Argumentation in Multi-Issue Debates*

Ran Spiegler†
School of Economics, Tel Aviv University

May 2, 2004

Abstract

An axiomatic modeling approach to multi-issue debates is proposed. A debate is viewed as a decision procedure consisting of two stages: (1) an "argumentation rule" determines what arguments are admissible for each party, given the "raw data", depending on the issue or set of issues under discussion; (2) a "persuasion rule" determines the strength of the admissible arguments and selects the winning party. Persuasion rules are characterized for various alternative specifications of the argumentation rule. These characterizations capture rhetorical effects that we sometimes encounter in real-life multi-issue debates.

KEYWORDS: Debates, Rhetoric, Argumentation, Procedural Invariance. JEL No. D79.

*I am grateful to Nuffield College and the Institute for Advanced Study at Princeton for their stimulating hospitality while this project was undertaken. I also thank Elhanan Ben-Porath, Eddie Dekel, Michele Piccione and the referees for useful comments. I am especially indebted to Kfir Eliaz, Barton Lipman and Ariel Rubinstein for detailed comments on an earlier draft. Financial support from the Arthur Goodhart Fund is gratefully acknowledged.

†School of Economics, Tel Aviv University, Tel Aviv, Israel 69978. E-mail: rani@post.tau.ac.il. URL: http://www.tau.ac.il/~rani.

1 Introduction

Deliberation over collective decisions often takes the form of debates. What distinguishes debates from other collective-choice procedures is the element of rhetoric. Like voting mechanisms, debates aggregate "raw data" such as individual motives or differential information. However, unlike voting, the resolution of debates depends not only on the raw data elicited from debaters, but also on *quality of the arguments* in which the data are couched. "Irrelevant" information can appear convincing in the hands of a skilled rhetorician, whereas high-quality information can turn ineffective in the hands of an inept one.

Elements of rhetoric play an important role in various decision processes, such as litigation, parliamentary legislation and public-opinion formation. Nevertheless, except for two very recent works, Glazer and Rubinstein (2000) and Aragones et. al. (2001), I know of no attempt by economic theorists to formalize aspects of rhetoric and argumentation. For a survey of these papers, as well as other works in the intersection of economics and language, see Lipman (2002). (In the large literature on strategic information transmission, communication games are sometimes described as "debates", but this literature is not concerned with rhetorical aspects of communication.)

What makes debates especially hard to model is their relative lack of explicit structure, comparing with mechanisms such as voting or even bargaining. The "laws of rhetoric", which determine the legitimacy and strength of arguments are seldom clear-cut. This suggests that an axiomatic approach, which does not commit to a particular game-theoretic structure, might be useful at this stage of the research program.

This paper is a first step in an attempt to develop an axiomatic modeling approach to debates, somewhat in the spirit of social choice theory. Here are some of the questions that I aim to analyze using this approach: What determines the strength of arguments? Can certain decision biases originate from rhetorical effects? How does the procedure of debates affect rhetorical conventions?

Needless to say, I do not pretend to offer a "general model of debates". Rather, I will adopt the axiomatic approach to study a particular aspect of rhetoric that arise in the context of *debates over multiple issues.* People usually hold conflicting opinions over more than one issue. When they enter a debate they may discuss one or more of these issues. Moreover, the decision whether to introduce a new issue into the discussion is deliberate, as it may affect the debate's outcome. My task in this paper will be to analyze the implications of this consideration on the outcome of multi-issue debates.

I will view multi-issue debates as two-stage procedures for selecting a winning party on the basis of raw data. An *argumentation rule* transforms raw data into sets of available arguments for each party, depending on the issue or set of issues under discussion. In other words, it captures the rhetorical convention that determines which pieces of information count as admissible arguments in a particular debate. A *persuasion rule* is a function that selects a winning party, given the parties' admissible arguments. In other words, it captures the rhetorical convention that determines the strength of different arguments.

To motivate our discussion, imagine a political debate between a "right-wing" party and a "left-wing" party, who disagree over a pair of issues: death penalty and abortion rights. There are four possible "worldviews" (combinations of yes/no opinions on these issues). Two of these worldviews are held by the debating parties. The right-wing (left-wing) party approves (disapproves) of the death penalty and disapproves (approves) of abortion rights. "Raw data" consists of desirable attributes that each of the *four* possible worldviews may or may not possess. "Crime reduction", "sanctity of human life", or "consistency with constitutional law" are examples of such attributes.

As mentioned above, the debaters can discuss these two issues independently or conjointly. To make this distinction concrete, imagine a conference devoted to the public debate between the right-wing and left-wing parties over death penalty and abortion rights. The conference can be conducted in two ways:

1. "*Parallel sessions*": Different issues are discussed separately inside different halls in front of different audiences.

2. "*Plenary session*": both issues are discussed inside one large hall in front of all conference attendants.

The conference thus generates three possible debates from the same raw data: one debate exclusively devoted to death penalty, one debate exclusively devoted to abortion rights, and one debate over both subjects in conjunction. The raw data underlying each of these debates is the same. Nevertheless, it may endow the parties with different sets of available arguments in the different sessions.

For example, the "sanctity of human life" argument may be available only to the left-wing (right-wing) party in the death-penalty (abortion-rights) parallel session. Thus, when people disagree over multiple issues, the notion of a supporting argument is not obvious because the same raw data can provide a particular supporting argument for either party, depending on the "session" - i.e., the issue or combination of issues in question.

This is where the dichotomy between argumentation rules and persuasion rules turns out to be useful. The argumentation rule determines the set of arguments that are available to each party in each of the three sessions. The persuasion rule then selects a winning party for each session on the basis of the arguments that are available to each party in that session. The core assumption in the model will be a consistency condition that links the resolution of the parallel sessions to the resolution of the plenary session.

Formally, the model can be outlined as follows. To keep things simple, I assume that there are only two relevant issues. Every issue admits a yes/no opinion. Two parties hold opposite views on both issues. There is a universal set M of desirable attributes. A state (i.e., the "raw data") is a function that assigns a subset of M to each of the four possible combinations of yes/no opinions on the two issues. A debate is a pair of subsets of M, one for each party. An argumentation rule D is a function that assigns a triple of debates (one parallel session per issue, as well as one plenary session) to every state. A persuasion rule r is a function that assigns a winning party to every debate. The pair (D, r) is called a *multi-issue debate model* (MDM).

This formalism contains an implicit assumption: the persuasion rule is invariant to the issue or combination of issues in dispute. In other words, the rhetorical convention that determines the strength of arguments is the same for all sessions. All the differences among sessions are captured by the argumentation rule.

