Consistent Judgement Aggregation:
The Truth-Functional Case[1]

Klaus Nehring
Department of Economics, University of California at Davis
Davis, CA 95616, U.S.A.
kdnehring@ucdavis.edu

and

Clemens Puppe
Department of Economics, University of Karlsruhe
D – 76128 Karlsruhe, Germany
puppe@wior.uni-karlsruhe.de

October 2006

[1] Acknowledgements to be added

Abstract. *Generalizing the celebrated "discursive dilemma," we analyze judgement aggregation problems in which a group of agents independently votes on a set of complex propositions (the "conclusions") and on a set of "premises" by which the conclusions are truth-functionally determined. We show that for conclusion- and premise-based aggregation rules to be mutually consistent, the aggregation must always be "oligarchic," that is: unanimous within a subset of agents, and typically even be dictatorial. We characterize exactly when consistent non-dictatorial (or anonymous) aggregation rules exist, allowing for arbitrary conclusions and arbitrary interdependencies among premises.*

1 Introduction

A new kind of aggregation problem has recently emerged from the Law and Economics literature, the problem of aggregating individual judgements on sets of logically interrelated propositions. The leading example is the following so-called "doctrinal paradox" (Kornhauser and Sager (1986), or "discursive dilemma" (Pettit (2001)). A court of three judges has to decide on an agenda consisting of the following three propositions: "the defendant broke the contract" (proposition a), "the contract was legally valid" (proposition b) and "the defendant is liable" (by legal doctrine equivalent to the conjunction of a and b, i.e. to the proposition $a \wedge b$). Suppose that the first judge affirms a but not b and therefore, for reasons of logical consistency, also not $a \wedge b$. The second judge affirms b but not a and thus neither $a \wedge b$. Finally, the third judge affirms all three propositions. Can we aggregate these individual judgements into a logically consistent "social" judgement on the case? The doctrinal paradox consists in the observation that majority voting on the "premises" a and b, as well as on the "conclusion" $a \wedge b$ leads to an inconsistent set of judgments, since both premises are affirmed by a respective majority of two judges, but at the same time a majority of two judges rejects the conclusion.

The purpose of the present paper is to demonstrate the robustness of the discursive dilemma for general "truth-functional" agendas under an appropriate independence condition.[2] An agenda is referred to as *truth-functional* if it contains a set of complex propositions (the "conclusions") whose truth-values are determined by the truth-values of a set of atomic propositions (the "premises"). While the bulk of the first-generation literature on judgement aggregation inspired by List and Pettit's (2002) seminal contribution has dealt with the truth-functional case, contrary to appearance, these contributions fail to properly demonstrate the robustness of the discursive dilemma. In particular, to date there does not exist any reasonably general result for *simple* agendas, that is: for agendas with a single conclusion. This case is clearly fundamental as it corresponds to a single binary group decision, as exemplified by the original discursive dilemma. By contrast, the strongest extant results (see the literature review below) apply to agendas with a sufficiently rich set of conclusions, and are thus mathematically and conceptually much more restrictive.

In contrast to the extant literature, the present paper shows that in many cases strong impossibilities (dictatorships) apply already for simple agendas. While non-dictatorial consistent aggregation is sometimes possible, the possibilities remain tightly confined to oligarchies; but, at least, in these cases anonymous consistent aggregation satisfying independence is possible in the form of unanimity rules. This is demonstrated in the basic result of the paper, Theorem 1, which also shows exactly when such non-dictatorial aggregation is possible, namely in the case of "conjunctive" agendas. A simple truth-functional agenda is called "conjunctive" if either the conclusion or its negation can be written as the conjunction of the premises (or their negations); for example, the original doctrinal paradox given by the agenda $\{a, b, a \wedge b\}$ is conjunctive, as are the agendas $\{a, b, a \vee b\}$ and $\{a, b, a \rightarrow b\}$. By contrast, the agenda

[2]We require that the social judgement on each proposition only depends on the individual judgements on *that* proposition, and that this dependence is monotonic ("monotonic independence"). It has been noted in the literature that monotonic independence is equivalent to strategy-proofness if individuals' preferences over judgements are "single-peaked," see Nehring and Puppe (2002/2006a), Dietrich and List (2004).

