Pro Arguments, Con Arguments and Status Quo Bias in Multi-Issue Decision Problems

Ran Spiegler (*)

May 2000

ABSTRACT: *The public faces a choice between two alternatives: the status quo and a "comprehensive reform" proposal that departs from the status quo in several dimensions. Deliberation over the problem takes the form of a public multi-issue debate. The "reformists" argue that the proposed reform satisfies desirable features lacked by the status quo. The status quo supporters counter-argue that some of these features are obtainable by a reform that departs from the status quo in a single dimension only. This "modest reform" is not a feasible alternative in the debate and is used by the status quo camp merely as an argument. This type of argument is familiar from real-life debates and usually viewed as a (possibly decisive) pro-status-quo argument. Two questions arise: First, how could an argument that is based on an "irrelevant alternative" such as the modest reform be considered relevant to the debate? And second, how could this argument be construed as favoring the status quo (perhaps even decisively so)? I analyze these questions using a model of multi-issue debates, in which alternatives are represented by a repertory of pro and con arguments. The main results are: (1) Con arguments are more effective than pro arguments in determining the debate's resolution; (2) A status quo bias: certain con arguments against both the status quo and the "comprehensive reform" nevertheless decide in favor of the status quo.*

(*) Nuffield College, Oxford OX1 1NF, UK. ran.spiegler@nuffield.oxford.ac.uk ; http://www.nuff.ox.ac.uk/economics/people/spiegler.htm.

(**) I am grateful to Elchanan Ben-Porat, Eddie Dekel, Kfir Eliaz, Jacob Glazer, Gilat Levy, Barton Lipman, Volker Nocke, Michele Piccione, Kevin Roberts and Ariel Rubinstein, as well as seminar participants at Nuffield College, Boston University and Columbia University, for valuable conversations and comments. Financial support from the Arthur Goodhart Fund is gratefully acknowledged.

1. Introduction

ECONOMIC theory views preferences and information as the “raw material” of decision. When deliberating over this “raw material”, decision makers (especially in the case of collective choice) often go through a process of raising arguments and counter-arguments. The consequences of political, legal or academic debates depend to a large extent on the quality of the arguments involved. This latter notion seems to be at least partly independent of any aspect of raw preferences or information: a potentially “good” piece of information may be couched in a “bad” argument and lose its effectiveness, while apparently “irrelevant” information may have a decisive rhetorical impact. Nevertheless, the argument-based aspect of decision making has been ignored by economic theorists. This paper is a first attempt to address this question: what determines the effectiveness of various types of arguments and what is their effect on actual decisions?[1] It is a modest first step — I study argument-based elements of choice in the specific context of multi-issue decision problems. Real-life debates often involve disagreement over several issues. This aspect turns out to raise questions of argument-based choice which that not have arisen in a single-issue debate.

The main ideas of the paper can be illustrated by a simple example, a stylized version of a tax policy debate. The debate begins when Politician 1 proposes a comprehensive tax reform which departs from the current, long-existing tax system in two dimensions: capital income and inheritance. She argues that her proposal can fix the current system’s inefficiency. Politician 2, an upholder of the status quo, counter-argues as follows: “But this inefficiency could be fixed by a more modest reform which retains the current method of capital income taxation — so why do we need your drastic reform”?

For another real-life example, consider the following (stylized) academic debate. Scientist 1 confronts a seminar audience with a theory that relaxes several assumptions of a standard model. She argues that her theory can explain (under suitable choice of parameter values) some empirical facts that are anomalous from the standard model’s point of view. Scientist 2, an upholder of the status quo, counter-argues: “But at least part of these anomalies could be explained by a theory that relaxed only a single assumption of the standard model — so why do we need your more drastic departure”?

Both examples present a debate between two alternatives: status quo versus "drastic reform". The "reformists" argue for their proposal on the grounds that it has desirable features lacked by the status quo. The status quo camp responds with the counter-argument that some of these features could be satisfied by a more modest reform from the status quo. Now, this "modest reform" serves a purely rhetorical role; it is not a feasible alternative in the debate and is represented as such by neither the status quo nor the reformist camps;[2] it is therefore, in a sense, "irrelevant". Yet, this kind of counter-argument is familiar from real-life debates, in which it is used by status quo supporters to obstruct proposals for comprehensive reforms. We also have a strong intuition that the debaters' audience views it as a pro-status-quo argument (the status quo camp would not have used it otherwise!), perhaps even a decisive one. This raises two puzzles: How could an "irrelevant alternative" serve an effective argument? Why is it so naturally interpreted as a pro-status-quo argument?

I address these questions by constructing a simple model of two-issue debates, for which the tax policy debate serves as a leading example. The model is completely non-strategic — I study the effect certain arguments have on decisions, not the strategic use of these arguments. Denote two propositions on tax policy, "capital income tax should be imposed" and "inheritance tax should be imposed", by p and q. Let ~p and ~q be these propositions' negations and define conjunction and disjunction as usual (e.g., $p\wedge{\sim}q$ denotes "capital income tax should be imposed and inheritance tax should not", $p\vee q$ denotes "either capital income tax or inheritance tax should be imposed", etc.). We will look at three bilateral debates that can be constructed from these propositions — two single-issue debates, p vs. ~p and q vs. ~q, and the induced multi-issue debate, $p\wedge q$ vs. ${\sim}p\wedge{\sim}q$. The former and latter propositions in each debate are represented by a "reformist" and a "status quo" parties.

In this model, decision between propositions in a debate is argument-based. More specifically, it is based on argumentation by examples, which is pervasive in real-life debate. In the context of our tax debate policy, describing a possible scenario in which inheritance taxation will be efficient; referring to an already-existing efficient tax system with inheritance taxation; or citing a study that "shows" the efficiency of inheritance taxation; are all ways of arguing by examples in favor of inheritance taxation. When such examples exist, we say that q has an efficiency pro argument. A pro argument for a proposition is thus an example for an element in the class of tax policies advocated by the proposition, which satisfies a desirable feature. The idea of argumentation by examples carries an important formal implication: for every pair of propositions x and y, their disjunction $x\vee y$ has an efficiency pro argument if and only if x has one or y has one (see also footnote 5). This is referred to as the "union" property of the examples structure.

For expositional simplicity, assume that only two desirable features are relevant to the tax policy debate: efficiency and being a status quo (a feature dubbed "status-quo-ness"). Figure 1 illustrates tow possible situations that differ in their assignment of pro arguments to propositions. The left/middle/right matrices present the pro arguments' assignment to $p\wedge q$, $p\wedge\sim q$, $\sim p\wedge q$ and $\sim p\wedge\sim q$ / q and ~q / p and ~p. Efficiency and status-quo-ness are abbreviated by E and S. Note that the "union" property is preserved. E.g., in figure 1(b), efficiency is assigned to p,~p, q, $p\vee q$ and $\sim p\vee\sim q$, but not to ~q, whereas "status-quo-ness" is assigned to ~p, ~q and $\sim p\vee\sim q$, but not to p, q and $p\vee q$.