In addition, the MDM is required to satisfy two axioms: (1) "Procedural Invariance" - for every state, if the same party wins both parallel sessions, he must win the plenary session as well; (2) "Free Disposal" - a debater cannot be harmed by having more available arguments. Procedural Invariance is the key axiom in the model. It captures a sort of "procedural equilibrium" in the way people conduct multi-issue debates. It means that parties are able to discuss a particular issue, say the death penalty, such that no party would have a strong incentive to bring the other issue (abortion rights) into the discussion and turn it into a grand debate about the parties' entire worldviews.

The main exercise that I carry out is to characterize the persuasion rules that satisfy these axioms for two alternative specifications of the argumentation rule D:

"**Positive argumentation**". According to this argumentation rule, parties argue by listing desirable attributes of worldviews that are consistent with their position in a debate. Under this rule, only a constant persuasion rule satisfies the two axioms - i.e., the same party must win all debates.

"**Negative argumentation**". According to this argumentation rule, parties argue by listing desirable attributes of worldviews that are inconsistent with the opponent's position in the debate. Under this rule, the two axioms allow for non-constant persuasion rules, but these involve the following bias. There exists a party k and a subset of attributes M^*, such that party k wins any debate whenever his argument set intersects M^*. Moreover, M^* is uniquely characterized as the smallest subset, whose exclusion from k's argument set guarantees the victory of k's opponent whenever the latter's argument set is M.

These results demonstrate how different kinds of argumentation can lead to different resolutions of debates, given the same raw data. As we shall see, the characterization results, and the reasoning behind them, highlight rhetorical effects that we sometimes encounter in real-life multi-issue debates. I should emphasize that I do not view the model as a predictive theory of how real-life debates are resolved. Rather, the axiomatic method enables me to highlight some rhetorical considerations that we observe in real-life debates.

The general idea of breaking up rhetorical conventions into a rule that determines the admissibility of arguments and a rule that determines their persuasiveness has a precedent in Glazer and Rubinstein (2001). However, the methodology of Glazer and Rubinstein (2001) is game-theoretic, whereas I adopt an axiomatic approach. The particular problems addressed in the two papers are also very different. The fact that such different formal outlooks rely on this general idea suggests its fruitfulness for thinking about the subject of rhetoric and argumentation.

The paper proceeds as follows. Section 2 presents the MDM. Section 3 presents the characterization results. In Section 4, I analyze an extended MDM, in which debaters can use more than just one type of argument. I show that when parties can use both positive and negative argumentation, the resolution of debates is solely determined by the negative arguments. This result demonstrates the extended model's potential for analyzing the problem of what makes certain types of argument more effective rhetorically than others. Section 5 offers concluding remarks and discusses related literature.

2 A Model of Multi-Issue Debates

I begin with a bit of terminology. Let α and β be a pair of *issues*. There are four possible yes/no opinions on these issues: (yes to α, yes to β), (no to α, no to β), (yes to α, no to β) and (no to α, yes to β). I refer to each of these four multi-issue opinions as a *worldview*. Two parties, 1 and 2 (party j's rival is referred to as party $-j$) hold opposite opinions on both issues. E.g., party 1's worldview is (yes to α, no to β) and party 2's worldview is (no to α, yes to β). The worldview that agrees with party j on issue α and with party k on issue β is denoted jk.

Let M be a finite set of desirable *attributes*. A *debate* $d = (A_1, A_2)$ is a pair of attribute subsets, where $A_k \subseteq M$ is party k's *argument set* in the debate. Since the model is not game-theoretic, there is no explicit distinction between the arguments that are available to the parties and the arguments that they actually raise in the course of the debate. However, I will consistently use the former interpretation.

A *state* is a function $\omega : \{1,2\}^2 \to 2^M$ that assigns a subset of attributes to every worldview. For every $j, k \in \{1,2\}$, ω_{jk} is the set of attributes assigned to the worldview that agrees with party j over issue α and with party k over issue β. (The function ω need not be partitional: an attribute can be assigned to no worldview, or to several worldviews at the same time.) The set of all states is denoted Ω.

So far, our terminology contains two levels: debates (i.e., the arguments available to parties) and states (i.e., the raw data underlying the debates), but these two levels are not linked together in any way. The model we are about to construct will provide the link between raw data and the arguments that parties can raise in debates on the basis of the raw data.

A *multi-issue debate model* (MDM) is a pair (D, r), where D and r are referred to as the *argumentation rule* and *persuasion rule*, respectively:

The argumentation rule is a function $D : \Omega \to (2^M)^2 \times (2^M)^2 \times (2^M)^2$, which assigns a triple of debates $D(\omega) = (d_\alpha(\omega), d_\beta(\omega), d_{\alpha\circ\beta}(\omega))$ to every state $\omega \in \Omega$. Why a *triple* of debates? When two parties disagree over a pair of issues, they may discuss them independently or conjointly. Thus, every state generates three possible debates: a "parallel session" $d_\alpha(\omega)$ devoted to issue α, a "parallel session" $d_\beta(\omega)$ devoted to issue β, and a "plenary session" $d_{\alpha\circ\beta}(\omega)$ on both issues in conjunction. In the parallel session devoted to issue $a \in \{\alpha, \beta\}$, the parties argue their opinions on a. In the plenary session, they argue their entire worldviews.

The Persuasion rule is a function $r : (2^M)^2 \to \{1, 2\}$, which assigns a winning party $r(d) \in \{1, 2\}$ to every debate $d = (A_1, A_2)$. Note that r need not be symmetric - i.e., it is not required that $r(A, B) = -r(B, A)$ whenever $A \neq B$. An asymmetric persuasion rule may be sensitive to the parties' names, in addition to their argument sets. The role of this extra degree of freedom will become clear in the sequel.

The argumentation rule captures the rhetorical convention that determines which arguments are admissible for every party in every debate, given the raw data. The persuasion rule captures the rhetorical convention that determines the relative strength of the parties' admissible arguments. In other words, an MDM is a two-stage procedure. The first stage carries us from raw data to arguments, and the second stage carries us from arguments to the outcome of the debate.

An important feature of this model is that *the two stages are independent.* The persuasion rule is the same for every debate, whether the debate is a plenary session or a parallel session on any particular issue. All differences between sessions are reflected in the argumentation rule: different sessions can admit different argument sets, given the same underlying state. Of course, this is a strong assumption - in real-life debates, the strength of a particular argument sometimes seems to depend on the debated issue.

Note that the outcome of debates in this model is defined as the identity of the winning *party.* Thus, this is an adversarial model of debates, which fits televised debates between political candidates, or debating clubs. In many real-life debates, however, the important thing about the outcome is not who wins the debate, but what position wins. In other situations, debaters do not care about winning at all, and they use the debate format to achieve a better understanding of the underlying issues. Finally, the objective of debating parties is often changing the opponent's convictions, rather than just winning. This aspect of real-life debates is not captured at all by the present model.

