and if n is odd, there are $\frac{n-1}{2}$ times $p\bar{q}$ and $\frac{n+1}{2}$ times $\bar{p}q$ in $\mathbf{v}$. It follows that $f(\mathbf{v}) = \bar{p}\bar{q}$ since $n \geq 2$, which is not efficient for any given type profile. Hence, a contradiction. □

Chapter 5

Judgment aggregation in search for the truth: the case of continuous information

5.1 Introduction

Judgment aggregation problems arise in situations where a group needs to form a 'yes' or 'no' judgment on different issues, based on individuals' judgments on these issues. The 'epistemic' approach in judgment aggregation has recently received some attention. This approach generally aims to track the truth. In the very well-known court trial problem, the aim of the jury seems to be to find out two independent facts (whether the defendant has broken the contract and whether the contract is legally valid). So, the jury's voting rule should be optimised with respect to the goal that the resulting group judgments are *true*.

This chapter takes the epistemic approach in a setting where voters with common interests and private information face a judgment aggregation problem with two issues, and asks questions about efficient information aggregation and strategic voting in a Bayesian voting game setting. If voters had not just common preferences, but also common information or beliefs, then no disagreement would arise. We allow for informational asymmetry: each voter has a continuum of types which represents private information about what the truth might be. Bozbay, Dietrich and Peters (2011) consider a similar problem where the private information is in the form of 'true' or 'false' for each issue. The model considered here departs from their model by assuming that private information of a voter can't be summarized as indicating only truth or falsity. We assume that a voter's private information is drawn from a continuous distribution. Duggan and Martinelli (2001) take an epistemic binary collective choice problem, i.e., a single-issue judgment aggregation problem, and they allow voters to receive a continuum of types. They characterize the equilibrium of the game while leaving open the question of efficient information aggregation. In this work, we deal with efficient information aggregation and want to find when voting rules lead to simple-minded and truthful behaviour as well as efficient outcomes.

By simple-minded and truthful behaviour, one might mean 'informative behaviour' or 'sincere behaviour'. Informative behaviour means that voters who receive types indicating a higher likelihood of truth about an issue judge that the issue is 'true' while voters who receive types indicating a lower likelihood of truth about an issue judge that the issue is 'not true'. Simple-minded behaviour might alternatively be represented by 'sincere voting', which means voting without taking other voters into consideration. In the classical Condorcet jury theorem framework, it is generally assumed that informative behaviour and sincere behaviour coincide and voters vote informatively and sincerely. Austen-Smith and Banks (1996) show that this assumption is inconsistent with a game-theoretic view of collective behaviour. Following their track, we give a characterization of all problems where informative voting and sincere voting coincide. For an extended survey of the two literatures to which this chapter connects to, please see Section 3.1.1.

This chapter proceeds as follows. Section 5.2 introduces the model. In Section 5.3, we characterize the conditions where informative behaviour and sincere behaviour are equivalent. Section 5.4 focuses on the key question of efficient information aggregation, characterizing the problems where informative voting is efficient with a quota rule – a rule which takes separate votes on each issue using acceptance thresholds. We conclude by discussing the possible extensions of this work in Section 5.5.

5.2 The Model

We consider a group of voters, labelled $i = 1, ..., n$, where $n \geq 2$. This group needs a collective judgment on whether some proposition p or its negation $\bar{p}$ is true, and whether some other proposition q or its negation $\bar{q}$ is true. The four possible judgment sets are $\{p, q\}, \{p, \bar{q}\}, \{\bar{p}, q\}$ and $\{\bar{p}, \bar{q}\}$; we abbreviate them by pq, $p\bar{q}$, $\bar{p}q$ and $\bar{p}\bar{q}$, respectively. For instance, $p\bar{q}$ means accepting p but not q. Each voter votes for a judgment set in $\mathcal{J} = \{pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$. After all voters cast their votes, a collective decision in $\mathcal{J}$ is taken using a voting rule. Formally, a voting rule is a function $f : \mathcal{J}^n \to \mathcal{J}$, mapping each voting profile $\mathbf{v} = (v_1, ..., v_n)$ to a decision $d \equiv f(\mathbf{v})$. Among the various voting rules, *quota rules* stand out as particularly natural and common. A quota rule is given by two thresholds $m_p, m_q \in \{1, ..., n\}$, and for each voting profile it accepts p $[q]$ if and only if at least m_p $[m_q]$ voters accept it in the profile.

Exactly one judgment set in $\mathcal{J}$ is 'correct', i.e., contains propositions which are factually true. It is called the *state (of the world)* and is generically denoted by s. For instance, the state might be $p\bar{q}$, so that p and $\bar{q}$ are true (and $\bar{p}$ and q are false). Voters have identical preferences, captured by a common utility function $u : \mathcal{J} \times \mathcal{J} \to \mathbb{R}$ which maps any decision-state pair (d, s) to its utility $u(d, s)$. The utility function is given by

$$u(d,s) = \begin{cases} 1 & \text{if } d = s \text{ (correct decision)} \\ 0 & \text{if } d \neq s \text{ (incorrect decision).} \end{cases} \tag{5.2.1}$$

Such preferences are the simplest candidate for truth-tracking preferences. Correct decisions are preferred to incorrect ones, without further sophistication.[1]

Each voter has a *type*, representing private information or evidence.[2] A voter's type is denoted by $t = (t_p, t_q) \in [0,1]^2$. When we refer to the type of particular voter i, we write $t_i = (t_{ip}, t_{iq})$. For each $r \in \{p, q\}$, t_{ir} is distributed randomly according to the state conditional density f_r or $f_{\bar{r}}$ depending on whether r is true or the negation $\bar{r}$ is true. For instance, the probability that $t_p \leq \bar{\tau}$ conditional on p being true is given by the expression $\int_0^{\bar{\tau}} f_p(t_p)dt_p$. By assumption, the prior probability that r ($\in \{p, \bar{p}, q, \bar{q}\}$) is true is denoted

$$\pi_r = \Pr(r),$$

and belongs to $(0, 1)$. Note that $\pi_{\bar{r}} = 1 - \pi_r$. We write $\mathbf{t} = (t_1, ..., t_n) \in ([0,1]^2)^n$ for a vector of voters' types. We denote by $T = (T_p, T_q)$ the random variable generating voters' types and by $S = (S_p, S_q)$ the random variable generating the state s. We make the following assumptions for each $r \in \{p, q\}$.