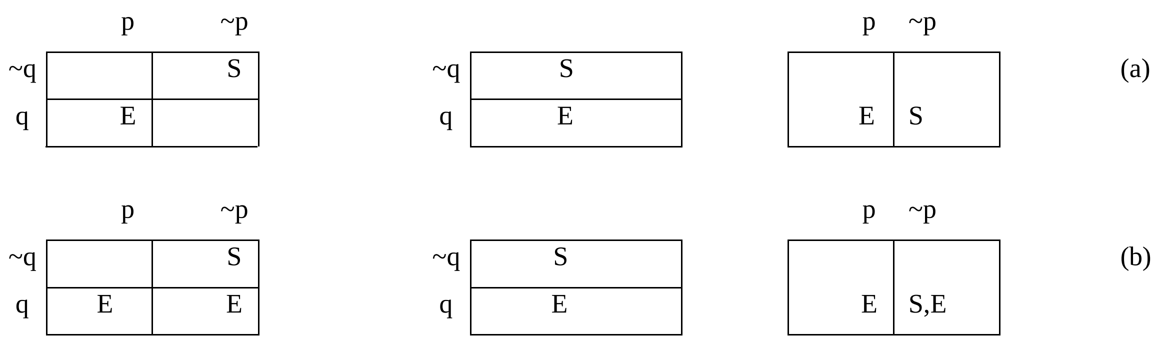

Figure 1

Now, a standard decision-theoretic approach would represent each claim in a debate by its set of pro arguments, thus reducing the debate to a problem of ranking two sets of features. E.g., in figure 1(b), p and ~p would be represented by {efficiency} and {efficiency, status-quo-ness}. This paper departs from standard approach by extending the repertory of arguments that represent claims, so as to include con arguments as well as pro arguments. In a real-life x vs. y debate, the x-party often argues not only for x but also against the rival claim, y. One common way of doing the latter is arguing for the negation of the rival claim's, ~y, even when $\sim y\neq x$. Thus, in the present model, the claims in the multi-issue debate, $p\wedge q$ and $\sim p\wedge\sim q$, are represented by their pro arguments as well as by those of their negations, $\sim p\vee\sim q$

and p∨q. Note that this extension is meaningful only in multi-issue debates — it is redundant in the single-issue debates, p vs. ~p and q vs. ~q, where each claim is the other's negation.

To sum up at this stage, a debate consists of two claims (either the reformist and status quo multi-issue claims p∧q vs. ~p∧~q, or their single-issue claims, p vs. ~p and q vs. ~q) and a function that assigns pro arguments to all relevant propositions and satisfies the "union" property. Whereas standard approach would consider only the parties' claims as "relevant propositions", in the present model their negations are relevant as well. An outcome function assigns a winner to every debate (either the reformist party or the status quo party).

The final major component in the model is also a consequence of the multiplicity of issues in the decision problem. Suppose that for a given assignment of pro arguments to propositions, the same party wins both single-issue debates. In this case, I assume that the same party wins the multi-issue debate as well. This "debate splitting invariance" property is particularly natural in legal debates: courts often justify a ruling in multi-issue cases by drawing on relevant single-issue precedents. In general, it is a minimal requirement of consistency between two ways to resolve a multi-issue debate: discussing each issue separately and confronting the parties' opinions on all issues simultaneously.

As it turns out, the combination of these components — argumentation by examples, the pro-argument/con-argument distinction, and debate splitting invariance (as well as a weak tie-breaking rule described below) — implies the status quo bias discussed above. This result can be illustrated by Figure 1. In Figure 1(a), p, q and p∧q are represented by an efficiency pro argument and a status-quo-ness con argument, whereas ~p, ~q and ~p∧~q are represented by a status-quo-ness pro argument and an efficiency con argument. Suppose that in all these debates, the reformist party beats the status quo party. This is in accordance with debate splitting invariance. Moving to Figure 1(b), an efficiency pro argument is now assigned to ~p∧q (a "modest reform") as well, so that an efficiency con argument exists against both p∧q and ~p∧~q. This is the only way representation of multi-issue claims by pro/con arguments in Figure 1(b) differs from that in Figure 1(a). Nevertheless, it is shown that in the case of Figure 1(b), ~p∧~q beats p∧q. This constitutes a status quo bias. The modest reform's efficiency is as much a con argument against the status quo multi-issue claim ~p∨~q, as it is against the reformist multi-issue claim p∧q. However, the debate's resolution necessarily treats it as a decisive argument in favor of the status quo. The same effect, suitably reformulated and somewhat subtler, holds for any number of relevant features.

Note that in this example, the multi-issue claims' pro arguments play no role in the resolution of the multi-issue debate — only the con arguments are relevant. In fact, we could change the multi-issue claims' pro arguments without changing their con arguments, and the multi-issue debate's outcome would be the same. This superiority of con arguments over pro arguments in determining the multi-issue debate's outcome (the pro/con distinction being irrelevant in single-issue debates) holds for every number of relevant features. It leads to a form of "independence of irrelevant alternatives" because every pro argument for a "modest reform" (say, $p \wedge \sim q$) is a con argument against $p \wedge q$ and $\sim p \wedge \sim q$. The reason for the relevance of "modest reforms" to the multi-issue debate is their totally unproblematic relevance to the single-issue debates.[3] The ranking of con arguments above pro arguments is completely independent of the features invoked by the arguments. In other words, the model fully separates the question of "quality of the arguments" from any aspect of the "raw material these arguments manipulate".

The hierarchy of arguments and status quo effect are the paper's main results. The remainder explores two modifications of the basic model. The first recovers a class of outcome functions (and partially characterizes it), that depend purely on pro arguments, by identifying an implicit self-inconsistency of argumentation in the basic model, and ruling it out. The second introduces an additional type of con argument and derives a hierarchy of arguments for the extended model. The basic model's main results continue to hold. As can be seen throughout, I prefer to interpret the model as pertaining to debates — i.e., to collective, rather than individual choice. However, the lack of any strategic element means that the model may be relevant to individual argument-based choice. In fact, the model's predictions recall similar effects that have been observed in individual behavior and classified in the psychology literature as "reason-based choice" effects (see Shafir, Simonson and Tversky (1993) and Shafir and Redelmeier (1993)).

The paper proceeds as follows. In Section 2 I present the model of a multi-issue debate with a status quo. In Section 3 I discuss the implications of self-consistent argumentation. Section 4 concludes. The extended argumentation repertory model appears in the appendix.

2. A Model of Multi-Issue Debates with a Status Quo

The key feature of the model is disagreement over multiple issues. Let p and q denote two statements of opinion, or propositions, each on a different issue (e.g., "our tax system should include inheritance tax" and "our tax system should include capital tax — the "issues" being inheritance taxation and capital income taxation), and let ~p and ~q denote their negations.[4] The restriction to just two issues is without loss of generality. Define conjunction and disjunction of statements as usual (e.g., $p\wedge\sim q$ denotes "our tax system should include inheritance tax and should not include capital tax)". Define $P=\{p, q, \sim p, \sim q, p\wedge q, p\wedge\sim q, \sim p\wedge q, \sim p\wedge\sim q, p\vee q, p\vee\sim q, \sim p\vee q, \sim p\vee\sim q\}$. I refer to p, q, ~p and ~q as single-issue propositions, and to the other elements in P as multi-issue propositions. Part of these statements of opinion are the claims of two individual parties in a bilateral debate who entertain opposite views on both issues: p (q) and q (~q) are party A's (B's) single-issue claims and $p\wedge q$ ($\sim p\wedge\sim q$) is party A's (B's) multi-issue claim.

Let M be a finite set of desirable features (e.g., efficiency, fiscal soundness and distributive justice are desirable features that pertain to a tax policy debate). A features assignment is a function $f : P\rightarrow 2^M$ that assigns a subset of desirable features to every proposition in P. Thus, features are binary — a proposition may either "possess" it or "lack" it. Let F be some class of features assignments. A debate is a triple (x,y,f), where $(x,y)\in\{(p,q),(\sim p,\sim q),(p\wedge q,\sim p\wedge\sim q)\}$ denotes the parties' claims and $f\in F$. I refer to $(p,\sim p,f)$ and $(q,\sim q,f)$ as the single-issue debates and to $(p\wedge q,\sim p\wedge\sim q,f)$ as the multi-issue debate. This reflects the disagreement between two parties (a status quo supporter and a reformist) over multiple issues, which can be debated singly or in combination. The outcome function, defined by $d(x,y,f)\in\{x,y\}$, assigns a winner to every debate.[1]

The assumptions in this model are divided into properties of features assignments and conditions on the outcome function. Let us first consider the former. An important ingredient in the model is the assignment of features to propositions by examples. This requires some explanation. A substantial part of our argumentation in real-life debates consists of examples. Sometimes (for example, in this sentence), the use of examples is

[1] The reason for excluding the possibility of a draw will become clear shortly.

explicit. Occasionally, however, the example-like nature of an argument is somewhat disguised by phrases such as: “the tax reform will be efficient under a certain scenario”, “the prediction of this model is consistent with an empirical fact, under a suitable choice of parameter values”, or “as was shown in a certain article, this economic policy is appalling”. These are essentially examples because they can in principle be contradicted by counter-examples: different scenarios, parameter values and articles.