To illustrate the concept of an argumentation rule, let us return to the "political debate" example of the introduction. Let α and β stand for "death penalty" and "abortion rights". Party 1 (right-wing) approves of the former and disapproves of the latter. Party 2 (left-wing) holds opposite opinions on both issues. Suppose that in the state ω, the attribute "sanctity of human life", denoted $\bar{m}$, is assigned only to the worldview that agrees with the left-wing party on the death penalty and with the right-wing party on abortion rights. That is, $\bar{m} \in \omega_{21}$ and $\bar{m} \notin \omega_{11}, \omega_{12}, \omega_{22}$.

In this case, it is not obvious a priori that the left-wing party can use "sanctity of human life" as an argument in the parallel session on the death penalty. Can people, when discussing the death penalty, list attributes that are somehow consonant with their opinion on the death penalty, but not with their entire worldview (which also consists of their opinion on abortion rights)? That depends on the rhetorical convention for what counts as an admissible argument in the debate. This is where the argumentation rule D comes into play: it determines whether $\bar{m}$ belongs to party 2's argument set in the parallel session $d_\alpha(\omega)$.

I shall discuss further the concept of an argumentation rule at the beginning of Section 3.

The possibility of two different procedures for conducting multi-issue debates - "parallel sessions" vs. "plenary session" - naturally raises the question of how the debates' procedure affects their outcome.

Let us impose two axioms on the MDM (D, r). The first axiom requires the MDM to satisfy a "free disposal" property:

Axiom 1 (Free Disposal) *If* $r(A_1, A_2) = k$, $A_k \subseteq B_k$ *and* $B_{-k} \subseteq A_{-k}$, *then* $r(B_1, B_2) = k$.

Free Disposal (**FD** henceforth) means that if a certain party wins a debate, then expanding his argument set or shrinking his opponent's cannot reverse the outcome. This axiom is particularly attractive if we insist on interpreting A_k as the set of arguments that are *available* to party k in the debate $d = (A_1, A_2)$, rather than as the set of arguments that he actually raises during the debate. Under the former interpretation, FD means that parties never raise arguments that would make them lose the debate. (Of course, since the MDM is not a game-theoretic model, the distinction between available and actually used arguments is quite artificial.)

The second axiom I shall impose on the MDM (D, r) requires a certain form of "Procedural Invariance":

Axiom 2 (Procedural Invariance) *For every* $\omega \in \Omega$, *if* $r[d_\alpha(\omega)] = r[d_\beta(\omega)] = k$, *then* $r[d_{\alpha\circ\beta}(\omega)] = k$.

Procedural Invariance (**PI** henceforth) is a consistency requirement, somewhat in the spirit of "single-profile conditions" in social choice theory. It means that given a state of the world, if the same party wins both "parallel sessions", he would win the "plenary session" as well.

PI is a *stability* property of the rhetorical conventions captured by (D, r). If PI were violated, then there would exist a state, in which one party prefers to discuss the issues separately, while his opponent prefers to discuss them conjointly. The parties would thus fight over procedure instead of substance and in this sense, the rhetorical conventions would cease to be stable. I will be interested in multi-issue debates models that do not suffer from this instability.

PI captures a sort of "procedural equilibrium" in multi-issue debates. People normally disagree over many issues. What allows them to discuss a specific issue independently is some norm, according to which the debaters' views on other issues are irrelevant to the current debate. In order for this norm to be stable, no party should have a strong incentive to introduce other issues into the discussion and turn it into a grand debate over the parties' entire worldviews.

The role of multiple issues

The MDM can be viewed as an extension of a more standard attribute-based model, in which the two parties disagree over a single yes/no issue; a state is a function ω that assigns subsets of M to each every opinion on the issue, where ω_j is the set of attributes of party j's opinion; a debate continues to be a pair of attribute sets (A_1, A_2), but now $A_j = \omega_j$, such that a debate and a state are the same thing.

In such a model, the "persuasion rule" r is a function that acts directly on the "raw data" ω and assigns a winning party $r(\omega) \in \{1, 2\}$ to every state ω. Since there is no distinction between raw data and arguments, there is no need for an "argumentation rule". Note that if we impose symmetry on this model (i.e., $r(A, B) = r(B, A)$ whenever $A \neq B$), it is reduced to a complete binary relation on 2^M.

As we saw earlier, the distinction between raw data and arguments becomes necessary when we turn from single-issue to multi-issue debates, because the notion of a supporting argument is not clear-cut anymore: different "sessions" are characterized by different sets of arguments, given the same raw data. Adding the argumentation-rule component to the standard attribute-based model allows us to deal with this complication created by the multiplicity of issues.

The general idea of viewing debates as two-stage procedures, consisting of an argumentation rule and a persuasion rule is not peculiar to multi-issue debates. One could surely construct interesting models of single-issue debates on the basis of the same idea. Multi-issue debates simply provide an environment, in which the need for a two-stage procedure arises naturally.

In principle, the MDM can easily be generalized for any number $I > 2$ of issues in disagreement between the two parties. In such an extended model, however, there are several ways to generalize the Procedural Invariance axiom. For example, we may insist on "plenary session" and "parallel sessions" as the only relevant procedures. In this case, each state generates $I + 1$ different sessions. (I parallel sessions on each issue independently, and one plenary session on all issues simultaneously.) Alternatively, we may wish to consider debates over any combination of issues. In this case, each state generates $2^I - 1$ different sessions. Such generalizations are left to be pursued by future research.

3 Analysis

This section characterizes persuasion rules that satisfy the Free Disposal and Procedural Invariance axioms, for two alternative specifications of the argumentation rule D.

Define the argumentation rule D^{pos} as follows:

1. $d^{pos}_{\alpha\circ\beta}(\omega) = (\omega_{11}, \omega_{22})$
2. $d^{pos}_{\alpha}(\omega) = (\omega_{11} \cup \omega_{12}, \omega_{21} \cup \omega_{22})$
3. $d^{pos}_{\beta}(\omega) = (\omega_{11} \cup \omega_{21}, \omega_{12} \cup \omega_{22})$

I will refer to D^{pos} as the "*positive argumentation*" rule. It captures the following rhetorical convention: an attribute serves as an argument supporting a party's position in a "session" if it is assigned at least to one worldview, which is consistent with this position. In other words, parties argue by saying what is good about the worldviews that are consistent with their position in the debate.

In the "plenary session", party k's argument set is simply ω_{kk}, the set of attributes assigned to his worldview. In the "parallel session devoted to issue α", party k's argument set is the union of ω_{kk} and ω_{kj}, as both attribute sets are assigned to worldviews that agree with party k on the issue α. Similarly, in the "parallel session devoted to issue β", party k's argument set is the union of ω_{kk} and ω_{jk}, as both attribute sets are assigned to worldviews that agree with party k on the issue β.