- The state and types relative to p are independent of the state and types relative to q. [3]
- Each type for each proposition is independently drawn given the true state of the proposition.[4]
- The state conditional densities f_r and $f_{\bar{r}}$ are piecewise continuous, and $f_r(t_r), f_{\bar{r}}(t_r) > 0$ for all $t_r \in [0, 1]$.

[1]This kind of preferences is called 'simple preferences' in the remaining chapters of this part where we consider two different kinds of preferences; simple and consequentialist preferences. In this chapter, we only focus on simple preferences.

[2]The type could represent information that is not shared with other voters because of a lack of deliberation or limits of deliberation. More generally, a voter i's type could represent uncertainty of other voters about i's beliefs.

[3]Formally, $S_p, T_{1p}, ..., T_{np}$ are independent of $S_q, T_{1q}, ..., T_{nq}$.

[4]Formally, $T_{1p}, ..., T_{np}$ are conditionally independent given S_p and $T_{1q}, ..., T_{nq}$ are conditionally independent given S_q.

- $\frac{f_{\bar{r}}}{f_r}$ is weakly decreasing on $[0,1]$.[5]

The two independence assumptions allow us to characterize the joint distribution of the state and types. The last assumption ensures that types convey information about the state: the higher values of types are stronger indications of propositions' being true. This assumption also implies that for each proposition, the distribution of types conditional on 'true' first-order stochastically dominates the distribution of types conditional on 'false'.[6]

Voter i's belief about proposition p conditional on part of his type t_p is represented by the expression $\Pr(S_p = p|T_p = t_p)$. We abuse notation for simplicity and we write $\Pr(p|t_p)$. This expression is estimated as follows:

$$\begin{aligned}
\Pr(p|t_p) &= \lim_{\Delta x \to 0} \Pr(p|t_p < T_p \leqslant t_p + \Delta x) \\
&= \lim_{\Delta x \to 0} \frac{\Pr(p, t_p < T_p \leqslant t_p + \Delta x)}{\Pr(t_p < T_p \leqslant t_p + \Delta x)} \\
&= \lim_{\Delta x \to 0} \frac{\pi_p f_p(t_p)\Delta x}{(\pi_p f_p(t_p) + (1-\pi_p) f_{\bar{p}}(t_p))\Delta x} \\
&= \frac{\pi_p f_p(t_p)}{\pi_p f_p(t_p) + (1-\pi_p) f_{\bar{p}}(t_p)}.
\end{aligned}$$

Similarly, one can derive the expressions for $\Pr(q|t_q)$, $\Pr(\bar{p}|t_p)$ and $\Pr(\bar{q}|t_q)$. The expressions for the probability of each state conditional on a given type are in the Appendix.

After a voter receives his type, he submits a vote in $\mathcal{J}$. A voting strategy is a function $\sigma : [0,1]^2 \to \mathcal{J}$, mapping each possible type $t \in [0,1]^2$ to that type's vote $\sigma(t) = v$. A strategy profile is a vector $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ of strategies across voters. We denote the vector of votes by $\mathbf{v} = (v_1, ..., v_n)$.

With a voting rule f and a common utility function u, we now have a well-defined Bayesian game. For a type profile $\mathbf{t}$, we call a decision d *efficient* if it has maximal expected utility conditional on the full information $\mathbf{t}$, i.e., if it maximizes $E(u(d,S)|\mathbf{t})$. Some common notions of voting behaviour can now be adapted to our framework:

- A strategy σ of a voter is *informative* if there exists a *cut-point* $\hat{t} = (\hat{t}_p, \hat{t}_q) \in (0,1)^2$ such that $r \in \sigma(t)$ if $t_r \geqslant \hat{t}_r$ and $\bar{r} \in \sigma(t)$ if $t_r < \hat{t}_r$ for all $t \in [0,1]^2$ and $r \in \{p,q\}$. Hence, an informative voter votes according to the relation

[5] This assumption is commonly used especially in mechanism design problems and is called monotone likelihood ratio property (MLRP).

[6] Formally, for each $r \in \{p,q\}$, the conditional cumulative distribution function F_r has first-order stochastic dominance over the conditional cumulative distribution function $F_{\bar{r}}$; hence, $F_{\bar{r}}(t_r) \geq F_r(t_r)$ for all $t_r \in [0,1]$.

between her type and the cut-point. [7]

- A strategy σ of a voter is *sincere* if for every type t, the vote $\sigma(t)$ maximizes the expected utility conditional on the information t. A sincere voter acts as if her vote alone determined the decision, which amounts to optimal behaviour in a hypothetical single-player decision problem.
- A strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ is *rational* if each strategy is a best response to the other strategies, i.e., if the profile is a Nash equilibrium of the corresponding Bayesian game.
- A strategy profile $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_n)$ is *efficient* if for every type profile $\mathbf{t} = (t_1, ..., t_n)$ the resulting decision $d = f(\sigma_1(t_1), ..., \sigma_n(t_n))$ is efficient (i.e., has maximal expected utility conditional on full information $\mathbf{t}$). Hence, all the information spread across the group is used efficiently.