$\{a, b, a \leftrightarrow b\}$ is non-conjunctive and thus forces dictatorships. In the case of non-simple agendas, the formal definition of conjunctiveness requires the conclusions fit together in an appropriate, tightly circumscribed way. As a result, only very little "richness" is required to force dictatorship.

In the second part of the paper, we go beyond the formulations of the existing literature by allowing for logical interdependencies among the premises. These interdependencies may arise from purely logical interrelations, or from an underlying "background consensus" in that all agents may agree that certain combinations of premises are ruled out. While not yet studied in the literature, this case seems to be quite important for applications. Consider, for instance, the multi-criteria evaluation of a candidate in a committee, say a hiring decision in an academic department where candidates are evaluated according to their research potential on the one hand and their teaching skills on the other. Here, the "premises" are the evaluations according to the different criteria, such as "good research," "excellent research," or "fair teaching abilities." The relevant "conclusion" is the hiring decision which is determined truth-functionally by the hiring criteria. Premises are interdependent here since, for instance, asserting a particular grade on one criterion implies not asserting any other grade on the same criterion.

As another example, consider the following version of the discursive dilemma. Suppose that the defendant actually signed two contracts and that there are doubts whether he committed a breach of contract in each case. For $i = 1, 2$, let a_i stand for the proposition "the defendant broke contract i." In analogy to the original example, suppose that the judges have to assess the truth of both a_1 and a_2, and whether the two contracts were legally valid (propositions b_1 and b_2). Moreover, the judges agree that the defendant is to be held liable if and only if he acted illegally in at least one case, i.e. if and only if $(a_1 \wedge b_1) \vee (a_2 \wedge b_2)$ is true. However, now suppose that due to the structure of the contracts it is commonly agreed upon that legal validity of the first contract implies legal validity of the second contract, so that the implication $b_1 \rightarrow b_2$ is commonly held to be true. Evidently, the premises are no longer independent in this case as well.

Generalizing Theorem 1, we show that in the case of interdependent premises consistent yet non-dictatorial aggregation continues to be possible only in the form of oligarchic aggregation rules, and only for conjunctive agendas. Somewhat surprisingly, it turns out that non-dictatorial aggregation is possible only when the premises are effectively logically independent, see Theorem 2 below.

The remainder of the paper is organized as follows. In the following Section 2, we characterize the class of all aggregation rules in the truth-functional case with independent premises. In Section 3, we generalize this to the case of interdependent premises. After stating our main result, we discuss, in Subsection 3.2, an application to an enhanced version of the classical preference aggregation problem. Specifically, we interpret the premises as binary preference judgements between social alternatives, and show how different rationality requirements (e.g. transitivity, asymmetry, completeness) can be described as specific contexts, i.e. as specific interdependencies among premises. Moving beyond the classical formulation, individuals' judgements on the optimality of alternatives in a given feasible set are also aggregated independently. These represent the "conclusions" that are truth-functionally determined by the premises. This additional aggregation requirement leads to a novel impossibility theorem for the aggregation of tournaments (complete and asymmetric relations). Section 4 provides further discussion and concludes.