Now consider the following alternative argumentation rule D^{neg}:

1. $d^{neg}_{\alpha\circ\beta}(\omega) = (\omega_{11} \cup \omega_{12} \cup \omega_{21}, \omega_{22} \cup \omega_{12} \cup \omega_{21})$
2. $d^{neg}_{\alpha}(\omega) = d^{pos}_{\alpha}(\omega)$
3. $d^{neg}_{\beta}(\omega) = d^{pos}_{\beta}(\omega)$

I will refer to D^{neg} as the "*negative argumentation*" rule. It captures a different rhetorical convention than D^{pos}: parties do not argue by saying what is desirable about their own position in the debate, but rather what is desirable about the *negation of the opponent's position*. In other words, an attribute serves as a negative argument supporting party k's position in a session if it is assigned to a worldview that is inconsistent with party $-k$'s position in that session.

In the parallel sessions, D^{pos} and D^{neg} coincide because the parties argue about a single yes/no issue, wherein party k's position is precisely the negation of party $-k$'s position. In the plenary session, D^{pos} and D^{neg} differ because party k's worldview is *not* the negation of party $-k$'s worldview.

To illustrate these argumentation rules, let $M = \{m, n, p\}$ and construct the state ω as follows: $\omega_{11} = \omega_{12} = \{m\}$, $\omega_{21} = \{m, p\}$, $\omega_{22} = \{n\}$. It may be useful to present the state in the form of a matrix (ω_{jk}):

$$\begin{matrix} \{m\} & \{m\} \\ \{m,p\} & \{n\} \end{matrix} \qquad \text{(Table 1)}$$

Then, $d_{\alpha}^{pos}(\omega) = d_{\alpha}^{neg}(\omega) = (\{m\}, M)$, $d_{\beta}^{pos}(\omega) = d_{\beta}^{neg}(\omega) = (M, \{m, n\})$, $d_{\alpha\circ\beta}^{pos}(\omega) = (\{m\}, \{n\})$ and $d_{\alpha\circ\beta}^{neg}(\omega) = (\{m, p\}, M)$.

We see that the same raw data gives rise to different argument sets under different argumentation rules. Each argumentation rule captures a different kind of rhetorical manipulation of raw data. Under D^{pos}, each party argues by saying what is good about the worldviews that are consistent with his position in the debate. Under D^{neg}, each party argues by saying what is good about the worldviews that are inconsistent with the opponent's position in the debate.

3.1 Characterizing r under "Positive Argumentation"

Let us first explore the implications of FD and PI on the persuasion rule, when the argumentation rule is $D = D^{pos}$.

Proposition 1 *If (D^{pos}, r) satisfies PI and FD, then r must be constant - i.e., there exists $k \in \{1, 2\}$, such that $r(A, B) = k$ for every $A, B \subseteq M$.*

Proof. For notational ease, denote $D^{pos} = D$. Let $r(A, B) = 1$, w.l.o.g. Consider the following state ω: $\omega_{22} = B$ and $\omega_{11} = \omega_{12} = \omega_{21} = A$. Then, $d_{\alpha\circ\beta}(\omega) = (A, B)$ and $d_{\alpha}(\omega) = d_{\beta}(\omega) = (A, A \cup B)$, hence $r[d_{\alpha}(\omega)] = r[d_{\beta}(\omega)] \equiv k$. By PI, $k = 1$. Thus, $r(A, A \cup B) = 1$. By FD, $r(A \cup B, A) = 1$.

Now construct the state ψ as follows: $\psi_{11} = B$ and $\psi_{22} = \psi_{12} = \psi_{21} = A$. Then, $d_{\alpha\circ\beta}(\psi) = (B, A)$ and $d_{\alpha}(\psi) = d_{\beta}(\psi) = (A \cup B, A)$. Since $r[d_{\alpha}(\psi)] = r[d_{\beta}(\psi)] = 1$, PI implies $r[d_{\alpha\circ\beta}(\psi)] = 1$. We have thus shown that $r(A, B) =$

$r(B, A)$ for every $A, B \subseteq M$. In particular, $r(\emptyset, M) = r(M, \emptyset) = 1$. By FD, $r(A, B) = 1$ for all $A, B \subseteq M$. ■

This is an impossibility result: under D^{pos}, PI and FD rule out a non-trivial persuasion rule. The intuition behind this result is simple. Positive Argumentation means that even when an argument does not support a party's entire worldview, it can support his position in each of the parallel sessions. This is what happens in the state ω given by Table 1: the attribute m is excluded from party 2's argument set in the plenary session ($m \notin \omega_{22}$), but it is included in his argument set in each of the parallel sessions. ($m \in \omega_{22} \cup \omega_{12}$ and $m \in \omega_{22} \cup \omega_{21}$.)

Suppose that r is not constant. Then, there is a state ω and a party k, such that k loses the plenary session yet wins each of the parallel sessions, thanks to an argument that supports his opinion on any issue in isolation, while failing to support his entire worldview. Therefore, Procedural Invariance is violated: party k strictly prefers to discuss issues in isolation, whereas party $-k$ prefers to discuss them simultaneously.

For instance, let $M = \{m, n\}$ and suppose that the persuasion rule satisfies $r(\{m\}, \{n\}) = 1$ and $r(\{m\}, \{m, n\}) = 2$. That is, adding m to party 2's argument set overturns the outcome in his favor. When $\omega_{11} = \omega_{12} = \omega_{21} = \{m\}$ and $\omega_{22} = \{n\}$, $d_{\alpha}^{pos}(\omega) = d_{\beta}^{pos}(\omega) = (\{m\}, \{m, n\})$ and $d_{\alpha\circ\beta}^{pos}(\omega) = (\{m\}, \{n\})$, hence party 2 loses the plenary session yet wins each of the parallel sessions, in contradiction to PI.

The rhetorical effect involved here can be likened to *promising the same dollar to two different people*: this trick works as long as the two people are kept apart, but not when they are both present in the same room. Similarly, in the parallel sessions, parties can raise positive arguments, which would be inadmissible in the plenary sessions because they do not support their entire worldviews. The availability of this rhetorical trick leads to the violation of Procedural Invariance.

3.2 Characterizing r under "Negative Argumentation"

The rhetorical trick that destabilizes non-trivial persuasion rules when the argumentation rule is D^{pos} stems from the discrepancy between what is arguable in parallel sessions and what is arguable in plenary sessions. Specifically, ω_{12} and ω_{21} (the attributes of the "mixed" worldviews that are held by none of the parties) enter into the parties' positive argument sets only in the parallel sessions. This discrepancy is attenuated under D^{neg} because ω_{12} and ω_{21} enter into the parties' argument sets in the plenary session, too.