Informativeness and sincerity are properties of a single strategy while efficiency and rationality refer to a whole profile of strategies. An *informative strategy profile* is a profile of informative strategies with some cut-point $\hat{t} \in (0,1)^2$ which is equal for all voters. To avoid distraction by special cases, we make two assumptions over the parameters of the model π_r, f_r and $f_{\bar{r}}$ where $r \in \{p, q, \bar{p}, \bar{q}\}$. First, we exclude the degenerate case in which some decision in $\mathcal{J}$ is not efficient for any type profile whatsoever. Second, we exclude efficiency ties, i.e., we exclude those special parameter combinations such that some type profile $\mathbf{t}$ leads to different efficient decisions.

5.3 When is informative voting sincere?

In the classical Condorcet Jury Theorem framework, voters are taken to vote sincerely. Moreover, it is generally assumed that informative voting is sincere. In our framework, the latter is not necessarily true as the following theorem shows:

Theorem 5.3.1 *The informative voting strategy with cut-point $\hat{t} = (\hat{t}_p, \hat{t}_q) \in (0,1)^2$ is sincere if and only if for each $r \in \{p, q\}$, $\frac{\pi_r}{1-\pi_r} \geq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r \geq \hat{t}_r$ and $\frac{\pi_r}{1-\pi_r} \leq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r < \hat{t}_r$.*

An immediate corollary of the theorem is as follows:

Corollary 5.3.1 *There exists an informative and sincere strategy if and only if there is $t = (t_p, t_q) \in (0,1)^2$ satisfying $\frac{\pi_r}{1-\pi_r} \geq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r \geq \hat{t}_r$ and $\frac{\pi_r}{1-\pi_r} \leq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r < \hat{t}_r$, $r \in \{p, q\}$.*

[7] We exclude the uninteresting cases of $t_r \in \{0, 1\}$ since then an informative voter would always accept or always reject the related proposition regardless of the type she received.

The following remark follows from the above corollary and $\frac{f_{\bar{r}}}{f_r}$ being a decreasing function.

Remark 5.3.1 *There exists an informative and sincere strategy if and only if for each* $r \in \{p, q\}$, $\frac{\pi_r}{1-\pi_r} \in (\frac{f_{\bar{r}}(0)}{f_r(0)}, \frac{f_{\bar{r}}(1)}{f_r(1)})$.

Note that if $\frac{f_{\bar{r}}}{f_r}$ was assumed to be *strictly* decreasing for each $r \in \{p, q\}$, then the informative and sincere strategy would be unique. Under our weaker assumption, there may be an interval of cut-points making the associated informative strategy sincere.

5.4 When is informative voting efficient?

We now turn to efficiency of informative voting. Our objective is to find out whether a voting rule which encourages informative voting behaviour and leads to efficient decisions exists. The former is reached if informative voting occurs in equilibrium; hence, if it is rational. By the following remark, our objective reduces to finding whether there are voting rules which make informative voting efficient.[8]

Remark 5.4.1 *For any voting rule, if a strategy profile is efficient, then it is rational.*

If voters had private information in the form of 'true' or 'false' for each proposition (so, in the form of an element of $\mathcal{J}$), then an informative strategy would be defined as direct revelation of the type. Under this information structure, there is always a quota rule which makes informative voting efficient (Bozbay, Dietrich, Peters 2011). Is efficient information aggregation possible with quota rules when voters' private information reflects a rich spectrum of possibilities instead of only indicating truth or not? The following theorem answers this question.

Theorem 5.4.1 *A quota rule with thresholds* m_p, m_q *makes informative voting with cut-point* $\hat{t}$ *efficient if and only if the following condition is satisfied for all* $\mathbf{t} \in ([0,1]^2)^n$:

$$|\{i : t_{ir} \geq \hat{t}_r\}| \geq m_r \Leftrightarrow \frac{\pi_r}{1-\pi_r} > \prod_{i \in \{1,\dots,n\}} \frac{f_{\bar{r}}(t_{ir})}{f_r(t_{ir})} \text{ for each } r \in \{p, q\}. \tag{5.4.1}$$

Efficient information aggregation is possible under some condition on the model parameters. We now present two examples, starting with the one where (5.4.1) is satisfied and turning to another where (5.4.1) does not hold.

[8] By saying 'informative voting' without referring to a particular voter, we mean informative voting by all voters with a common cut-point.

Example 1:[9] Let $n = 3$, $\pi_p = \pi_q = 0.6$. Given $0.5 < \alpha < 1$, f_r and $f_{\bar{r}}$ are defined as follows for each $r \in \{p, q\}$ and $t_r \in [0, 1]$.

$$f_r(t_r) = \begin{cases} 1 - \alpha & \text{if } 0 \leq t_r \leq 0.5 \\ \alpha & \text{if } 0.5 < t_r \leq 1 \end{cases}$$

and

$$f_{\bar{r}}(t_r) = \begin{cases} \alpha & \text{if } 0 \leq t_r \leq 0.5 \\ 1 - \alpha & \text{if } 0.5 < t_r \leq 1 \end{cases}$$

Then, (5.4.1) is satisfied (only) for $m_p = m_q = 2$ and $\hat{t} = (0.5, 0.5)$ under this parametrization. Voters who receive $t_r \in [0, 0.5]$ accept r while voters who receive $t_r \in (0.5, 1]$ reject r. This is of course a special case. In the model where voters receive private information in the form of 'true' or 'false' for each proposition, this means that whenever probability of r being true conditional on receiving evidence that 'r is true' is above 0.5, one accepts r. Here, a continuum of types replaces each of the binary types, i.e., each of 'true' or 'false'.

Example 2: We only show for $r = p$ that (5.4.1) does not hold. Let $n = 3$, $\pi_p = 0.6$, $f_p(t_p) = 2t_p$ and $f_{\bar{p}}(t_p) = 1$ for each $t_p \in [0, 1]$. We consider the following three type profiles:

- $\mathbf{t}^1 = ((0.5, t^1_{1q}), (0.4, t^1_{2q}), (0.4, t^1_{3q}))$,
- $\mathbf{t}^2 = ((0.5, t^2_{1q}), (0.6, t^2_{2q}), (0.4, t^2_{3q}))$,
- $\mathbf{t}^3 = ((0.6, t^3_{1q}), (0.6, t^3_{2q}), (0.2, t^3_{3q}))$.