2 Truth-Functional Aggregation with Logically Independent Premises

2.1 Basic Notation and Definitions

By $L = \{p, q, r, ...\}$ we denote the set of all *propositions*, and by $L^0 = \{a, b, c, ...\}$ the subset of all *atomic* propositions. Any proposition is built up from a finite set of atomic propositions via the standard connectives of propositional logic. Typical elements of L are thus $p = a \wedge b$, or $p = a \wedge \neg b$, and so on. We use negation "$\neg$" and conjunction "$\wedge$" as the primitive logical operators and define the other logical connectives "$\vee$," "$\rightarrow$," and "$\leftrightarrow$" in the standard way. Throughout, we use standard two-valued propositional logic. We identify $\neg\neg p$ with p, and we write $p = q$ if p and q are logically equivalent, i.e. if $p \leftrightarrow q$ is a tautology. For all subsets $K \subseteq L$, denote by K^* the *negation closure* of K, i.e. the set of all propositions p such that $p \in K$ or $\neg p \in K$ (or both). The set $(L^0)^*$ of all atomic propositions and their negations is referred to as the set of *literals*. A proposition p is *truth-functionally determined* by the set $\{a_1, ..., a_m\}$ of atomic propositions if, for any $j \in \{1, ..., m\}$ and any selection $l_j \in \{a_j, \neg a_j\}$,

$$\{l_1, ..., l_m, p\} \text{ is consistent} \Leftrightarrow \{l_1, ..., l_m, \neg p\} \text{ is inconsistent.}$$

The *agenda*, i.e. the set of propositions to be decided upon, is denoted by Z. Throughout we assume that the agenda neither contains the tautology nor the contradiction. A *truth-functional* agenda consists of a set of atomic propositions $\{a_1, ..., a_m\}$ (the "premises") and a set of complex propositions $\{p_1, ..., p_k\}$ (the "conclusions") each of which is truth-functionally determined by the a_j; for obvious reasons, we assume $p_h \neq a_j$ for all h, j. A truth-functional agenda with only one conclusion will be called *simple*. Moreover, say that a truth-functional agenda $\{a_1, ..., a_m; p_1, ..., p_k\}$ is *conjunctive* if one can choose $l_j \in \{a_j, \neg a_j\}$ for all $j \in \{1, ..., m\}$, and $z_h \in \{p_h, \neg p_h\}$ for all $h \in \{1, ..., k\}$, such that

$$z_h = \bigwedge_{j \in I_h} l_j$$

for all $h \in \{1, ..., k\}$ and appropriate $I_h \subseteq \{1, ..., m\}$. In other words, a truth-functional agenda is conjunctive if each conclusion or its negation can be written as the conjunction of a subset of premises or negations thereof; crucially, it must be possible to choose the "sign" of the premises independently of the conclusion in which they might occur.

A truth-functional agenda $\{a_1, ..., a_m; p_1, ..., p_k\}$ is *irreducible* if there does not exist a strict subset $A \subset \{a_1, ..., a_m\}$ and a (not necessarily strict) subset $P \subseteq \{p_1, ..., p_k\}$ such that all $p \in P$ are truth-functionally determined by the atomic propositions in A and all $p \in \{p_1, ..., p_k\} \setminus P$ are truth-functionally determined by $\{a_1, ..., a_m\} \setminus A$.

A *judgement (relative to the agenda Z)* is a set $J \subseteq L$ of propositions such that (i) $J^* = Z^*$, and (ii) J is logically consistent (in the sense of propositional logic). Thus, a judgement is a consistent selection of the propositions in the agenda that is maximal in the sense that, for all $p \in Z$, either $p \in J$ or $\neg p \in J$ (but, of course, not both). Finally, denote by $\mathcal{J}$ the set of all judgements (relative to a fixed agenda Z).

Denote by $N = \{1, ..., n\}$ the set of individuals. An *aggregation rule of judgements* is a mapping

$$F : \begin{array}{ccc} \mathcal{J}^n & \rightarrow & \mathcal{J} \\ (J_1, ..., J_n) & \mapsto & J \end{array},$$

where J_i is the judgement of individual i. Note that since an aggregation rule is defined on $\mathcal{J}^n$ there is an implicit unrestricted domain condition. The following additional properties of an aggregation rule will play an important role in our analysis.

Sovereignty (S) The mapping F is onto, i.e. any $J \in \mathcal{J}$ is in the range of F.