For illustration, let $M = \{m, n\}$ and construct the state ω as follows: $\omega_{11} = \omega_{12} = \omega_{21} = \{m\}$ and $\omega_{22} = \{n\}$. Let us depict ω in matrix form:

$$\begin{matrix} \{m\} & \{m\} \\ \{m\} & \{n\} \end{matrix} \qquad \text{(Table 2)}$$

Then, $d^{pos}_{\alpha}(\omega) = d^{neg}_{\alpha}(\omega) = d^{pos}_{\beta}(\omega) = d^{neg}_{\beta}(\omega) = (\{m\}, M)$, $d^{pos}_{\alpha\circ\beta}(\omega) = (\{m\}, \{n\})$ and $d^{neg}_{\alpha\circ\beta}(\omega) = (\{m\}, M)$. Each party has different sets of positive arguments in the parallel and the plenary sessions, whereas each party has the same set of negative arguments in each session. Therefore, under the negative argumentation rule, parties do not have any procedural preference in this state because their arsenal of arguments is the same for each of the two procedures for running the multi-issue debate.

This example suggests that the forces that destabilize rhetorical conventions are weaker under negative argumentation than under positive argumentation. Consequently, our characterization of persuasion rules should be more permissive. This indeed turns out to be the case:

Proposition 2 *If (D^{neg}, r) satisfies PI and FD, and r is not a constant function, then there exists a unique party k (say, $k = 1$, w.l.o.g) and a unique non-empty subset $M^* \subset M$, such that:*

1. $r(A_1, A_2) = 1$ whenever $M^* \cap A_1 \neq \phi$.
2. $r(A_1, A_2) = 2$ whenever $M^* \cap A_1 = \phi$ and $A_2 = M$.

This result characterizes a class of biased persuasion rules. There exists a subset $M^* \subseteq M$, such that one party (say, party 1) wins any debate, as long as his argument set contains some attribute $m \in M^*$. This subset M^* is uniquely identified as the smallest set B, whose empty intersection with A_1 implies $r(A_1, M) = 2$. Note that it is possible that $r(A, B) = 1$ for some $B \subseteq M$, even when $M^* \cap A = \emptyset$. That is, empty intersection between party 1's argument set and M^* is a necessary, but insufficient condition for party 2's victory.

Proof. We will say that a non-empty subset $C \subseteq M$ is *decisive* in favor of party k if $r(A_1, A_2) = k$ whenever $A_k = M$ and $C \cap A_{-k} = \phi$. If C contains no proper non-empty subset, which is also decisive in favor of k, we will say that C is *minimally decisive* in favor of k.

If $r(\phi, M) = 1$, then by FD, $r(A, B) = 1$ for every $A, B \subseteq M$ - i.e., r is constant, a contradiction. Therefore, $r(\phi, M) = 2$, which means that M is decisive in favor of party 2. Thus, there exists a non-empty subset $M^* \subseteq M$, which is minimally decisive in favor of party 2. That is, $r(A, M) = 2$ whenever $M^* \cap A = \phi$; and for every $B \subset M^*$, there exists $A \subseteq M$, such that $B \cap A = \phi$, $(M^* - B) \cap A \neq \phi$ and $r(A, M) = 1$.

We have thus established the existence of a set M^*, which meets condition (2) in the statement of the proposition. Our next objective will be to establish that $r(A, M) = 1$ whenever $M^* \cap A \neq \phi$.

For notational ease, denote $D^{neg} = D$. Suppose that $r(A, M) = 2$ for some $A \subseteq M$, for which $M^* \cap A \equiv R \neq \phi$. By FD, $r(R, M) = 2$. Consider the state ω as follows: $\omega_{11} = \phi$, $\omega_{12} = M - M^*$, $\omega_{21} = R$ and $\omega_{22} = M$. It will be useful to present ω in matrix form:

$$\begin{matrix} \phi & M - M^* \\ R & M \end{matrix} \qquad \text{(Table 3)}$$

By the definition of D, $d_\alpha(\omega) = (M - M^*, M)$, $d_\beta(\omega) = (R, M)$ and $d_{\alpha\circ\beta}(\omega) = ((M - M^*) \cup R, M)$. Since M^* is decisive in favor of party 2, $r[d_\alpha(\omega)] = 2$. We already saw that $r[d_\beta(\omega)] = 2$ as well. By PI, $r[d_{\alpha\circ\beta}(\omega)] = 2$. By FD, $r(B, M) = 2$ for every $B \subseteq (M - M^*) \cup R$ - i.e., by every B satisfying $(M^* - R) \cap B = \phi$. Therefore, $M^* - R$ is decisive in favor of party 2, thus contradicting the fact that M^* is *minimally* decisive in favor of party 2.

It follows that whenever $M^* \cap A \neq \phi$, $r(A, M) = 1$ and by FD, $r(A, B) = 1$ for every $B \subseteq M$. We have thus shown that there exists a set M^*, which meets conditions (1) and (2) in the statement of the proposition. Two things remain to be shown: (i) there is no other minimally decisive set in favor of party 2, in addition to M^*; (ii) there is no decisive set in favor of party 1.

(i) Suppose that there are two different minimally decisive sets in favor of party 2, M_1^* and M_2^*. Neither set is a subset of the other set. Then, $r(M_1^* - M_2^*, M) = 2$ because M_2^* is decisive in favor of party 2. We have found a non-empty set $A = M_1^* - M_2^*$, such that $r(A, M) = 2$, although M_1^* is decisive in favor of party 2 and $A \cap M_1^* \neq \phi$ (by definition of A), a contradiction with what we have already proved.

(ii) Suppose that M_1^* and M_2^* are decisive in favor of parties 1 and 2, respectively. Consider the debate $(\{a\}, \{b\})$, where $a \in M_2^*$ and $b \in M_1^*$. By the above result, $a \in M_2^*$ implies $r(a, b) = 1$, whereas $b \in M_1^*$ implies $r(a, b) = 2$, a contradiction. ■

Proposition 2 implies that the resolution of plenary sessions is sensitive to the attributes of the "mixed" worldviews that are held by none of the parties. A change in ω_{12} or ω_{21} can reverse the outcome of the plenary session $d_{\alpha\circ\beta}^{neg}(\omega)$. To see why, suppose that $r[d_{\alpha\circ\beta}^{neg}(\omega)] = 2$, given some state ω. Then, $M^* \cap (\omega_{11} \cup \omega_{12} \cup \omega_{21}) = \phi$. Now, let ψ be identical to ω, except that $\psi_{12} = \omega_{12} \cup \{m\}$, where $m \in M^*$. Nothing in the attributes of the parties' own worldviews has changed. However, according to Proposition 2, $r[d_{\alpha\circ\beta}^{neg}] = 1$. Thus, there is a sense in which the resolution of plenary sessions depends on "irrelevant alternatives".

Interpreting Proposition 2

The bias identified by Proposition 2 recalls an interesting phenomenon that is sometimes encountered in real-life debates over proposals to depart from a status quo.

Imagine two parties, a "status-quo upholder" and a "reformer", who discuss in front of some audience the latter's proposal to reform university admission policy. The reformer's proposal abolishes the SAT and introduces affirmative action. This is a "closed-rule" debate: either the proposal is accepted in toto, or the status quo prevails. There is no room for compromise. The following exchange of arguments takes place. The reformer argues by listing desirable attributes of the proposed reform. The status-quo upholder counter-argues that a milder reform (which retains the SAT and only introduces affirmative action) would share some of these attributes. The audience judges that the debate has been won by status quo upholder.