For $\mathbf{t}^1$, we have $\frac{\pi_p}{1-\pi_p} < \prod_{i \in \{1,\ldots,n\}} \frac{f_{\bar{p}}(t^1_{ip})}{f_p(t^1_{ip})}$. Suppose there exist $m_p, \hat{t}_p$ such that $|\{i : t^1_{ip} \geq \hat{t}_p\}| < m_p$. Now, consider $\mathbf{t}^2$. Since $\frac{\pi_p}{1-\pi_p} > \prod_{i \in \{1,\ldots,n\}} \frac{f_{\bar{p}}(t^2_{ip})}{f_p(t^2_{ip})}$, for (5.4.1) to hold, we must have $|\{i : t^2_{ip} \geq \hat{t}_p\}| \geq m_p$ for $\mathbf{t}^2$. This implies that $0.6 \geq \hat{t}_p$. Then, for $\mathbf{t}^3$, we must have $|\{i : t^3_{ip} \geq \hat{t}_p\}| \geq m_p$. However, $\frac{\pi_p}{1-\pi_p} < \prod_{i \in \{1,\ldots,n\}} \frac{f_{\bar{p}}(t^3_{ip})}{f_p(t^3_{ip})}$.

5.5 Conclusion

We consider a model where a group of voters with common interests wants to make collective decision over two propositions each of which is factually true or false. Different from the model of Bozbay, Dietrich, Peters (2011) where each voter's type either takes the form of 'true' or 'false' for each proposition, we assume that voters' types are distributed from a state-dependent continuous distribution allowing for a

[9] This example is introduced for the single issue case by Duggan and Martinelli (2011). We hereby adapt it to our multi-issue problem.

more realistic model. In contrast to their results (where informative voting is always and only efficient with quota rules), efficient information aggregation with a quota rule is possible only under a strong condition in this work.

An interesting extension would be to consider a larger class of voting rules and analyse the possibilities of efficient information aggregation within this class. Fixing the mechanism and characterizing the equilibrium of the game would be another interesting extension. The multi-issue problem provides possibility for different preference specifications. Assuming that collective decisions have consequences, one can define 'consequentialist preferences' (see Bozbay, Dietrich and Peters [2011]) as giving some (positive) utility if and only if the consequence of the decision matches the consequence of the state. In the court trial example, suppose the defendant is convicted if and only if both issues are collectively accepted by the jury. Defining the utility function as a mapping $u : \mathcal{J} \times \mathcal{J} \to \mathbb{R}$ such that $u(d, s) = 1$ if the consequence of the decision is the same as the consequence of the state and $u(d, s) = 0$ otherwise, one question that immediately arises is whether these results persist under consequentialist preferences.

5.6 Appendix: proofs

5.6.1 Preliminary derivations

A voter's probability of a state $s \in \mathcal{J}$ conditional on her type $t \in [0, 1]^2$ is given by the following expressions.

$$\Pr(S = pq|T = t) = \frac{\pi_p \pi_q f_p(t_p) f_q(t_q)}{(\pi_p f_p(t_p) + \pi_{\bar{p}} f_{\bar{p}}(t_p))(\pi_q f_q(t_q) + \pi_{\bar{q}} f_{\bar{q}}(t_q))} \tag{5.6.1}$$

$$\Pr(S = p\bar{q}|T = t) = \frac{\pi_p \pi_{\bar{q}} f_p(t_p) f_{\bar{q}}(t_q)}{(\pi_p f_p(t_p) + \pi_{\bar{p}} f_{\bar{p}}(t_p))(\pi_q f_q(t_q) + \pi_{\bar{q}} f_{\bar{q}}(t_q))} \tag{5.6.2}$$

$$\Pr(S = \bar{p}q|T = t) = \frac{\pi_{\bar{p}} \pi_q f_{\bar{p}}(t_p) f_q(t_q)}{(\pi_p f_p(t_p) + \pi_{\bar{p}} f_{\bar{p}}(t_p))(\pi_q f_q(t_q) + \pi_{\bar{q}} f_{\bar{q}}(t_q))} \tag{5.6.3}$$

$$\Pr(S = \bar{p}\bar{q}|T = t) = \frac{\pi_{\bar{p}} \pi_{\bar{q}} f_{\bar{p}}(t_p) f_{\bar{q}}(t_q)}{(\pi_p f_p(t_p) + \pi_{\bar{p}} f_{\bar{p}}(t_p))(\pi_q f_q(t_q) + \pi_{\bar{q}} f_{\bar{q}}(t_q))} \tag{5.6.4}$$

The probability of the four states in $\mathcal{J}$ conditional on the full information $\mathbf{t} \in$

$([0,1]^2)^n$ is given as follows:

$$\Pr(pq|\mathbf{t}) = \frac{\pi_p \pi_q \prod_{i\in\{1,\dots,n\}} f_p(t_{ip}) f_q(t_{iq})}{\Pr(\mathbf{t})} \tag{5.6.5}$$

$$\Pr(p\bar{q}|\mathbf{t}) = \frac{\pi_p \pi_{\bar{q}} \prod_{i\in\{1,\dots,n\}} f_p(t_{ip}) f_{\bar{q}}(t_{iq})}{\Pr(\mathbf{t})} \tag{5.6.6}$$

$$\Pr(\bar{p}q|\mathbf{t}) = \frac{\pi_{\bar{p}} \pi_q \prod_{i\in\{1,\dots,n\}} f_{\bar{p}}(t_{ip}) f_q(t_{iq})}{\Pr(\mathbf{t})} \tag{5.6.7}$$

$$\Pr(\bar{p}\bar{q}|\mathbf{t}) = \frac{\pi_{\bar{p}} \pi_{\bar{q}} \prod_{i\in\{1,\dots,n\}} f_{\bar{p}}(t_{ip}) f_{\bar{q}}(t_{iq})}{\Pr(\mathbf{t})}. \tag{5.6.8}$$

5.6.2 Proofs

Proof of Theorem 5.3.1. Consider a type $t = (t_p, t_q) \in [0,1]^2$ and the informative strategy σ with cut-point $\hat{t} \in (0,1)^2$. We refer to the following fact throughout the proof.