Condition (S) is a very weak requirement; for instance, it is satisfied whenever the aggregation respects unanimous consent, i.e. whenever, for all $J \in \mathcal{J}$, $F(J, ..., J) = J$. The following independence condition is central.

Monotone Independence (MI) Consider $(J_1, ..., J_n)$ and $(J'_1, ..., J'_n)$ such that, for some p and all i, $p \in J_i \Rightarrow p \in J'_i$. Then, $p \in F(J_1, ..., J_n) \Rightarrow p \in F(J'_1, ..., J'_n)$.

Monotone Independence states that if the aggregate judgement entails p, and if the individual support for p increases, then p must remain in the aggregate judgement. Condition (MI) can be thought of as being composed of an Arrovian independence condition ("the aggregate judgement on a proposition may only depend on the individual judgements on *that* proposition") and a condition of non-negative responsiveness.

An aggregation rule is called *oligarchic with default* J_0 if there exists $M \subseteq N$ such that $F(J_1, ..., J_n) = J_0$ whenever $J_i = J_0$ for some $i \in M$. Under (MI) there is exactly one oligarchic rule with default J_0 and oligarchy M. It is given as follows. If $p \in J_0$, then $p \in F(J_1, ..., J_n)$ if and only if $p \in J_i$ for some $i \in M$; if $p \notin J_0$, then $p \in F(J_1, ..., J_n)$ if and only if $p \in J_i$ for all $i \in M$. Thus, each proposition in the default judgement is collectively affirmed as soon as at least one oligarch affirms it; by contrast, the negation of a proposition in the default judgement requires unanimous consent among the members of the oligarchy. Special cases of oligarchic rules are the unanimity rule (with $M = N$) and dictatorship (with $M = \{i\}$ for some i).

2.2 Characterization of all Non-Dictatorial Agendas

Say that an agenda Z is *dictatorial* if the only aggregation rules on Z satisfying (S) and (MI) are the dictatorial ones. Moreover, say that Z is *oligarchic* if all aggregation rules on Z satisfying (S) and (MI) are oligarchic and if some of them are non-dictatorial.

Theorem 1 *Let $Z = \{a_1, ..., a_m; p_1, ..., p_k\}$ be an irreducible truth-functional agenda. The following statements are equivalent.*
i) *Z is non-dictatorial.*
ii) *Z is oligarchic.*
iii) *Z is conjunctive.*
Moreover, if Z is conjunctive, all non-dictatorial oligarchic rules involve the same default judgement. If $z_h = \wedge_{j \in I_h} l_j$, for all $h = 1, ..., k$, the default judgement is given by $\{\neg l_1, ..., \neg l_m; \neg z_1, ..., \neg z_k\}$.

Theorem 1 follows from the more general Theorem 2 below. To illustrate the result in the simplest case with two premises, consider the agenda $\{a, b; p\}$. As is easily verified, such an agenda is conjunctive if $p = a \wedge b$, $p = a \vee b$, or $p = a \rightarrow b$, but not if $p = a \leftrightarrow b$. Thus, by Theorem 1, in the first three cases but not in the fourth case there exist non-dictatorial aggregation rules. In the case $p = a \wedge b$ (the original doctrinal paradox) all oligarchic aggregation rules involve the default judgement $\{\neg a, \neg b; \neg(a \wedge b)\}$; similarly, the default judgements corresponding to the cases $p = a \vee b$ and $p = a \rightarrow b$ are given by $\{a, b; a \vee b\}$ and $\{\neg a, b; a \rightarrow b)\}$, respectively.

More generally, a simple and irreducible truth-functional agenda $\{a_1, ..., a_m; p\}$ is conjunctive if and only if there is either only one truth-value assignment of the a_j that makes p true, or only one truth-value assignment of the a_j that makes p false. For instance, the agenda $\{a, b, c; p\}$ where $p = a \wedge (b \vee c)$ is not conjunctive.