An important reason for the use of examples in debates is the debaters’ limited ability to aggregate information. There may be space and time limits on the amount of information that can be elicited or on the complexity of the arguments that can be raised (e.g., televised political debates). Alternatively, the debaters may have bounded ability to imagine all the pieces of information that are relevant to the debate. Let $m\in f(x)$ signify that there exists an example for an element in the class advocated by the proposition x, which has the feature m. Whenever there exists an example for a tax system with inheritance taxation which is efficient, we will say that there is an efficiency pro argument for the proposition “our tax system should include inheritance taxation”. Accordingly, f(x) is referred to as the pro argument set of x.

The example-based interpretation of f carries an important formal implication:

(P1) Union: For every $f\in F$, $x,y\in P$, $f(x\vee y)=f(x)\cup f(y)$.

The rationale behind this property should be clear: an example for an efficient tax system with either inheritance or capital taxation exists if and only if there is an example for an efficient tax system with inheritance taxation or there is an example for an efficient tax system with capital taxation. Note that the only distinction made here is between existence and non-existence of an example that invokes a given feature. It makes no difference whether or not there is more than a single example for efficient tax systems with inheritance taxation. I conjecture that results in the same spirit as this paper’s can be obtained in a model that relaxes (P1) and allows further distinction between different numbers of examples, as long as the number of examples that can be raised for each proposition and each feature is effectively bounded.[5]

The second restriction on F involves the notion of a status quo. In many real-life bilateral debates, one claim enjoys a preferred position with respect to the other, often because it is associated with the status quo. The model captures this aspect by identifying a specific feature $m^*\in M$ as "status-quo-ness" and assigning it to party B's claims but not to party A's. Thus, parties A and B are categorized as "reformist" and "status quo" camps in every debate. To sharpen the message of the second result of this section, I focus on an extreme case, in which "status-quo-ness" is the only feature assigned to $\sim p\wedge\sim q$. All other desirable features are assigned to other propositions in P.

(P2) Status-quo: For all $f\in F$, $m^*\in f(\sim p\wedge\sim q),f(\sim p),f(\sim q)$ and $m^*\notin f(x)$ for all other $x\in P$. Moreover, $m\notin f(\sim p\wedge\sim q)$ for all $m\neq m^*$.

The reader can verify that (P1) and (P2) are mutually consistent. The implication that $m^*\in f(\sim p),f(\sim q)$ and $m^*\notin f(p),f(q)$ follows from (P1). Of course, (P2) fits other notions of a preferred position than status quo (e.g., one claim is a default). To sum up, two restrictions are imposed on F: the first stems from the idea of argumentation by examples and the second from the existence of a status quo. Let us turn to the outcome function.

The outcome function d satisfies two assumptions. The first concerns the question of argument-based choice. A standard features-based choice model would represent claims in a debate by their pro argument sets, thus reducing a debate (x,y,f) into a problem of ranking two subsets of M, f(x) and f(y). Characterizing the debate's outcome would then amount to characterizing a complete relation on 2^M. The present model departs from standard approach by extending the repertory of arguments that represent claims in a debate, so as to include a certain type of con arguments. In real-life debates, we commonly use not only examples that support our own claim, but also ones that mitigate against the rival claim. Sometimes these two notions are equivalent, sometimes they are not — a con argument against x need not be a pro argument for y, and vice versa. The point is that whether they are equivalent or not, both pro and con argumentation takes place in the debate. The multiplicity of issues in the $p\wedge q$ vs. $\sim p\wedge\sim q$ debate suggests a familiar type of con argument against a claim: a pro argument for its

<u>negation</u>. The repertory of arguments that represent claims is thus enriched in the following way:

<u>(A1) Argument-based resolution</u>: Let (x,y,f) and (x',y',f') be two debates, $\alpha(z)=(f(z),f(\sim z))$ and $\alpha'(z)=(f'(z),f'(\sim z))$ for all $z\in P$, and $\sigma:\{x',y'\}\to\{x',y'\}$ be a permutation, such that $\alpha(x)=\alpha(\sigma(x'))$ and $\alpha(y)=\alpha(\sigma(y'))$. Then, $d(x,y,f)=\sigma(d(x',y',f'))$.

Assumption (A1) says that each claim in a debate is represented by its own pro argument set as well as by the pro argument set of its negation. The debate's outcome is determined solely on this basis.[6] In a single-issue debate, say (p,~p,f), the pro/con distinction is totally redundant because ~p is both the rival claim and negation of p. It is not redundant in the multi-issue debate because $\sim(\sim p\wedge\sim q)\neq p\wedge q$ and $\sim(p\wedge q)\neq\sim p\wedge\sim q$.[7] This means that the multi-issue debate's outcome depends on the features assigned to $(p\wedge\sim q)\vee(\sim p\wedge q)$ — a proposition which is not itself represented by any party. Put differently, (A1) opens the door for "dependence on irrelevant alternatives" in the multi-issue debate's resolution.

The second assumption I impose on F concerns the relation between the outcomes of single- and multi-issue debates:

<u>(A2) Debate splitting invariance</u>: For every $f\in F$: (1) If d(p,~p,f)=p and d(q,~q,f)=q, then $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$; (2) If d(p,~p,f)=~p and d(q,~q,f)=~q, then $d(p\wedge q,\sim p\wedge\sim q,f)=\sim p\wedge\sim q$.

In other words, for every $f\in F$, if the same party wins both single-issue debates, she wins the multi-issue debate as well. Real-life multi-issue debates are sometimes resolved issue-by-issue; at other times, the parties' views on the entire range of issues are confronted all at once. Debate splitting invariance fits situations in which these two procedures do not yield contradictory results: if the same party wins both single-issue debates, she is expected to win the multi-issue debate as well. For instance, it is natural for courts to resolve a complicated multi-issue case by drawing on single-issue precedents having analogous argumentation structure, such that if all single-issue precedents point to favor of the same party, the court will decide in her favor. Thus, (A2) captures an aspect of the procedure of relying on precedents in a legal system. Both the idea of debate splitting and the pro-argument/con-argument distinction are a consequence of the multiplicity of issues and are irrelevant in single-issue decision problems.

To understand better (A1) and (A2), let us consider outcome functions which violate them. First, consider the class of functions that prescribe a winner in a debate on the basis of $(f(p\wedge q),f(\sim p\wedge\sim q),f(p\wedge\sim q),f(\sim p\wedge q))$. It can be easily shown that such functions can violate (A1) in a way that none of this paper's results survive. As to (A2), let $M=\{1,\ldots,n\}$ and consider an outcome function that prescribes a winner in any debate (x,y,f) purely on the basis of f(x) and f(y) (i.e., the claims' pro argument sets). This function is lexicographic in the usual sense: $d(x,y,f)=x$ if and only if $1\in f(x)$ and $1\notin f(y)$, or if $1\in f(x),f(y)$, $1\in f(x)$ and $2\notin f(y)$, and so forth. This rule clearly satisfies (A1). However, it violates (A2). Let $n=2$, $m^*=2$ (i.e., $2\notin f(p\vee q)$ and $2\in f(\sim p\wedge\sim q)$), $1\in f(p\wedge\sim q),f(\sim p\wedge q),f(p\wedge q)$, $1\notin f(\sim p\wedge\sim q)$. By the lexicographic rule and (P1), $d(p,\sim p,f)=\sim p$ and $d(q,\sim q,f)=\sim q$, yet $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$.

Results

Our first result derives a "hierarchy of arguments" in the multi-issue debate model. In other words, we can rank the types of argument according to their effectiveness in the debate's resolution. This result also facilitates discussing our subsequent result. I usually substitute $\sim p\vee\sim q$ and $p\vee q$ for $\sim(p\wedge q)$ and $\sim(\sim p\wedge\sim q)$, respectively, for notational ease.