Fact 1: $E(u(d,S)|t) = \Pr(d|t)$ for each $d \in \mathcal{J}$.

Suppose σ is sincere. We want to show that $(*)$ for each $r \in \{p,q\}$, $\frac{\pi_r}{1-\pi_r} \geq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r \geq \hat{t}_r$ and $\frac{\pi_r}{1-\pi_r} \leq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r < \hat{t}_r$. Let $t_r \geq \hat{t}_r$ for each $r \in \{p,q\}$.[10] Then, $\sigma(t) = pq$. Since σ is sincere, $E(u(d,S)|t)$ is maximized if $d = pq$. Using (5.6.1)[11], (5.6.2) and Fact 1, we write:

$$\begin{aligned}
&E(u(pq,S)|t) \geq E(u(p\bar{q},S)|t) \\
\Leftrightarrow&\pi_p\pi_q f_p(t_p) f_q(t_q) \geq \pi_p(1-\pi_q) f_p(t_p) f_{\bar{q}}(t_q) \\
\Leftrightarrow&\frac{\pi_q}{1-\pi_q} \geq \frac{f_{\bar{q}}(t_q)}{f_q(t_q)}.
\end{aligned}$$

Similarly, using (5.6.1), (5.6.3) and Fact 1,

$$\begin{aligned}
&E(u(pq,S)|t) \geq E(u(\bar{p}q,S)|t) \\
\Leftrightarrow&\pi_p\pi_q f_p(t_p) f_q(t_q) \geq (1-\pi_p)\pi_q f_p(t_p) f_{\bar{q}}(t_q) \\
\Leftrightarrow&\frac{\pi_p}{1-\pi_p} \geq \frac{f_{\bar{p}}(t_p)}{f_p(t_p)}.
\end{aligned}$$

Next, let $t_r < \hat{t}_r$ for each $r \in \{p,q\}$. So, $\sigma(t) = \bar{p}\bar{q}$. Since σ is sincere, $E(u(\bar{p}\bar{q},S)|t)$ is maximal. Then, $E(u(\bar{p}\bar{q},S)|t) \geq E(u(p\bar{q},S)|t)$, which leads to $\frac{\pi_p}{1-\pi_p} \leq \frac{f_{\bar{p}}(t_p)}{f_p(t_p)}$ by (5.6.2) and (5.6.4). Similarly, from $E(u(\bar{p}\bar{q},S)|t) \geq E(u(\bar{p}q,S)|t)$, we have $\frac{\pi_q}{1-\pi_q} \leq \frac{f_{\bar{q}}(t_q)}{f_q(t_q)}$ by (5.6.4) and (5.6.3). Thus, $(*)$ is true.

[10] Note that $\Pr(T_r = \hat{t}_r)$ is a 0-probability event.

[11] See the Appendix for this expression and other expressions referred throughout the proof.

Conversely, suppose for each $r \in \{p,q\}$, $\frac{\pi_r}{1-\pi_r} \geq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r \geq \hat{t}_r$ and $\frac{\pi_r}{1-\pi_r} \leq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for all $t_r < \hat{t}_r$. Let $\sigma(t) = v$. We want to show that σ is sincere, hence, $E(u(d,S)|t)$ is maximized at $d = v$. Let $v = pq$. This means $t_r \geq \hat{t}_r$ for each $r \in \{p,q\}$. Then, $\frac{\pi_r}{1-\pi_r} \geq \frac{f_{\bar{r}}(t_r)}{f_r(t_r)}$ for each $r \in \{p,q\}$. By (5.6.1), (5.6.2), (5.6.3) and (5.6.4), $E(u(pq,S)|t)$ is maximal. One can similarly show for the remaining cases ($\mathbf{v} \in \{p\bar{q}, \bar{p}q, \bar{p}\bar{q}\}$) that $E(u(d,S)|t)$ is maximized at $d = v$. Thus, σ is sincere. □

Proof of Theorem 5.4.1. We start the proof by introducing the following lemma.

Lemma 5.6.1 *For all type profiles $\mathbf{t} = (t_1, ..., t_n) \in ([0,1]^2)^n$, all $r \in \{p,q\}$ and all decisions $d, d' \in \mathcal{J}$ such that d but not d' contains r and d and d' share the other proposition or its negation,*

$$E(u(d,S)|\mathbf{t}) > E(u(d',S)|\mathbf{t}) \Leftrightarrow \frac{\pi_r}{1-\pi_r} > \prod_{i\in\{1,\dots,n\}} \frac{f_{\bar{r}}(t_{ir})}{f_r(t_{ir})}. \tag{5.6.9}$$

Proof of Lemma. We first prove the equivalence for $r = p$, $d = pq$ and $d' = \bar{p}q$. Since $E(u(d,S)|\mathbf{t}) = \Pr(d|\mathbf{t})$ for each $d \in \mathcal{J}$, $E(u(pq,S)|\mathbf{t}) > E(u(\bar{p}q,S)|\mathbf{t}) \Leftrightarrow \Pr(pq|\mathbf{t}) > \Pr(\bar{p}q|\mathbf{t})$. By (5.6.5) and (5.6.7), we can write

$$\pi_p \pi_q \prod_{i\in\{1,\dots,n\}} f_p(t_{ip}) f_q(t_{iq}) > (1-\pi_p)\pi_q \prod_{i\in\{1,\dots,n\}} f_{\bar{p}}(t_{ip}) f_q(t_{iq}).$$