What is going on here? Given the rules of the debate, the status-quo upholder's utterance cannot be construed as a suggested compromise, but rather as a rebuttal of the reformer's arguments. The meaning of his counter-argument is that the attributes mentioned by the reformer provide insufficient reason for abandoning the status quo in favor of his proposal for a drastic reform, because they could also support a milder reform. The audience's judgment means that it was a smashing counter-argument.

Let us try to interpret this effect in terms of the MDM. The two parties conduct a debate over two issues, SAT and affirmative action, under a "plenary session" procedure. The argumentation rule is D^{neg}. The reformer's worldview (against SAT, in favor of affirmative action) has several desirable attributes, but another worldview ("the milder reform") shares some of these attributes. Therefore, the status quo upholder's set of negative arguments contains some of the attributes that are assigned to the reformer's worldview. Some of these attributes belong to the subset M^*. Therefore, the status quo upholder wins the plenary session, even though his own worldview (in favor of SAT, against affirmative action) may fail to possess any desirable attribute.

Non-Emptiness

To see that the class of persuasion rules identified by Proposition 2 is not vacuous, construct the following persuasion rule: $r(A, B) = 2$ if and only if $M^* \subseteq B - A$, where M^* is a proper non-empty subset of M. Thus, *party 2 wins a debate if and only if M^* is exclusively contained in his argument set.* In the plenary session, this means that each of the attributes in M^* must be assigned to party 2's worldview, and to none of the other three worldviews.

It is easy to verify that (D^{neg}, r) satisfies FD. Let us check that (D^{neg}, r) also satisfies PI. Suppose that given ω, party 1 wins both parallel sessions. In particular, he wins the session devoted to issue α. Then, by the definition of r, $M^* \cap (\omega_{11} \cup \omega_{12}) \neq \phi$ or $M^* \subsetneq \omega_{22} \cup \omega_{21}$. Either possibility implies that party 1 wins the plenary session as well. Alternatively, suppose that given ω, party 2 wins both parallel sessions. Then, by the definition of r, $M^* \subseteq (\omega_{22} \cup \omega_{21}) - (\omega_{11} \cup \omega_{12})$ and $M^* \subseteq (\omega_{22} \cup \omega_{12}) - (\omega_{11} \cup \omega_{21})$. It follows that $M^* \subseteq \omega_{22}$ and $M^* \cap (\omega_{11} \cup \omega_{12} \cup \omega_{21}) = \phi$, hence party 2 wins the plenary session as well.

Illustrating the proof

Let us use an example with a three-element attribute set M, in order to illustrate how the proof of Proposition 2 works. Let $M = \{m, n, p\}$ and suppose that the subset $\{m, n\}$ is minimally decisive in favor of party 2 - i.e., $r(A, M) = 2$ whenever $B \cap A = \phi$, and there exists no proper non-empty subset of B with the same property. Suppose, contrary to Proposition 2, that $r(\{n\}, M) = 2$. Construct the state ψ as follows: $\psi_{11} = \phi$, $\psi_{12} = \{p\}$, $\psi_{21} = \{n\}$ and $\psi_{22} = M$. Once again, it is helpful to present the state in the form of a matrix:

$$\begin{matrix} \phi & \{p\} \\ \{n\} & \{m, n, p\} \end{matrix} \qquad \text{(Table 4)}$$

Then, $d_{\alpha}^{neg}(\psi) = (\{p\}, M)$, $d_{\beta}^{neg}(\psi) = (\{n\}, M)$ and $d_{\alpha\circ\beta}^{neg}(\psi) = (\{n, p\}, M)$. Since $r(A, M) = 2$ whenever $\{m, n\} \cap A = \phi$, $r[d_{\alpha}^{neg}(\psi)] = 2$. By assumption, $r[d_{\beta}^{neg}(\psi)] = 2$. By PI, $r[d_{\alpha\circ\beta}^{neg}(\psi)] = 2$. By FD, $r(A, M) = 2$ whenever $m \notin A$, in contradiction to the definition of M^* as a *minimally* decisive set in favor of party 2.

3.3 Summary of the Characterization Results

Let us summarize the results of this section. Only constant persuasion rules satisfy the Procedural Invariance and Free Disposal axioms when parties use "positive arguments". In contrast, when parties use "negative arguments", the axioms allow for non-trivial persuasion rules. The main accomplishment of Propositions 1 and 2 is that they capture familiar rhetorical effects and trace them to particular argumentation rules.

Proposition 1 relies on a "rhetorical trick" similar to "promising the same dollar to different people". Parties can raise arguments in the parallel sessions, which would be unavailable to them in the plenary session because they do not support their entire worldview. This effect is traced to the argumentation rule D^{pos} ("positive argumentation").

The biased persuasion rules characterized by Proposition 2 have a natural interpretation. One party can win a debate only if a certain set of arguments *exclusively* supports his position in the debate. In the plenary session, this allows his opponent to win even when the latter's own worldview lacks any desirable attribute. This effect is traced to the argumentation rule D^{neg} ("negative argumentation").

Proposition 2 is suggestive of asymmetric burden-of-proof assignments in real-life debates between reformers and status-quo upholders. The status-quo upholder can argue convincingly against a proposed reform by showing that some of the proposal's desirable attributes could be achieved by a milder reform, even when the status quo itself does not satisfy any desirable attribute. In contrast, the reformer can beat the status quo only if he shows that there is no milder way than his own proposal to meet certain desiderata. We saw that this effect can be interpreted in terms of Proposition 2.

It should be emphasized that I do not view the axiomatic model as a predictive theory. At this stage of the research agenda, it would be excessively bold to try fitting the results of this paper to concrete, real-life debates. In particular, the Procedural Invariance axiom is probably violated in real-life debates. For example, political candidates spend considerable energy on procedural strategizing before public debates. Nevertheless, I believe that our analysis of the logical implications of Procedural Invariance is instrumental. It highlights rhetorical considerations such as the trick of "promising the same dollar to different people", and how the availability of such a trick depends on whether one is using positive or negative argumentation.

4 Extension to Multiple Argument Types

So far, the MDM has allowed parties to use a single type of argument in debates. In real-life debates, however, we normally apply a multitude of argument types - positive and negative arguments, proofs, examples, analogies, and so forth - to the same raw data. This section extends the original multi-issue debate model, so as to accommodate multiple argument types.