Proposition 1 (Hierarchy of arguments): Let $f,f'\in F$ satisfy $f(p\vee q)\neq f(\sim p\vee\sim q)$, $f(p\vee q)=f'(p\vee q)$ and $f(\sim p\vee\sim q)=f'(\sim p\vee\sim q)$. Then, $d(p\wedge q,\sim p\wedge\sim q,f)=d(p\wedge q,\sim p\wedge\sim q,f')$.

Proof – Suppose, without loss of generality, that $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$. Let $g\in F$ satisfy $g(p\wedge\sim q)=g(\sim p\wedge q)$, while $g(p\wedge q)=f(p\wedge q)$ and $g(\sim p\wedge\sim q)= f(\sim p\wedge\sim q)$. By (P1), $g(w)=f(w)$ for all $w\in\{p\wedge q,p\vee q,\sim p\wedge\sim q,\sim p\vee\sim q\}$ and therefore, by (A1), $d(p\wedge q,\sim p\wedge\sim q,g)=d(p\wedge q,\sim p\wedge\sim q,f)$. Also by (P1), $g(p)=g(q)$ and $g(\sim p)=g(\sim q)$. Therefore, by (A1), $d(p,\sim p,g)=p$ if and only if $(q,\sim q,g)=q$. By (A2), $d(p,\sim p,g)=p$ and $d(q,\sim q,g)=q$ (otherwise, we would get a contradiction with $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$). Let f^* satisfy $f^*(p\vee q)=g(p\vee q)$ and $f^*(\sim p\vee\sim q)=g(\sim p\vee\sim q)$ (thus,

$f^*(p\vee q)=f(p\vee q)$ and $f^*(\sim p\vee\sim q)=f(\sim p\vee\sim q)$. By (P1), $f^*(z)=g(z)$ for all $z\in\{p,\sim p,q,\sim q\}$. By (A2), $d(p\wedge q,\sim p\wedge\sim q,f^*)=d(p\wedge q,\sim p\wedge\sim q,f)$.

Proposition 1 states that in order to resolve a multi-issue debate with a status quo, we only have to consider the claims' con argument sets — their pro arguments are irrelevant. It should be stressed that information is strictly lost by ignoring the pro arguments: $f(p\vee q)$ and $f(\sim p\vee\sim q)$ do not pin down $f(p\wedge q)$ and $f(\sim p\wedge\sim q)$. This property holds trivially in single-issue debates, where the pro/con distinction is totally redundant.[8] Note that the result provides a hierarchy between the different argument types, independently of any trade-off that may exist different argument sets of the same type. This seems to be an essential feature of the model (as illustrated by the appendix). It yields insights about relative effectiveness of arguments, without entering at all the question of the relative importance of the features manipulated by these arguments. Given the paper's objective, this should be considered a merit.

The superiority of con arguments implies that the "dependence on irrelevant alternatives" allowed by (A1) is in fact a necessary property of the debate's resolution. The only difference between each multi-issue claim's pro argument and con argument sets lies in $f((p\wedge\sim q)\vee(\sim p\vee q))$. The reason $p\wedge\sim q$ or $\sim p\wedge q$ are relevant to the multi-issue debate is that they are relevant to the single-issue debates. There is no conceptual problem with the latter — $f(p\wedge\sim q)$ and $f(\sim p\wedge q)$ cannot be considered "irrelevant" to the single-issue debates; they are as relevant as $f(p\wedge q)$ and $f(\sim p\wedge\sim q)$ because they affect both $f(p)$, $f(\sim q)$, $f(\sim p)$ and $f(q)$. Their relevance poses a conceptual problem only in the multi-issue debate. However, by debate-splitting invariance, their single-issue relevance is carried to the multi-issue debate.

The proposition's proof reveals an interesting property implicit in the model: parties can, when debating about one issue, raise pro arguments which are at the same time con arguments against their own claims in the other single-issue debate. Thus, for instance, in the p vs. ~p debate, a pro argument for p, which is at the same time a con argument against q, is nonetheless a legitimate member of the pro argument set of p. This is the whole idea of debate splitting: we can legitimately use an argument when discussing one subject, even when the argument runs contrary to our view on other subject. We do not put on the table our attitude towards all subjects in each individual debate. The implication of this argumentation norm, however, is that it allows party B, say, to use efficiency pro arguments for p and q in the single-issue debates, although her multi-issue claim ($\sim p\wedge\sim q$) has none (see Figure 2).

Proposition 1 implies that claims in a debate are reducible to their con argument sets. I take this implication for granted in the rest of the section. For our next result, two further assumptions and one definition are necessary. Let $x\in\{p,q,p\wedge q\}$ and $y\in\{\sim p,\sim q,\sim p\wedge\sim q\}$:

(A3) Tie-breaking: Whenever $f(\sim x)-\{m^*\}=f(\sim y)$, $d(x,y,f)=y$.

A4 (Monotonicity): (1) Let $d(x,y,f)=x$, $f'(\sim x)\subseteq f(\sim x)$ and $f(\sim y)\subseteq f'(\sim y)$. Then, $d(x,y,f')=x$.
(2) Let $d(x,y,f)=y$, $f'(\sim y)$)$\subseteq f(\sim y)$ and $f(\sim x)\subseteq f'(\sim x)$. Then, $d(x,y,f')=y$.

Definition (Decisiveness and minimal decisiveness): A set of features $M^*\subset M/\{m^*\}$ is decisive if $M^*\cap f(\sim x)=\phi$ implies $d(x,y,f)=x$ and $M^*\cap f(\sim y)=\phi$ implies $d(x,y,f)=y$. It is minimally decisive if there exists no $M'\subset M^*$ which is decisive.

Assumption (A3) means that whenever two claims are indistinguishable except for the "status-quo-ness" of one of them, then this claim wins the debate. It is the minimal condition one could impose on the role of status quo considerations in a debate. Assumption (A4) means that expanding (shrinking) a claim's con argument set, or shrinking (expanding) its rival claim's con argument set, cannot reverse the debate's outcome and turn it into a winner (loser).[9] A decisive set is simply a set of features whose "unequivocal support of a claim" (i.e., a situation in which no feature in the set belongs to the claim's con argument set) guarantees winning the debate, regardless of the other features' assignment.

Proposition 2: Either $d(p\wedge q,\sim p\wedge\sim q,f)=\sim p\wedge\sim q$ for all $f\in F$, or there exists a unique nonempty minimally decisive set M^*, such that $d(p\wedge q,\sim p\wedge\sim q,f)=\sim p\wedge\sim q$ whenever $M^*\cap f(\sim p\vee\sim q)\neq\phi$.

Proof – If $d(p\wedge q,\sim p\wedge\sim q,f)=\sim p\wedge\sim q$ when $f(p\wedge q)=M-\{m^*\}$ and $(M-\{m^*\})\cap f(\sim p\vee\sim q)=\phi$, then by (A4), $d(p\wedge q,\sim p\wedge\sim q,f')=\sim p\wedge\sim q$ for all $f'\in F$. If $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$, then trivially, M-{m*} is a decisive set. Now, let M* be in this case a minimally decisive set, but assume, in contradiction, that there exist $f\in F$ and non-empty $B\subseteq M^*$ such that $M\in f(\sim p\vee\sim q)$ for all $B\in M'$ and yet $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$. Now, consider a features assignment $g\in F$ which satisfies: $g(p\wedge q)=M^*-B$, $g(p\wedge\sim q)=B$ and $g(\sim p\wedge q)=M-(M^*\cup\{m^*\})$. By (P2), $g(\sim p\wedge\sim q)=\{m^*\}$. Now, by (P1), $M^*\cap g(\sim p)=\phi$. Since M^* is decisive, this implies that $d(p,\sim p,g)=p$. Also by (P1), $g(\sim q)=B\cup\{m^*\}$ and $g(q)=M-(B\cup\{m^*\})$. By monotonicity and the contra-hypothesis, $d(q,\sim q,g)=q$. By debate-splitting invariance, $d(p\wedge q,\sim p\wedge\sim q,g)=p\wedge q$. But by monotonicity, this means that $d(p\wedge q,\sim p\wedge\sim q,g')=p\wedge q$ for every g' such that $(M^*-B)\cap g'(\sim p\vee\sim q)=\phi$. Thus, M^*-B is decisive, in contradiction with the assumption that M^* is a minimally decisive set.