When there are several complex propositions conjunctiveness not only requires that each of these can be represented as a conjunction of its premises appropriately "signed," but also that the sign of a premise is the same in all conclusions in which that premise occurs. To illustrate, consider the agenda $\{a, b, c; p, q\}$, where $p = b \to c$. First, assume that the second conclusion is given by $q = a \to b$. As is easily verified, the agenda is not conjunctive in this case; hence, by Theorem 1, only dictatorial rules emerge. Intuitively, the reason is that the default candidates corresponding to the two conclusions are not compatible with each other. Indeed, the default judgement corresponding to $a \to b$ entails the assertion of b, while the default judgement corresponding to $b \to c$ entails the negation of b. By contrast, if $q = a \to c$ the above agenda is conjunctive, hence non-dictatorial aggregation becomes possible. The consistent oligarchic rules all involve the default judgement $\{\neg a, \neg b, c; p, q\}$.

Finally, note that, up to isomorphism, there exists a unique minimal conjunctive agenda on m premises (obviously $\{a_1, ..., a_m; \wedge_{j=1}^m a_j\}$), and a unique maximal conjunctive agenda which adds all subconjunctions of the form $\wedge_J a_j$ where $J \subseteq \{1, ..., m\}$.

2.3 Related Literature

The first formal impossibility result in the context of judgement aggregation is due to List and Pettit (2002).[3] Their basic result has been strengthened in Pauly and van Hees (2006, Theorem 4) by giving up anonymity. These two contributions are in fact the only ones in the extant literature deriving impossibility results for agendas with a single complex proposition as in the original discursive dilemma. However, both results use a very strong independence condition, called "systematicity." In conjunction with closedness under negation, this condition is in fact equivalent to an Arrovian independence of irrelevant alternatives condition plus neutrality. In particular, on the domain $\{p, \neg p\}$, majority voting is the *only* positively responsive rule satisfying systematicity and anonymity. Moreover, there is no anonymous rule satisfying systematicity with an even number of individuals.[4]

The impossibility results of Dietrich (2006a) and the other results in Pauly and van Hees (2006) use independence instead of systematicity but rely on strong richness assumptions on the agenda. For instance, the richness condition imposed by Pauly and van Hees (2006) implies that an agenda contains *both* the conjunction and the disjunction of any two premises. In other words, a possibility result would require that the same aggregation rule for the premises yields a conclusion that is *simultaneously* consistent with the direct aggregation of individual judgements regarding both the

[3] see also List and Pettit (2004), and List (2003, 2004) for extensions.

[4] As noted above, our condition of monotone independence is equivalent to the Arrovian independence of irrelevant alternatives condition plus non-negative responsiveness. Some of the impossibility results obtained in the literature do not assume non-negative responsiveness (see, e.g., List and Pettit (2002) and Dietrich (2004)). However, conceptually, the step from independence to monotone independence seems very small; indeed, we are not aware of any interesting aggregation method that would satisfy independence but not non-negative responsiveness. For a characterization of the agendas (truth-functional or not) that admit non-dictatorial aggregation rules satisfying independence but not necessarily non-negative responsiveness, see Dokow and Holzman (2005).

conjunction and the disjunction. This is a strong, systematicity-like requirement. And it matters: as noted above, both the conjunction and the disjunction admit of consistent oligarchic aggregation; however, these will be mutually incompatible unless they are dictatorial.

Pauly and van Hees (2006, Theorem 6) also observe that conjunctive agendas admit consistent unanimity rules and force a veto; however, they do not establish the full oligarchic characterization and, more importantly, do not establish the necessity of conjunctiveness for non-dictatorial aggregation.

Throughout, our framework assumes standard two-valued propositional logic. Moreover, we require both individual and collective judgements to be *complete* in the sense that each proposition in the agenda has either to be affirmed or rejected. Gärdenfors (2006) and, subsequently, Dietrich and List (2006) study judgement aggregation without the completeness requirement. Pauly and van Hees (2006) and van Hees (2004) analyze judgement aggregation in multi-valued logics. Dietrich (2004) shows that certain results extend to general "monotone" logics, including modal logics.