Simplifying and rearranging, we obtain

$$\frac{\pi_p}{1-\pi_p} > \prod_{i\in\{1,\dots,n\}} \frac{f_{\bar{p}}(t_{ip})}{f_p(t_{ip})}.$$

Hence, (5.6.9) holds for $r = p$, $d = pq$ and $d' = \bar{p}q$. Next, suppose $r = p$, $d = p\bar{q}$ and $d' = \bar{p}\bar{q}$. By (5.6.6) and (5.6.8), this gives

$$\begin{aligned} &E(u(p\bar{q},S)|\mathbf{t}) > E(u(\bar{p}\bar{q},S)|\mathbf{t}) \\ \Leftrightarrow &\pi_p \pi_q \prod_{i\in\{1,\dots,n\}} f_p(t_{ip}) f_{\bar{q}}(t_{iq}) > (1-\pi_p)(1-\pi_q) \prod_{i\in\{1,\dots,n\}} f_{\bar{p}}(t_{ip}) f_{\bar{q}}(t_{iq}) \\ \Leftrightarrow &\frac{\pi_p}{1-\pi_p} > \prod_{i\in\{1,\dots,n\}} \frac{f_{\bar{p}}(t_{ip})}{f_p(t_{ip})}. \end{aligned}$$

The proof for the remaining two cases is analogous. □

To prove the theorem, suppose informative voting with some cut-point $\hat{t} \in (0,1)^2$ is efficient. Consider a quota rule f with thresholds m_p, m_q. We take any $\mathbf{v} \in \mathcal{J}$ and consider the following four cases.

Case 1: $f(\mathbf{v}) = pq$. This means for the type profile $\mathbf{t}$ which leads to $\mathbf{v}$ under informative voting, $(*)$ $|i : t_{ir} \geq \hat{t}_r| \geq m_r$ for each $r \in \{p, q\}$. Since informative voting is efficient, pq is efficient given any type profile satisfying $(*)$. By Lemma 5.6.1, we have $\frac{\pi_r}{1-\pi_r} > \prod_{i\in\{1,...,n\}} \frac{f_{\bar{r}}(t_{ir})}{f_r(t_{ir})}$ for each $r \in \{p, q\}$, which holds for all $\mathbf{t}$ satisfying $(*)$. (We use $>$ rather than $=$ because of the no efficiency ties assumption.)

Case 2: $f(\mathbf{v}) = p\bar{q}$. This means for the type profile $\mathbf{t}$ which leads to $\mathbf{v}$ under informative voting, $(**)$ $|i : t_{ip} \geq \hat{t}_p| \geq m_p$ and $|i : t_{iq} \geq \hat{t}_q| < m_q$. Since informative voting is efficient, $p\bar{q}$ is efficient given any type profile satisfying $(**)$. By Lemma 5.6.1, we have $\frac{\pi_p}{1-\pi_p} > \prod_{i\in\{1,...,n\}} \frac{f_{\bar{p}}(t_{ip})}{f_p(t_{ip})}$ and $\frac{\pi_q}{1-\pi_q} < \prod_{i\in\{1,...,n\}} \frac{f_{\bar{q}}(t_{iq})}{f_q(t_{iq})}$ which is true for all $\mathbf{t}$ with $(**)$.

Case 3: $f(\mathbf{v}) = \bar{p}q$. This case is analogous to Case 2.

Case 4: $f(\mathbf{v}) = \bar{p}\bar{q}$. Then, for all $\mathbf{t}$ which leads to $\mathbf{v}$ under informative voting, $(***)$ $|i : t_{ir} \geq \hat{t}_r| < m_r$ for each $r \in \{p, q\}$. By efficiency of informative voting, $\bar{p}\bar{q}$ is efficient given any type profile satisfying $(***)$. By Lemma 5.6.1, $\frac{\pi_r}{1-\pi_r} < \prod_{i\in\{1,...,n\}} \frac{f_{\bar{r}}(t_{ir})}{f_r(t_{ir})}$ for each $r \in \{p, q\}$.

Conversely, let (5.4.1) hold for some $m_p, m_q \in \{1, ..., n\}$ and some $\hat{t} \in (0, 1)^2$. Consider the quota rule f with thresholds m_p, m_q and a given type profile $\mathbf{t} \in ([0, 1]^2)^n$. Supposing informative voting with cut-point $\hat{t} \in (0, 1)^2$, let the resulting voting profile be $\mathbf{v}$. We want to show that $f(\mathbf{v})$ is efficient for $\mathbf{t}$. We consider four cases:

Case 1: $|i : t_{ir} \geq \hat{t}_r| \geq m_r$ for each $r \in \{p, q\}$. Then, $f(\mathbf{v}) = pq$. Since (5.4.1) holds, $\frac{\pi_r}{1-\pi_r} > \prod_{i\in\{1,...,n\}} \frac{f_{\bar{r}}(t_{ir})}{f_r(t_{ir})}$ for each r which by Lemma 5.6.1 implies $E(u(pq, S)|\mathbf{t}) > E(u(p\bar{q}, S)|\mathbf{t}) > E(u(\bar{p}\bar{q}, S)|\mathbf{t})$ and $E(u(pq, S)|\mathbf{t}) > E(u(\bar{p}q, S)|\mathbf{t}) > E(u(\bar{p}\bar{q}, S)|\mathbf{t})$. So, pq is efficient given $\mathbf{t}$.