It remains to show that M^* is unique. Assume the contrary, and let B^* be another minimally decisive set. If $M^*\cap B^*=\phi$, then consider the features assignment $g\in F$ which satisfies: $M^*\subset g(p\wedge\sim q)$ and $B^*\subset g(\sim p\wedge q)$. By (P1), $M^*\cap g(\sim p)=\phi$ and $B^*\cap g(p)=\phi$. As both M^* and B^* are decisive, the former implies $d(p,\sim p,g)=p$ whereas the latter implies $d(p,\sim p,g)=\sim p$, a contradiction. Now, assume that $M^*\cap B^*\neq\phi$. Consider the features assignment $f\in F$ which satisfies: $f(p\wedge q)=M^*\cap B^*$, $M^*-B^*\subset f(p\wedge\sim q)$ and $B^*-M^*\subset f(\sim p\wedge q)$. By (P1), $M^*\cap f(\sim p)=\phi$ and $B^*\cap f(\sim q)=\phi$. By decisiveness, $d(p,\sim p,f)=p$ and $d(q,\sim q,f)=q$. By debate-splitting invariance, $d(p\wedge q,\sim p\wedge\sim q,f)=p\wedge q$. By monotonicity, $M^*\cap B^*$ is a decisive set, contradicting the assumption that M^* and B^* are minimally decisive sets.

Thus, in order for the status quo to win the multi-issue debate, it suffices to have a single con argument against the reformist multi-issue claim. Now, such an argument is at the same time a con argument against the status quo because it belongs to both $f(p\vee q)$ and $f(\sim p\vee\sim q)$. Nevertheless, the outcome function necessarily interprets it as a decisive pro-status-quo argument in the multi-issue debate. We have thus derived an <u>status quo bias</u> from argument-based considerations. The bias is subtle — the feature invoked by the decisive con argument must belong to the minimal set of features, whose unequivocal support of the reformist multi-issue claim would have otherwise guaranteed its victory. If no decisive set exists, the status quo wins every debate. Note that the bias can be viewed as a sort of <u>preference reversal</u> — a supposedly "irrelevant" change in the primitives of the choice problem (i.e., $f(\sim p\wedge q)$) reverses the debate's resolution. I believe it is a merit of the model that its predictions identify two well-known biases in the psychological literature as different aspects of the same choice function.

This argument-based status quo bias is familiar from real-life multi-issue debates. Proposals for comprehensive reforms meet a heavier burden of proof than proposals for minor reforms, let alone the status quo. Of course, such an effect could be rationalized in many ways — e.g., the further a proposal for reform moves away from the status quo, the less certain we are of its potential consequences, hence the more cautious we are towards it. However, in the present model, such a bias exists even after all relevant features are taken into account. Thus, status quo bias can have purely argument-based causes, irrespective of any rationalization.

Let us return to the simple tax policy debate example of Section 1 and prove the status quo bias: $p\wedge q$ ($\sim p\wedge\sim q$) beats $\sim p\wedge\sim q$ ($p\wedge q$) in Figure 1(a) (Figure 1(a)). Consider Figure 2:

	P	~p
~q	E	S
q	E	E

~q	S,E
q	E

p	~p
E	S,E

Figure 2

In all three debates, both the status quo and reformist claims have an efficiency con arguments (by Proposition 1, we do not need to look at their pro arguments). The claims differ only in that the latter has a status quo con argument whereas the former does not. By (A3), the status quo wins in such cases. Since the con argument sets of both $p\wedge q$ and $\sim p\wedge\sim q$ are identical across Figures 1(b) and 2, it follows that $\sim p\wedge\sim q$ beats $p\wedge q$ in both cases.

This example involved only two features. As can be seen, the status quo bias exists in this case. The monotonicity assumption is required only for its extension to the case of more than two features. Suppose that the relevant features in the tax policy debate are: status-quo-ness, efficiency and fiscal soundness (abbreviated by S, E and F). Figure 3 represents two situations which differ in their features assignments:

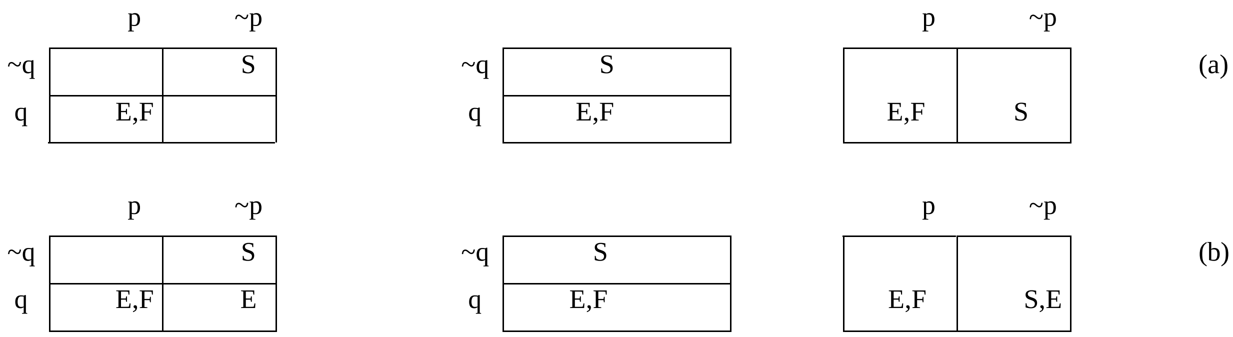

Figure 3

Figure 3(a) represents a situation in which the efficiency and fiscal soundness features "unequivocally support" the reformist multi-issue claim (i.e., $E,F \in f(p \wedge q)$ and $E,F \notin f(\sim p \vee \sim q)$). Figure 3(b), on the other hand, represents a situation in which efficiency only equivocally supports the reform, in the sense that $E \in f(p \wedge q), f(\sim p \vee \sim q)$. If $\{E,F\}$ is a minimally decisive set, then $p \wedge q$ beats $\sim p \wedge \sim q$ in the former situation but $\sim p \wedge \sim q$ beats $p \wedge q$ in the latter. Thus, a "modest reform" ($\sim p \wedge q$ in Figure 3(b) having an efficiency pro argument (but no fiscal soundness argument) suffices to guarantee the status quo's victory.

Proposition 2 states necessary conditions on the outcome function. It can be easily verified that any outcome function which satisfies $d(x,y,f)=x$ if and only if $M^* \cap f(\sim x)=\phi$ (where $M^* \subseteq M-\{m^*\}$ and x is party A's single- or multi-issue claim) is consistent with assumptions (A1)-(A4). The uniqueness of the minimally decisive set (when it exists) rules out some outcome functions, such as those which depend on the <u>number</u> of con arguments against each claim (e.g., the reformist claim wins if and only if at least one feature, other than status-quo-ness, "unequivocally supports" it). However, Proposition 2 does not characterize the minimally decisive set. It does not tell us how many features, or which features, need to "unequivocally support" the reform in order to beat the status quo. As in the case of Proposition 1, the model's capability of yielding insights on aspects of argumentation is totally independent of the more standard questions of how to rank different features sets.