Dietrich (2006b) and Nehring and Puppe (2006b) derive possibility results on agendas with a premises/conclusion-like structure but without truth-functionality. Finally, Mongin (2005) and Nehring (2006) investigate the consequences of weakening the independence condition.

3 Interdependent Premises

In many cases, the premises of an aggregation problem will not be logically independent as assumed so far. A simple example already mentioned above is the evaluation of candidates in a hiring decision, say at an academic department. Concretely, suppose that each candidate can receive 1 to m points, and consider the set $\{a_1, ..., a_m\}$, where each a_j is interpreted as "the candidate receives at least j points." Evidently, the "premises" a_j are no longer independent since, for all $j \geq k$, the implication $a_j \to a_k$ is always true under this interpretation.

The interdependence of premises can formally described by a "context" as follows. A *context* is a consistent set $C \subseteq L$ of propositions. The intended interpretation is that the propositions contained in C are agreed upon by everyone and thus constrain both individual and social judgements. An agenda Z is compatible with the context C if, for each $q \in Z$, both q and $\neg q$ are consistent with C. By $\mathcal{D} \subseteq \mathcal{J}$ we denote the set of all judgements on Z that are consistent with C, i.e. $J \in \mathcal{D}$ if and only if $J \cup C$ is consistent. The interdependence of premises thus leads to a domain restriction to some subset of judgements $\mathcal{D} \subseteq \mathcal{J}$. Domain restrictions have, in general, a double-edged effect on the possibility of consistent aggregation. By making the domain of individual judgements smaller, consistent aggregation becomes easier. On the other hand, by making the co-domain of social judgements smaller as well, consistency becomes harder to achieve. It turns out that, as a consequence of the maintained truth-functional structure of judgement aggregation, the first effect is inoperative while the second is not. Interdependence of premises thus further tilts the scale towards impossibility.[5]

[5] In this way, our results differ from those of List (2004) who shows that domain restrictions may lead to possibility results, in analogy to the case of Arrovian preference aggregation. In contrast to the present work, List gives up the assumption of truth-functionality that characterizes the original discursive dilemma and is maintained here.

3.1 The Main Result

To formulate the main result, we first need to generalize the definitions in the previous section to the case of interdependent premises. A proposition p is *truth-functionally determined* by $\{a_1, ..., a_m\}$ *given* C, if for any $j \in \{1, ..., m\}$ and any selection $l_j \in \{a_j, \neg a_j\}$,

$$\{l_1, ..., l_m, p\} \cup C \text{ is consistent} \Leftrightarrow \{l_1, ..., l_m, \neg p\} \cup C \text{ is inconsistent.}$$

In the following, we write $p =_C q$ if p and q are equivalent given C, i.e. if $\bigwedge C \wedge p \leftrightarrow \bigwedge C \wedge q$ is a tautology. A truth-functional agenda $\{a_1, ..., a_m; p_1, ..., p_k\}$ is *independently conjunctive* if if one can choose $l_j \in \{a_j, \neg a_j\}$ for all $j \in \{1, ..., m\}$, and $z_h \in \{p_h, \neg p_h\}$ for all $h \in \{1, ..., k\}$, such that (i)

$$z_h =_C \bigwedge_{j \in I_h} l_j$$

for all $h \in \{1, ..., k\}$ and appropriate $I_h \subseteq \{1, ..., m\}$, (ii) the set $\{l_j\}_{j \in I_h}$ is logically independent given the context C, and (iii) $\neg l_1 \wedge ... \wedge \neg l_m \cup C$ is consistent. Thus, the definition of *independent* conjunctiveness adds to the previous definition the logical independence of the premises (clause (ii)) and the consistency of the default with the context (clause (iii)).