Case 2: $|\{i : t_{ip} \geq \hat{t}_p\}| \geq m_p$ and $|i : t_{iq} \geq \hat{t}_q| < m_q$. Then, $f(\mathbf{v}) = p\bar{q}$. Since (5.4.1) holds, $\frac{\pi_p}{1-\pi_p} > \prod_{i\in\{1,...,n\}} \frac{f_{\bar{p}}(t_{ip})}{f_p(t_{ip})}$ and $\frac{\pi_q}{1-\pi_q} < \prod_{i\in\{1,...,n\}} \frac{f_{\bar{q}}(t_{iq})}{f_q(t_{iq})}$. By Lemma 5.6.1, $E(u(p\bar{q}, S)|\mathbf{t}) > E(u(\bar{p}\bar{q}, S)|\mathbf{t})$, $E(u(p\bar{q}, S)|\mathbf{t}) > E(u(pq, S)|\mathbf{t}) > E(u(\bar{p}q, S)|\mathbf{t})$. So, $p\bar{q}$ is efficient given $\mathbf{t}$.

Case 3: $|\{i : t_{ip} \geq \hat{t}_p\}| < m_p$ and $|\{i : t_{iq} \geq \hat{t}_q\}| \geq m_q$. This case is analogous to case 2.

Case 4: $|\{i : t_{ir} \geq \hat{t}_r\}| < m_r$ for each $r \in \{p, q\}$. Then, $f(\mathbf{v}) = \bar{p}\bar{q}$. Since (5.4.1) holds, $\frac{\pi_r}{1-\pi_r} < \prod_{i\in\{1,...,n\}} \frac{f_{\bar{r}}(t_{ir})}{f_r(t_{ir})}$ for each r which by Lemma 5.6.1 implies that $E(u(\bar{p}\bar{q}, S)|\mathbf{t})$ is maximal. Then, $\bar{p}\bar{q}$ is efficient for $\mathbf{t}$. $\square$

References to Part II

Ahn, D. S., Oliveros, S. (2012) Combinatorial Voting. *Econometrica* 80 (1): 89-141

Ahn, D. S., and S. Oliveros, (2011). The Condorcet Jur(ies) Theorem. Mimeo, Berkeley University.

Austen-Smith, D., Banks, J. (1996) Information aggregation, rationality, and the Condorcet jury theorem. *The American Political Science Review* 90: 34-45

Austen-Smith, D., Feddersen, T. (2006) Deliberation, preference uncertainty and voting rules. *American Political Science Review* 100(2): 209-217

Austen-Smith, D., Feddersen, T. (2005) Deliberation and Voting Rules, in Austen-Smith, D. and J. Duggan (eds) *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, Berlin: Springer

Billingsley, P. (1995) *Probability and measure* Wiley, New York

Bovens, L., Rabinowicz, W. (2006) Democratic answers to complex questions: an epistemic perspective. *Synthese* 150(1): 131-153

Bozbay, I., Doetrich, F., Peters, H. (2011) Judgment aggregation in search for the truth. Mimeo, Maastricht University

Bozbay, I. (2012) Judgment aggregation in search for the truth: the case of interconnections. Mimeo, Maastricht University

Coughlan, P. (2000) In defense of unanimous jury verdicts: mistrials, communication and strategic voting. *The American Political Science Review* 94(2): 375-393

De Clippel G., Eliaz K. (2011) Premise versus outcome-based information aggregation. Mimeo, Brown University

Dietrich, F. (2006) Judgment aggregation: (im)possibility theorems. *Journal of Economic Theory* 126(1): 286-298

Dietrich, F. (2007) A generalised model of judgment aggregation. *Social Choice and Welfare* 28(4): 529-565

Dietrich, F. (2010) The possibility of judgment aggregation on agendas with subjunctive implications. *Journal of Economic Theory* 145(2): 603-638

Dietrich, F., List, C. (2007a) Arrow's theorem in judgment aggregation. *Social Choice and Welfare* 29(1): 19-33

Dietrich, F., List, C. (2007b) Strategy-proof judgment aggregation. *Economics and Philosophy* 23: 269-300

Dietrich, F., List, C. (2008) Judgment aggregation without full rationality. *Social Choice and Welfare* 31: 15-39

Dietrich, F., Mongin, P. (2010) The premise-based approach to judgment aggregation. *Journal of Economic Theory* 145(2): 562-582

Dokow, E., Holzman, R. (2010a) Aggregation of binary evaluations. *Journal of Economic Theory* 145(2): 495-511

Dokow, E., Holzman, R. (2010b) Aggregation of binary evaluations with abstentions. *Journal of Economic Theory* 145(2): 544-561

Duggan, J., Martinelli, C. (2001) A Bayesian model of voting in juries. *Games and Economic Behavior* 37(2): 259-294

Feddersen, T., Pesendorfer, W.(1997) Information aggregation and voting behaviour in elections. *Econometrica* 65(5): 1029-1058

Feddersen, T., Pesendorfer, W. (1998) Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. *The American Political Science Review* 92(1): 23-15

Gerardi, D., (2000) Jury verdicts and preference diversity. *American Political Science Review* 94: 395-406

Goertz, J. M., Maniquet, F. (2011) On the informational efficiency of simple scoring rules. *Journal of Economic Theory*, forthcoming

Guilbaud, G. (1952) Les théories de l'intérêt général et le problème logique de l'agrégation. *Économie Appliquée* 5: 501-584

List, C. (2005) The probability of inconsistencies in complex collective decisions. *Social Choice and Welfare* 24(1): 3-32

List, C., Pettit, P. (2002) Aggregating sets of judgments: an impossibility result. *Economics and Philosophy* 18(1): 89-110

List, C., Pettit, P. (2011) *Group Agency: The Possibility, Design and Status of Corporate Agents*. Oxford University Press

List, C., Polak, B. (2010) Introduction to judgment aggregation. *Journal of Economic Theory* 145(2): 441-466

McLennan, A. (1998) Consequences of the Condorcet Jury Theorem for beneficial information aggregation by rational agents. *American Political Science Review* 92 (2): 413418

Meirowitz, A. (2002) Informative voting and Condorcet jury theorems with a continuum of types. *Social Choice and Welfare* 19:219–236

Nehring, K., Puppe, C. (2002) Strategy-proof social choice on single-peaked domains: possibility, impossibility and the space between. *Working paper*, University of California at Davis