3. Self-Consistent Argumentation

This section presents an argument-based consideration which, as in the case of property (P2), is formalized as a restriction on features assignments. This time, however, the restriction leads to decision rules that depend purely on the claims' pro argument sets, ignoring "irrelevant" alternatives (i.e., $f(p\wedge\sim q)$ and $f(\sim p\wedge q)$ play no role in determining the outcome of the $p\wedge q$ vs. $\sim p\wedge\sim q$ debate). To motivate the discussion, let us return to the tax policy debate example, and in particular to Figure 2. Let us try to imagine more concretely how the debate splitting procedure can actually take place. The single-issue debates are conducted independently. Each single-issue debate follows a sequential protocol: Party A (who supports p and q) moves first by providing an example for every feature she can. Party B (who supports ~p and ~q) responds with an example for every feature she can. The distinction between pro and con examples is irrelevant because we are dealing with single-issue debate. It will be useful to think of party A and B as "he" and "she", respectively.

According to the protocol, given the features assignment of Figure 2, party B is able to provide an efficiency example in both single-issue debates. But now, suppose that party A knew the specific examples party B was going to use in both debates. He could then use in the p vs. ~p debate the same efficiency example that party B uses in the q vs. ~q debate. By doing this, he creates a situation in which party B can only respond by arguing against her own argument in the other debate — she uses one example in the q vs. ~q debate and counter-argues against party A when he uses this very example in the p vs. ~p debate. Thus, given this protocol, party A can force upon party B — if he knows her argumentation "strategy" — a sort of self-inconsistent argumentation. Of course, this story should be qualified: it is just one possible debating protocol; other protocols may not give rise to such self-inconsistent argumentation. The basic idea of the story, however, is that there are single-issue debating protocols that enable one party to put the other party in a position where the latter has to choose between arguing self-inconsistently and not arguing at all.

Now, this type of self-inconsistency is different from a superficially similar element discussed in Section 2, namely the legitimacy of a pro argument for a party's claim in one single-issue debate, which is at the same time a con argument against the party's own claim in the other single-issue debate. The sort of self-inconsistency I discuss here, in contrast, applies to the debater's arguing in one single-issue debate against an argument she herself raises in the other single-issue debate. These two aspects differ in their formal implications as well. Ruling out the former annihilates the distinction between multi-issue and single-issue debates and empties the model from all content. In contrast, ruling out self-inconsistent argumentation does not trivialize the model.

Self-inconsistent argumentation may or may not be legitimate within given debating conventions. In everyday debates, it seems illegitimate to raise an example in one debate and attack it in another, related debate. In legal debates under the adversarial system, however, this kind of self-inconsistency appears legitimate. The discussion elsewhere in the paper allows self-inconsistent argumentation. In this subsection I adopt an opposite convention: a debater in one single-issue debate cannot use those arguments, against which she argues in the other single-issue debate.

In the preceding story, the source of self-inconsistent argumentation was the existence of an efficiency pro argument for both $p\wedge\sim q$ and $\sim p\wedge q$. Self-consistency means that (for some single-issue debating protocols) even if both arguments are potentially available, the argumentation norms prohibit their simultaneous use in the two single-issue debates: only one can be used. Thus, we will simply rule out features assignments for which pro arguments for $p\wedge\sim q$ and $\sim p\wedge q$ that invoke the same feature exist simultaneously.

P3 (Self-consistent argumentation): For every $f\in F$, $f(p\wedge\sim q)\cap f(\sim p\wedge q)=\phi$.

This is the only restriction on features assignments in this subsection: (P2) is dropped and thus, I do not deal with debates with a status quo. The question I address is whether under this new restriction, there exist "standard" outcome functions that depend only on the claims' pro arguments (and thus do not rely on "irrelevant alternatives").

This question can be reduced to an order-extension problem. An outcome function d(x,y,f) which depends only on f(x) and f(y) induces a relation $\succ$ on 2^M, such that for every debate (x,y,f), d(x,y,f)=x if and only if $f(x)\succ f(y)$. Assume further that $\succ$ is a linear ordering on 2^M, hence an extension of a primitive ordering > on M={1,…,n} (say, $n>n-1>\cdots>1$, without loss of generality), which is interpreted as a ranking of all relevant features by their relative importance. Finally, assume that $B \succ \phi$ for every $B \subseteq M$. This simply means that every combination of features is desirable. Given the model's reduction, debate splitting invariance can be translated into the following assumption:

A5 ("Translated" debate-splitting invariance): For every $A,B,C,D \subset M$ such that $C \cap D = \phi$ and $A \neq B$, if $A \cup C \succ B \cup D$ and $A \cup D \succ B \cup C$, then $A \succ B$.

Spiegler (1999) studies order extensions which satisfy (A5) in a very different context[10]; shows their existence; and proves the following result. For every $A \subseteq M$, denote $a \in A$ by max(A) if $a > b$ for all $b \in A$, $b \neq a$.

Proposition 3 (Spiegler, 1999): For every $A,B \subseteq M$, $\max(A) > \max(B)$ implies $A \succ B$.

Thus, under self-consistent argumentation, it is possible to obtain a standard resolution of multi-issue debates, which is purely a function of pro arguments and independent of "irrelevant alternatives". However, there are restrictions: under mild auxiliary assumptions (order extension and desirability), the outcome function must be reduced to a max-max rule. Proposition 5 does not say anything about the debate's resolution when max(A)=max(B). The main point of the result is to show that the non-standard effects of the basic model — supremacy of con arguments, dependence on "irrelevant" alternatives, preference reversal and the status quo bias — can be traced to an implicit argumentation norm which allows a kind of self-inconsistent argumentation.

4. Discussion

This paper has two main contributions. First, it presents a formal approach to argument-based choice. Our starting point was the intuition that the notion of "quality of an argument" is an influential, and largely distinct notion, from this or that aspect of the raw pieces of information or preferences manipulated by the argument. In other words, being persuasive in a debate depends on finding the "right" argument, on top of using the "right" information. Our results were consistent with this intuition: insights on the relative effectiveness of various pro and con arguments have been derived, independently of the specific features they invoke. The paper's second contribution is to show that status quo bias may have a purely argument-based origin. The bias emerges even after all the considerations that could potentially rationalize it are taken into account. Specifically, the sources of the bias are multiplicity of issues in a debate (and the considerations it raises: pro/con argumentation and debate splitting) and argumentation by examples. In this section, I discuss the interpretation of the model.

4.1. Is the "modest reform" an "irrelevant" alternative?

As emphasized in Section 2, both the hierarchy-of-arguments and status quo bias results exhibit "dependence on irrelevant alternatives". I would like to elaborate on this important aspect. It will be simplest to use the tax policy debate example for this purpose. The "modest reform" referred to by Politician 2 is "irrelevant" to the multi-issue debate in the sense of not being a feasible alternative. It is, however, "relevant" in the sense of serving well Politician 2's argument. The model developed in this paper provides an explanation for this discrepancy: the modest reform clearly is relevant to the single-issue debates, and debate splitting invariance carries this relevance to the multi-issue debate. But if the modest reform is relevant as an argument, what prevents it from being a feasible alternative in the debate?

One explanation can be given in terms of the model itself. Consider the situation represented by Figure 2(b): the "modest reform" ($\sim p \wedge q$) has an efficiency pro argument but lacks a fiscal soundness pro argument. Suppose $\sim p \wedge q$ were suggested as an alternative to the status quo, $\sim p \wedge \sim q$. The modest reform and the status quo claims would then agree on the p vs. ~p issue. Therefore, the appropriate way to present the conflict between the modest reform and the status quo is as a single-issue debate, in which each alternative is the other's negation:

q	S
~q	E

Figure 4

In this debate, $\alpha(q)=(\{S\},\{E\})$ and $\alpha(\sim q)=(\{E\},\{S\})$. Assuming that q beats ~q under this arguments profile is fully consistent with the outcomes of the debates represented by Figure 3. If the modest reform also had a fiscal soundness pro argument, it would win the debate represented by Figure 4 as well (see Figure 3(a)). Thus, the modest reform can serve a decisive pro-status-quo argument in the status quo vs. drastic reform debate, but in a single-issue debate against the status quo, it would not have enough arguments on its own behalf and would lose the debate. The crucial part of this argument is the view of the conflict between the modest reform and the status quo as a single-issue debate. To conclude, the possibility that an alternative is too weak to confront the status quo, yet sufficiently relevant to affect a debate in which it is not a feasible alternative, is entirely consistent with the model.