An extended multi-issue debate model (D, r) introduces a single modification into the model of Section 2. A debate is now an *array* of pairs of argument sets $d = (d^h)_{h \in H}$, where H is a set of argument types and $d^h = (A_1^h, A_2^h)$ for every $h \in H$. Thus, $A_k^h \subseteq M$ is the set of type-h arguments in support of party k. Apart from this modification, the model is left untouched. The argumentation rule $D(\omega) = (D^h(\omega))_{h \in H}$ continues to assign a triple of debates (two "parallel sessions" and one "plenary session") to every state ω.

As to the axioms imposed on extended MDM's, the PI axiom remains intact and we only need to rewrite the FD axiom, the content of which remains essentially the same:

Axiom 3 (Modified-FD) *If* $r[(A_1^h, A_2^h)_{h \in H}] = k$, *and* $[A_k^h \subseteq B_k^h$ *and* $B_{-k}^h \subseteq A_{-k}^h]$ *for all* $h \in H$, *then* $r[(B_1^h, B_2^h)_{h \in H}] = k$.

The remainder of this section demonstrates the extended model's capacity to deliver insights into the question of relative rhetorical impact of different argument types. In Section 3, we characterized the persuasion rule for two alternative argumentation rules, "positive argumentation" and "negative argumentation". Now, suppose that parties can use both positive and negative arguments at the same time. I.e., $D = (D^{pos}, D^{neg})$, where D^{pos} and D^{neg} are as defined in Section 3.

Proposition 3 *Let $D = (D^{pos}, D^{neg})$ and assume that the extended multi-issue debate model (D, r) satisfies PI and Modified-FD. Then, $d^{neg}_{\alpha\circ\beta}(\omega) = d^{neg}_{\alpha\circ\beta}(\psi)$ implies $r[d_{\alpha\circ\beta}(\omega)] = r[d_{\alpha\circ\beta}(\psi)]$ for every $\omega, \psi \in \Omega$.*

Proof. Let $r[d_{\alpha\circ\beta}(\omega)] = 1$, w.l.o.g. Construct the state ω' as follows: $\omega'_{11} = \omega_{11}$, $\omega'_{22} = \omega_{22}$, $\omega'_{12} = \omega'_{21} = \omega_{12} \cup \omega_{21}$. Then, by the definition of D, $d^{pos}_{\alpha\circ\beta}(\omega') = d^{pos}_{\alpha\circ\beta}(\omega)$ and $d^{neg}_{\alpha\circ\beta}(\omega') = d^{neg}_{\alpha\circ\beta}(\omega)$. Thus, $d_{\alpha\circ\beta}(\omega') = d_{\alpha\circ\beta}(\omega)$, such that $r[d_{\alpha\circ\beta}(\omega')] = 1$. Again, by the definition of D, $d_\alpha(\omega') = d_\beta(\omega')$, such that $r[d_\alpha(\omega')] = r[d_\beta(\omega')] \equiv k$. By PI, $k = 1$. By the definition of d^{neg}, $d^{neg}_{\alpha\circ\beta}(\omega) = d^{neg}_{\alpha\circ\beta}(\psi)$ means that $\omega_{11} \cup \omega_{12} \cup \omega_{21} = \psi_{11} \cup \psi_{12} \cup \psi_{21}$ and $\omega_{22} \cup \omega_{12} \cup \omega_{21} = \psi_{22} \cup \psi_{12} \cup \psi_{21}$. Therefore, by construction, $d_\alpha(\omega') = d_\alpha(\psi)$ and $d_\beta(\omega') = d_\beta(\psi)$, such that $r[d_\alpha(\psi)] = r[d_\beta(\psi)] = 1$. By PI, $r[d_{\alpha\circ\beta}(\psi)] = 1$. ■

The rhetorical convention implied by this result is that negative arguments are more effective than positive arguments. In plenary sessions, the parties' negative arguments determine who wins the debate and the positive arguments are ignored. (In parallel sessions, positive and negative arguments are equivalent, hence the question of relative strength is meaningless for parallel sessions.) Carrying out similar exercises for other argument types is a challenge for future research.

5 Concluding Remarks

This paper proposed to model debates as a two-stage procedure for selecting a winning party on the basis of "raw data". First, an argumentation rule determines which arguments are admissible, given the raw data. Second, a persuasion rule determines the winning party, given the set of admissible arguments. Of course, there are many ways in which this general idea can be practiced; this paper merely applied it to the specific context of debates over a pair of issues.

In the model constructed for this special case, "raw data" is an assignment of desirable attributes to all possible worldviews. A debate is a pair of argument sets, one for each of the two debating parties. Multi-issue debates can be carried out in several ways: the issues can be discussed independently (in "parallel sessions") or conjointly (in a "plenary session"). Therefore, an argumentation rule is a function that maps raw data into argument sets in three debates: two parallel sessions (one per issue) and one plenary session. A persuasion rule assigns a winner to every debate, independently of whether the debate is a plenary session or a parallel session on any particular issue.

The multitude of procedures for discussing multi-issue disagreements naturally raises the question of whether the outcomes of the three sessions are somehow linked. A "Procedural Invariance" axiom postulates such a link. Coupled with a "Free Disposal" axiom, this leads to strong characterizations of the persuasion rule for various specifications of the argumentation rule. In an extended model, which incorporates multiple argument types, the characterizations are capable of ranking the different argument types by their rhetorical impact.

More examples of D

The framework proposed in this paper is capable of accommodating additional types of argumentation rules. Consider, for example, the following argumentation rule D^*:

1. $d^*_{\alpha\circ\beta}(\omega) = (\omega_{11}, \omega_{22})$
2. $d^*_{\alpha}(\omega) = (\omega_{11} \cap \omega_{12}, \omega_{21} \cap \omega_{22})$
3. $d^*_{\beta}(\omega) = (\omega_{11} \cap \omega_{21}, \omega_{12} \cap \omega_{22})$

This argumentation rule captures the following rhetorical convention: an attribute serves as an argument supporting a party's position in a "session" if it is assigned to *all the worldviews that are consistent with this position.* In the "plenary session", party k's argument set is simply ω_{kk}, the set of attributes assigned to his worldview. In the "parallel session devoted to issue α", party

k's argument set is the intersection of ω_{kk} and ω_{kj}. Similarly, in the "parallel session devoted to issue β", party k's argument set is the intersection of ω_{kk} and ω_{jk}. Thus, D^* is the same as D^{pos}, except that unions are replaced by intersections.

The argumentation rule D^* leads to the same impossibility result as D^{pos}.