A truth-functional agenda $\{a_1, ..., a_m; p_1, ..., p_k\}$ is *irreducible* if there does not exist a strict subset $A \subset \{a_1, ..., a_m\}$ and a (not necessarily strict) subset $P \subseteq \{p_1, ..., p_k\}$ such that all $p \in P$ are truth-functionally determined given C by the atomic propositions in A and all $p \in \{p_1, ..., p_k\} \setminus P$ are truth-functionally determined given C by $\{a_1, ..., a_m\} \setminus A$.

The following result generalizes Theorem 1 to the case of possibly interdependent premises.

Theorem 2 *Let $Z = \{a_1, ..., a_m; p_1, ..., p_k\}$ be an irreducible truth-functional agenda (given a context C). The following statements are equivalent.*
i) *Z is non-dictatorial.*
ii) *Z is oligarchic.*
iii) *Z is independently conjunctive.*
Moreover, if Z is independently conjunctive, all non-dictatorial oligarchic rules involve the same default judgement. If $z_h = \wedge_{j \in I_h} l_j$, for all $h = 1, ..., k$, the default judgement is given by $\{\neg l_1, ..., \neg l_m; \neg z_1, ..., \neg z_k\}$.

The starting point for the proof of Theorem 2 is the characterization of all aggregation rules satisfying (MI) and (S) for arbitrary agendas (truth-functional or not) provided in Nehring and Puppe (2006a). There, it is shown that the class of all such aggregation rules can be described as voting by propositions satisfying a simple combinatorial condition called the "Intersection Property" (see the appendix for details).

3.2 Examples and Applications

3.2.1 Performance-Based Hiring Decisions

To illustrate the content of Theorem 2, consider again the example of a hiring decision at an academic department. Assume now that the decision depends on two distinct criteria, say performance in research and in teaching. Concretely, let b_j stand for "the candidate receives at least j points for research," let c_j stand for "the candidate receives at least j points for teaching" and consider the set $\{b_1, ..., b_m, c_1, ..., c_m\}$. Evidently, the context is given by all implications of the form $b_j \to b_k$ and $c_j \to c_k$, for $j \geq k$. Now add to this set of "premises" (with $m = 10$, say) a hiring decision p that is a function of the candidates' points. First, consider the hiring decision according to which the candidate is employed if she receives at least 7 points in each criterion, i.e. $p = b_7 \wedge c_7$. Note that the only irreducible agenda is $\{b_7, c_7; p\}$ in this case.[6] Evidently, this agenda is conjunctive; moreover, the premises are independent given the context, and the negation of all premises is consistent with the context, hence the agenda is even independently conjunctive. In accordance with Theorem 2, any oligarchic rule with default $\{\neg b_7, \neg c_7; \neg p\}$ is consistent.

Next consider the hiring decision p_{sum}, where p_{sum} stands for "the sum of the grades for research and teaching is greater than, or equal to 10." In this case, the agenda $\{b_1, ..., b_{10}, c_1, ..., c_{10}, p_{sum}\}$ is the only irreducible one. It is easily seen that this agenda is not conjunctive since neither p_{sum} nor $\neg p_{sum}$ can be written as a conjunction of the premises or their negations. According to Theorem 2, only dictatorial rules emerge now. For instance, suppose that some individuals give maximal points $(4, 6)$ for research and teaching, respectively, while all others give the points $(6, 4)$. The unanimity rule with default $(1, 1)$ results in $(4, 4)$ despite the fact that p_{sum} is unanimously accepted.

Finally, suppose that, for some unfathomable reason, the academic department in question wants to hire a good but not too good candidate, say one with points ≥ 5 and ≤ 8. For simplicity, assume that only research matters, i.e. assume that the hiring decision is given by $\tilde{p} = b_5 \wedge \neg b_9$. In this case, the unique irreducible agenda is given by $\{b_5, b_9, \tilde{p}\}$. Evidently, this agenda is conjunctive but not *independently* conjunctive since b_5 and b_9 are not independent. Hence, due to the failure of clause (ii) in the definition of independent conjunctiveness only dictatorial rules emerge in this example. In the next subsection, we provide examples showing how the failure of clause (iii) can lead to dictatorship.