Other explanations are available on the basis of framing considerations. For instance, in a decision problem as complex as a tax policy debate, alternatives need to be represented and backed up by experts, lobbies, authorities, etc. — otherwise, the public will not consider them as proper alternatives. The public does not expect the same back up from tax policies that are mentioned only for the sake of the argument. This does not necessarily affect the seriousness with which it takes an argument that invokes them. Thus, it is possible for an opinion to be considered sufficiently relevant as part of an argument, yet insufficiently relevant as a feasible alternative.

4.2. Individual reason-based choice

As the behavioral decision making literature (e.g., Shafir, Tversky and Simonson (1993)) points out, several experimentally observed violations of standard rationality can be attributed to reason-based choice procedures. It is not at all clear how to conceptualize the idea of reason-based choice so as to render it applicable to formal economic analysis. Although I discussed the formalism in this paper in the context of collective argument-based choice (i.e., debates), I believe it is also potentially applicable to individual decision making.

In an experiment reported by Shafir and Redelmeier (1995), a population of physicians faced two hypothetical choice problems that required them to assign surgery priorities to patients on the basis of their "types" (i.e., diagnoses). The first experimental treatment involved choosing between one type-A patient and one type-B patient. The second experimental treatment involved choosing from among two type-A patients and one B-type patient. The physicians' propensity to give priority to type-B patients was significantly higher in the second treatment. I believe that this effect can be attributed to a similar decision procedure to the one studied in this paper. The physicians in the second treatment may have framed the decision problem not as a complete ranking problem, but as a "tournament", in which alternatives are compared pairwise. When a type-A patient is compared to a type-B patient, the former patient's diagnosis is a good pro argument for preferring her, but the existence of a second type-A patient having the same diagnosis is a con argument that weakens this preference. This increases the propensity to choose the type-B patient as the "winner of the tournament".

A related, more casual observation is the way pedestrians sometimes decide whether to spare change to a beggar. Their propensity to give anything to any beggar at all on their way to work is lower when there are two beggars than when there is just one. As well as in the physicians experiment, the search for good arguments is a key element in the choice procedure. The arguments for the alternatives are affected by an "irrelevant" third alternative. The reasoning behind this "dependence on irrelevant alternatives" therefore appears related to that studied in this paper. Further research is necessary to examine further this link.

4.3 Are there reducible elements in the model?

The model operates on two different levels: propositions and arguments. Assumption (A2) is formulated purely in terms of the former, whereas (A1) assumes that every debate can be reduced to the latter. Properties (P1) and (P2) mediate between the two levels. It is therefore tempting to dispense with the level of propositions altogether and formulate the entire model, including (A2), in terms of vectors of argument sets, such that the outcome function is reduced to a relation on these vectors. I would like to explain why such a priori reduction is not a good idea. If we tried to express (A2) in terms of vectors of argument sets, we would encounter some difficulties. First, the translation of (A2) into the languages of pro and con argument sets is not identical. Therefore, the rationale behind the assumption becomes unclear and its workability is in doubt. Second, as we saw in Section 3, it is possible to "turn off" the hierarchy-of-arguments result of the basic model by identifying an argument-based consideration that is implicit in it (self-inconsistent argumentation). This is impossible to do in a reduced model. Finally, as shown in the appendix, the unreduced model allows us to imagine additional types of arguments and derive new hierarchy-of-arguments results. Again, it is hard to see how this can be obtained in a reduced-from model. To conclude, working with the two levels may not be standard in features-based choice models, but I believe its benefits in terms of understanding argument-based choice outweigh the extra level's "cost".

4.4. Strategic argumentation

Strategic elements are absent from the present model, although implicit in the modification presented in Section 3. The model may nonetheless serve as a springboard for analyzing strategic considerations in the debaters' choice of claims and arguments. First, the difference between argumentation in single-issue and multi-issue debates may enter the debaters' considerations if they choose their opinions strategically. Second, we saw that certain con arguments can reverse the outcome of debates. However, if the rules of the debating game restrict the circumstances under which such con arguments can be made — e.g., allowing them only as counter-arguments to pro arguments — the decision whether to raise a pro argument becomes strategic. I hope to address these questions in future research.

References

1. J. Glazer and A. Rubinstein (1999). Debates and Decisions: On a Rationale of Argumentation Rules. Mimeo, Tel-Aviv University.
2. P. Grice (1989). Studies in the Ways of Words. Cambridge: Harvard University Press.
3. D. Redelmeier and E. Shafir (1995). Medical decision making in situations that offer multiple alternatives. *Journal of the American Medical Association,* **273**, 302-305.
4. E. Shafir, I. Simonson and A. Tversky (1993). "Reason-Based Choice." Cognition **49**, 11-36.
5. R. Spiegler (1999). Inferring a Ranking from Imperfect Observations as an Order Extension Problem. Mimeo, Tel-Aviv University.

Appendix: A model with an additional con argument

The basic model contains a single argument-based distinction, between a proposition's pro argument set represented by f(x) and its con argument set represented by f(~x). In this sub-section I introduce another type of con argument, which is independent of f. Let $h:P\rightarrow 2^M$ be a function which, like f, attaches a subset of features to every proposition in P. And like f, h satisfies the "union" property: for every $x,y\in P$, $h(x\vee y)=h(x)\cup h(y)$. The interpretation of h is the opposite of f: when $a\in h(x)$, this means that there exists an example for an element in the class advocated by x, which does not have a certain feature. E.g., in the context of the tax policy debate, such an example may be a possible scenario under which inheritance taxation is inefficient — in this case we will say that there exists an efficiency internal con argument against the claim "our tax system should include inheritance taxation". For every $x\in P$, h(x) will be referred to as the internal con argument set of x (and f(~x) will be referred to as the external argument set of x).

For every $x\in P$, f(x) and h(x) are completely independent. Thus, for instance, lack of efficiency pro examples for inheritance taxation does not imply the existence of an efficiency internal con example against it. This would not be possible in a world with logical omniscience — if there is no example for efficient tax policies with inheritance taxation, then there must be an example for inefficient ones. In contrast, in a world with bounded ability to think of examples of either sort, it could be that $a\notin f(x)$ and $a\notin h(x)$ at the same time. Thus, there is no a priori difference in the definitions of pro arguments and internal con arguments. One would therefore not expect them to be ranked by their effectiveness in the debate's resolution. The surprising result of this appendix is that they can be (almost completely) ranked. The reason is that the external con arguments against multi-issue claims do provide us with some information on the pro arguments of single-issue claims, and at the same time imply nothing regarding internal con arguments. In other words, the fact that external con arguments are derived from f not from h is the cause of this result.

The primitives of the extended model are thus enriched: a debate is a quadruple (x,y,f,h), where $(x,y)\in\{(p,q),(\sim p,\sim q),(p\wedge q,\sim p\wedge\sim q)\}$. The definition of "status-quo-ness" and argument-based resolution are modified accordingly. Debate-splitting invariance, being phrased only in terms in propositions and not of arguments, remains unchanged. Let H be a class of internal con argument functions.

P2’: For all $f\in F$, $h\in H$, $m^*\in f(\sim p\wedge\sim q), f(\sim p), f(\sim q)$, $m^*\notin f(x)$ for all other $x\in P$ and $m^*\notin h(\sim p\wedge\sim q)$. Moreover, $m\notin f(\sim p\wedge\sim q)$ for all $m\neq m^*$.

A1’: Let (x,y,f,h) and (x’,y’,f’,h’) be two debates, let $\alpha(z)=(f(z),f(\sim z),h(z))$ and $\alpha'(z)=(f'(z),f'(\sim z),(h))$ for all $z\in P$, and let $\sigma:\{x',y'\}\rightarrow\{x',y'\}$ be a permutation, such that $\alpha(x)=\alpha(\sigma(x'))$ and $\alpha(y)=\alpha(\sigma(y'))$. Then, $d(x,y,f,h)= \sigma(d(x',y',f',h'))$.