Proposition 4 *If (D^*, r) satisfies PI and FD, then r must be constant - i.e., there exists $k \in \{1, 2\}$, such that $r(A, B) = k$ for every $A, B \subseteq M$.*

Proof. Let $r(A, B) = 1$, w.l.o.g. Consider the following state ω: $\omega_{22} = B$ and $\omega_{11} = \omega_{12} = \omega_{21} = A$. Then, $d^*_{\alpha\circ\beta}(\omega) = (A, B)$ and $d^*_\alpha(\omega) = d^*_\beta(\omega) = (A, A\cap B)$. It follows that $r[d^*_{\alpha\circ\beta}(\omega)] = 1$, and by PI, $r[d^*_\alpha(\omega)] = r[d^*_\beta(\omega)] = 1$. Therefore, $r(A, A \cap B) = 1$. Now consider the state ψ: $\psi_{11} = A$ and $\psi_{22} = \psi_{12} = \psi_{21} = B$. Then, $d^*_{\alpha\circ\beta}(\psi) = (A, B)$ and $d^*_\alpha(\psi) = d^*_\beta(\psi) = (A \cap B, B)$. Recall that $r(A, B) = 1$, hence $r[d^*_{\alpha\circ\beta}(\psi)] = 1$. By PI, $r[d^*_\alpha(\psi)] = r[d^*_\beta(\psi)] = 1$. Therefore, $r(A \cap B, A) = 1$. We have thus shown that $r(A, A \cap B) = r(A \cap B, A)$ for every $A, B \subseteq M$. In particular, $r(M, \emptyset) = r(\emptyset, M)$. By FD, $r(A, B)$ is constant for all $A, B \subseteq M$. ■

The proof essentially mimics the proof of Proposition 1. In both cases, the impossibility result is a consequence of the *gap between the sets of arguments that support a party's position in parallel sessions and the set of arguments that support his entire worldview.* The difference is that under D^{pos}, the former set contains the latter, whereas under D^*, the latter set contains the former.[1]

The Domain of D

In the MDM, the function D acts on the full domain Ω. This means that a priori, any attribute can be assigned to any worldview in some state. This is a very strong assumption because some attributes are logically linked to certain worldviews. E.g., one cannot assign the attribute "a woman's right over her body" to an anti-abortionist opinion. (In contrast, we can easily

[1] I would like to thank a referee for suggesting the argumentation rule D^*, as well as the observation that D^* implies the same impossibility result as D^{pos}, using the same kind of proof.

imagine an attribute such as "practical experience in other countries" being assigned to any opinion on any issue.)

We may attempt to overcome this interpretational difficulty by assuming that α and β do not stand for a fixed pair of issues, such that the same MDM is applied to multiple pairs of issues α and β, where the exact issues vary but the parties' identity remains fixed. This interpretation is quite artificial, in that it requires the rhetorical conventions captured by the MDM to be invariant to the set of issues $\{\alpha, \beta\}$ under dispute.

Whenever we examine domain restrictions in an axiomatic model, the question is which elements in the full domain play a more important role in the strong results obtained under the full-domain assumption. For example, in classical choice theory, preference profiles that contain Condorcet cycles are the "villain" in Arrow's impossibility result. Similarly, in the present model, states of the world such as the one given by Table 2 play a crucial role in the results. Specifically, pairs of states ω and ψ, such that $\omega_{11} = \psi_{11} = A$, $\omega_{22} = \psi_{22} = B$, $\omega_{12} = \omega_{21} = A$ and $\psi_{12} = \psi_{21} = B$, cause the strongest tension between parallel and plenary sessions. I expect domain restrictions which rule out such pairs to loosen the characterization results of this paper significantly. I leave the task of analyzing the model under restricted domains for future research.

Related Literature

The basic view of debates adopted here shares some features with Glazer and Rubinstein (2001), despite many differences in the formal outlook. Glazer-Rubinstein adopt a game-theoretic, mechanism-design approach. In their model, a debate is a mechanism for eliciting information from interested parties. A state is a binary vector of odd-length. The planner's objective is to know whether there are more 1's or more 0's in the state. However, the two agents who know the state have conflicting interests.

The planner's goal is to construct a mechanism that minimizes the expected number of erroneous decisions that he makes in equilibrium. The crucial assumption is that the complexity of the mechanism is bounded: there is an upper bound on the total number of messages that can be transmitted in the course of the game. If there were no bound, the planner's problem would be trivial. A debate is thus a mechanism of bounded complexity. Formally, it is very close to the standard notion of a game form. It consists of a procedural rule and a persuasion rule. A procedural rule is a description of the messages that players are allowed to choose at any decision node. A persuasion rule is a function that selects a winner for every terminal node in the game.

Glazer-Rubinstein show that any constrained-optimal mechanism must be sequential. Furthermore, the persuasion rule must exhibit a property that a priori looks like a rhetorical fallacy: there are two arguments, x and y, such that x is a winning argument when raised as a counter-argument against y, and yet y wins when raised as a counter-argument against x.

The distinction between procedural rules and persuasion rules in Glazer and Rubinstein (2001) resembles the distinction between argumentation rules and persuasion rules in the present paper. Of course, due to the very different contexts of the models, there is no direct formal link between their respective components.(I avoid using Glazer and Rubinstein's exact terminology because the term "procedural rule" would be misleading in a non-game-theoretic framework.)

Aragones et. al. (2001) address the following question: why is argumentation by analogy so effective rhetorically? They construct two alternative models of analogies, prove their equivalence and show that the problem of finding analogies is NP-complete. They explain the rhetorical effectiveness of analogies by this aspect.

As mentioned in the introduction, there is a big literature on communication games, which are sometimes referred to as debates. However, none of the papers with which I am familiar deals with questions of rhetoric and argumentation. A partial exception is Lipman and Seppi (1995). Although not explicitly focusing on argumentation, they formalize various notions of "partial provability" in communication games. They characterize the "amount of provability" that is sufficient for robust inferences. Once again, the interested reader is referred to Lipman (2002), which contains a more extensive discussion of the above-mentioned papers.

The question of how multiplicity of issues affects aggregation has also been studied in the framework of conventional social choice theory. This problem is particularly interesting when the issues are logically interconnected. See List (2004), for example. In general, several political scientists have recently attempted to forge a synthesis between social-choice-theoretic methods and alternative views of the democratic process, which emphasize its discursive, deliberative aspect. (See Dryzek and List (2002).) The present paper shares this motivation, despite the different formal approaches.

6 References

1. Aragones, E., I. Gilboa, A. Postlewaite, and D. Schmeidler (2001). "Rhetoric and Analogies." Mimeo.

2. Dryzek, J. and C. List (2002). "Social Choice Theory and Deliberative Democracy: A Reconciliation", *British Journal of Political Science* **33**, 1-28.

3. Glazer, J. and A. Rubinstein (2001). "Debates and Decisions: On a Rationale of Argumentation Rules." *Games and Economic Behavior* **36**, 158-173.

4. List C. (2004). "A Model of Path Dependence in Decisions over Multiple Propositions", *American Political Science Review*, forthcoming.

5. Lipman, B. (2002). "Language and Economics," in M. Basili, N. Dimitri, and I. Gilboa, eds., *Cognitive Processes and Rationality in Economics*, forthcoming.

6. Lipman, B. and D. Seppi (1995). "Robust Inference in Communication Games with Partial Provability." *Journal of Economic Theory* **66**, 370-405.