Nehring, K., Puppe, C. (2007) The structure of strategy-proof social choice. Part I: General characterization and possibility results on median spaces. *Journal of Economic Theory* 135: 269-305

Nehring, K., Puppe, C. (2008) Consistent judgement aggregation:the truth-functional case. *Social Choice and Welfare* 31: 41-57

Nehring, K., Puppe, C. (2010) Abstract Arrovian aggregation. *Journal of Economic Theory* 145(2): 467-494

Nozick, R. (1981) *Philosophical explanations.* Oxford University Press

Pivato, M. (2011) Voting rules as statistical estimators. Mimeo, Trent University, Canada

Wilson R (1975) On the theory of aggregation. *Journal of Economic Theory* 10: 89-99

Samenvatting

Dit proefschrift is een verzameling van artikelen op twee verschillende onderzoeksgebieden in het veld van de economische theorie, namelijk onderhandelen en oordeelaggregatie.

Het eerste gedeelte van het proefschrift behandelt onderhandelingsproblemen. In het klassieke twee-persoons onderhandelingsprobleem van Nash (1950) proberen twee spelers tot een overeenkomst te komen over hoe een bepaald goed te verdelen, waarbij elke speler eenzijdig de onenigheidsuitkomst kan afdwingen als onderhandelingen mislukken. Navolgend op Vartiainen (2007) kijken we naar onderhandelingsproblemen waarbij geen exogene onenigheidsuitkomst bekend is. Een onderhandelingsuitkomst kent een tweetal uitkomsten toe aan zo'n probleem, namelijk een compromisuitkomst en een onenigheidsuitkomst. De onenigheidsuitkomst kan functioneren als een spelers' mentale referentiepunt voor de compromisuitkomst, maar andere interpretaties zijn ook mogelijk. Onderhandelingstheorie met endogene onenigheid maakt het mogelijk om bestaande vormen van de Alternatieve Onenigheids Oplossing te bestuderen; een methode die vaak in de Angelsaksische algemene rechtssystemen wordt gebruikt om dure en langdurige rechtszaken te vermijden door vooraf een compromis te sluiten. In dit raamwerk bestuderen we de klassieke Kalai-Smorodinsky onderhandelingsoplossing en stellen we een uitbreiding voor. We identificeren het (uitgebreide) domein waarop de oplossing single-waarde is. We presenteren twee axiomatische karakteristieken op subverzamelingen van dit domein. De eerste karakteristiek is gebaseerd op een axioma genaamd Onafhankelijkheid van Non-Utopie Informatie (INU). INU is een sterke voorwaarde en het dwingt de onderhandelingsoplossing in sommige gevallen tot het negeren van non-utopie informatie. Onze tweede karakteristiek vervangt INU door drie andere axiomas, waarvan elk axioma een tegenhanger heeft in de klassieke onderhandelingstheorie.

Het tweede gedeelte van dit proefschrift behandelt oordeelaggregatieproblemen. We analyseren oordeelaggregatie vanuit het waarheid-opsporings en strategisch-kiezen perspectief. We modelleren de privé informatie van kiezers en bestuderen efficiënte informatie aggregatie en strategisch kiezen in een Bayesiaanse keuzespel omgeving. Dit perspectief is nieuw op het gebied van oordeelaggregatie theorie, maar bekend in de literatuur op het gebied van kiezen tussen twee alternatieven,

waarbij de onenigheid tussen kiezers veelal voortkomt uit informatieconflicten en niet door verschillende belangen. We analyseren de resulteerde strategische drijfveren en bepalen welke kiesregels leiden tot collectieve oordelen die effectief alle privé informatie gebruiken, aangenomen dat de kiezers een gezamenlijke preferentie hebben voor werkelijke collectieve oordelen. Hoewel we een gezamenlijk doel hebben in alle hoofdstukken van deel II, analyseren we een verschillend kader in elk hoofdstuk.

In hoofdstuk 3 beginnen we met het aannemen van twee onafhankelijke kwesties. Het blijkt dat men in de meeste, maar niet alle, situaties een 'quota regel' dient te gebruiken, welke voor elke kwestie beslist door middel van of het aantal 'ja'-keuzes bij de specifieke kwestie een bepaald quota overschrijdt. De details hangen af van de specifieke soort waarheid-opsporings preferenties.

In hoofdstuk 4 analyseren we het geval waarin de twee kwesties op de agenda logisch zijn verbonden. Een kiezers privé informatie kan inconsistent zijn, en daarom niet doorslaggevend in dit raamwerk. We karakteriseren de (zeldzame) situaties waarin kiesregels bestaan welke leiden tot collectieve oordelen die efficiënt gebruik maken van alle privé informatie, en de omgeving van deze regels.

In hoofdstuk 5 gaan we terug naar het geval waar de twee kwesties op de agenda onafhankelijk zijn. Als innovatie nemen we aan dat een kiezers' privé informatie betreffende een voorstel continu is in plaats van binair. Dit keer analyseren we de mogelijkheid van efficiënte informatie aggregatie met quota regels, en we concluderen dat dit niet altijd mogelijk is. We karakteriseren de situaties waarin dit mogelijk is volledig.

About the author

İrem was born in Bursa, Turkey, on December 3, 1981. She studied Chemical Engineering at Boğaziçi University, İstanbul between 2000 and 2004. After working as a chemical engineer for two years, she quit her job to attend the Master of Science program in Economics at İstanbul Bilgi University in 2006. She worked as a research assistant for the same program during 2007. Right after she got her degree, she moved to Maastricht, The Netherlands for her doctoral studies at Maastricht University. She worked under the supervision of Prof. dr. Franz Dietrich and Prof. dr. Hans Peters in the Department of Quantitative Economics. She spent one term at London School of Economics as a visiting Ph.D. student in 2011. İrem focuses on topics in game theory and social choice theory. This dissertation presents the results of her research between September 2008 and December 2011.