The results of Section 2, suitably rephrased ((x,y,f) should be everywhere replaced with (x,y,f,h)) continue to hold in the extended model: in debates with a status quo, only external con arguments determine debates’ outcomes. When assumptions (A3) and (A4) are imposed, the same status quo bias holds. The proofs are practically identical and are therefore omitted. Thus, both the superiority of external con arguments and the status quo bias survive this extension of the argumentation repertory.

For the remainder of the appendix, the status-quo-ness property is dropped. No claim in either the single-issue or the multi-issue debates has any privileged position. No restriction on f and h except for the union property (which is motivated by the example-like nature of argumentation in this model) is imposed. It turns out that in this larger domain of debates, a complete hierarchy of arguments can be obtained, whenever the assignment of pro arguments to the multi-issue claims is sufficiently “rich”. First, the supremacy of external con arguments continues to hold in this extended model. That is, for any debate (x,y,f,h) such that $f(x)\neq f(y)$, the debate’s outcome is invariant with respect to changes in f and h that leave $f(\sim x)$ and $f(\sim y)$ intact. The proof is virtually the same as in Proposition 1 and is therefore omitted. In addition, we can derive a ranking between pro arguments and internal con arguments for a large class of features assignments for which $f(\sim x)=f(\sim y)$.

Proposition 4: Let (f,h) and (f',h') satisfy f(~x)=f'(~x)=f(~p∨~q)=f'(~y), f(x)≠f(y), f(x)=f'(y), and f(y)=f'(y). Also, let f(x) and f(y) contain at least one element each and let (f(x)∪f(y))-(f(x)∩f(y)) contain at least three elements. Then, d(x,y,f,h)=d(x,y,f',h').

Proof – Let (f,h) satisfy f(p∨q)=f(~p∨~q)=M. Denote M*(f)=(f(p∧q)∪f(~p∧~q))-(f(p∧q)∧f(~p∧~q)). By assumption, M*(f) can be partitioned into three subsets, {A,B,C}, such that A∪B=f(p∧q) and C=f(~p∧~q), without loss of generality. Now consider all pairs (f'',h'') which satisfy h''(p∧~q)=h''(~p∧q), h''(p∧q)=h(p∧q), h''(~p∧~q)=h(~p∧~q), f''(p∧q)=f(p∧q), f''(~p∧~q)=f(~p∧~q), f''(p∨q)=f(p∨q) and f''(~p∨~q)=f(~p∨~q). By (P1) (applied to both f'' and h''), f''(w)=f(w) and h''(w)=h(w) for all w∈{p∧q,p∨q,~p∧~q,~p∨~q}. Therefore, by (A1), d(p∧q,~p∧~q,f'',h'')=d(p∧q,~p∧~q,f,h). The following table is constructed in order to show that there must be at least one such pair (f'',h''), for which the same party wins both single-issue debates, (p,~p,f'',h'') and (q,~q,f'',h''). The table presents four combinations of f''(z)∩M*(f) for all z∈{p,~p,q,~q), all of which are consistent with the conditions on (f'',h'').

	f''_1	f''_2	f''_3	f''_4
f''(p) , f''(~p)	M*(f'') , B∪C	M*(f'') , B∪C	M*(f'') , B∪C	M*(f'') , A∪C
f''(q) , f''(~q)	A∪B , A∪C	A∪B , M*(f'')	M*(f'') , A∪C	A∪B , M*(f'')

Figure 5

By construction of h'', h''(p)=h''(q) and h''(~p)=h''(~q). As can be seen from Table 4, (A1) implies that the same party must win both single-issue debates for at least one (f'',h'') that meets the conditions above. Suppose, without loss of generality, that d(p,~p,f'',h'')=~p and d(q,~q,f'',h'')=~q. By (A2), d(p∧q,~p∧~q,f'',h'') =~p∧~q. The remainder of the proof is the same as in Proposition 1. Let (f*,h*) satisfy f*(p∨q)=f''(p∨q), f*(~p∨~q)=f''(~p∨~q), h*(p∨q)=h''(p∨q) and h*(~p∨~q)=h''(~p∨~q) (thus, f*(p∨q)=f(p∨q), f*(~p∨~q)=f(~p∨~q), h*(p∨q)=h(p∨q) and h*(~p∨~q)=h(~p∨~q)). By (P1), f*(z)=f(z) and h*(z)=h(z) for all z∈{p,~p,q,~q}. By (A2), d(p∧q,~p∧~q,f*,h*)= d(p∧q,~p∧~q,f,h). €

This result states an almost complete hierarchy of arguments. External con arguments are the most important in determining the debate's outcome. If claims in a debate have the same external con argument sets, the outcome is decided by the pro arguments, as long as there are not too few pro arguments for each claim. To put this result more figuratively, suppose that the debate's protocol is such that party A raises a pro argument and party B can respond only with one argument that invokes the same feature party A's argument did. Then, party B would prefer an external con argument (if she makes a pro argument for her claim that will be an external con argument as well, by definition) to an internal one.

---

[1] The only precedent I am aware of for an economic-theoretic study of argumentation is Glazer and Rubinstein (1997), in which an "argumentative" element of choice (the asymmetry between examples and counter-examples in debates) emerges as a property of a constrained optimal information elicitation mechanism. The difference between the present paper and Glazer and Rubinstein (1997) is that the latter develops the argument-based aspect in a strategic framework inspired by Gricean Pragmatics (Grice (1989)), whereas the present paper is decision-theoretic.

[2] Let us leave aside for the moment the question why it is not a feasible alternative in the first place. This question is discussed at length in Section 4.

[3] The superiority of con arguments also holds in debates without a status quo. In this class of debates, in contrast to debates with a status quo, it is possible for two claims to have identical con argument. Only in this case does the debate's resolution depend on the claim's pro arguments.

[4] Of course, reasonable application of the model would have the two issues related, even if logically independent: e.g., two aspects of a multi-dimensional debate about tax policy. I am not interested in situations in which p is a statement about inheritance taxation and q is a statement about scientific theories.

[5] One aspect of argumentation by examples is completely ignored in this paper. Consider two situations: in the first, a single example exists for a tax system with inheritance taxation which is both efficient and fiscally sound; in the second, two examples for a tax system with inheritance taxation exist, one efficient but fiscally unsound and the other fiscally sound but inefficient. These situations are considered equivalent in this paper, but clearly they do not have to be. Thus, the model fits debates in which all features are mutually exclusive (e.g., no tax system can possess both efficiency and distributive fairness at the same time), or situations in which time or space limits prevent debaters from mentioning more than one desirable feature per example, etc.. Of course, this entire problem can be solved by redefining the features set as 2^M, such that the redefined features are mutually exclusive.

[6] Other notions of a con argument against a claim other than a pro argument for the claim's negation could be imagined, of course (see the appendix).

[7] Note that (P2) guarantees $\alpha(x)\neq\alpha(y)$ for every (x,y,f). Thus, the requirement that every debate has a winner is tantamount to ruling out draws whenever claims have different α values.

[8] Property (P2) plays no direct role in the proof. Without (P2), however, we would not be able to assume that every debate has a winner (see footnote 7). In that case, Proposition 1 would have to be slightly modified: the multi-issue claims' pro arguments would be considered by the outcome function if and only if $f(p\vee q)=f(\sim p\vee\sim q)$.

[9] Normally, we tend to view monotonicity as a normative requirement. However, I believe that in a model whose business is to derive a decision bias, normative arguments are somewhat misleading. Rather, the interpretation of Proposition 2 should be that status quo bias can be avoided only at the price of admitting the possibility that adding con arguments against a claim (or removing con arguments against its rival claim) will actually improve its chances of winning. That is, status quo bias and violation of monotonicity are complementary phenomena in debates, and this paper simply focuses on the former. Whether monotonicity violation has interesting argument-based content is a question that I leave for future research.

[10] In Spiegler (1999), (A5) formalizes a non-Bayesian inference procedure.