3.2.2 Aggregation of Preference-plus-Optimality Judgements

Another important application of Theorem 2 is the *preference aggregation* problem. Concretely, interpret the premises $a_{(x,y)}$ as binary preference judgements of the form "state x is preferred to state y" for any pair x and y of distinct states in some universe X of social alternatives. Moreover, for any x in some non-empty subset $Y \subseteq X$, let the conclusion $p_{x;Y}$ stand for "state x is a best alternative in the feasible set Y." Thus, let the agenda Z be given by

$$Z = \{a_{(x,y)} : x, y \in X, x \neq y\} \cup \{p_{x;Y} : x \in Y, \emptyset \neq Y \subseteq X\}$$

Clearly, the conclusions are truth-functionally determined by the premises. Moreover, the resulting agenda is conjunctive since optimality of an alternative in a feasible set is logically equivalent to a conjunction of binary preference judgements between that alternative and all other feasible alternatives. Indeed, for all x, and all non-empty Y,

[6] A more comprehensive vote on the candidate's performance (also on the other premises) completes the picture but is not needed for the purpose of arriving at at a reason-based hiring decision.

one has

$$p_{x;Y} = \bigwedge_{y \in Y \setminus \{x\}} a_{(x,y)}.$$

Note that so far, we have made no assumptions about possible interrelations between the premises, i.e. between binary preference judgements. In other words, no rationality restrictions have yet been imposed. These can be naturally described using different contexts. For instance, the following three contexts describe asymmetry, transitivity and completeness of the underlying preference relation, respectively:

$$\begin{aligned} C_{asy} &:= \{a_{(x,y)} \to \neg a_{(y,x)} : x, y \in X, x \neq y\}, \\ C_{trans} &:= \{(a_{(x,y)} \wedge a_{(y,z)}) \to a_{(x,z)} : x, y, z \in X, x, y, z \text{ pairwise distinct}\}, \\ C_{comp} &:= \{a_{(x,y)} \vee a_{(y,x)} : x, y \in X, x \neq y\}. \end{aligned}$$

Clearly, these contexts can also be combined with each other. For instance, the context C_{trans} describes the case where preferences are preorders, the context $C_{trans} \cup C_{comp}$ the case of weak orders, and the context $C_{asy} \cup C_{comp}$ the case of tournament relations. As is easily verified, the three agendas $Z \cup C_{asy}$, $Z \cup C_{trans}$ and $Z \cup C_{asy} \cup C_{trans}$ are independently conjunctive. Indeed, clause (ii) in the definition of independent conjunctiveness is always satisfied here. Moreover, the default judgement is always given by the negation of all preference judgements and all optimality conclusions. Thus, clause (iii) in the definition of independent conjunctiveness (consistency of the default) is satisfied if and only if the context does not entail completeness. By Theorem 2, one obtains oligarchic possibility results for the contexts C_{asy}, C_{trans} (preorders) and $C_{trans} \cup C_{asy}$ (strict partial orders). Evidently, in the cases of preorders and strict partial orders, these results are closely related to Gibbard's well-known oligarchy theorem. Formally, however, the results are different since our present analysis pertains to the simultaneous aggregation of binary preference *and* optimality judgements.

For the contexts C_{comp}, $C_{asy} \cup C_{comp}$, $C_{trans} \cup C_{comp}$ and $C_{asy} \cup C_{trans} \cup C_{comp}$ one obtains strong impossibility results, i.e. dictatorships by Theorem 2 due to the failure of clause (iii) in the definition of independent conjunctiveness. The entailed impossibility in case of the latter two, i.e. in the case of weak orders and linear orders, respectively, could alternatively be derived from versions of Arrow's Theorem.[7] Probably the most interesting *novel* insight derived from applying Theorem 2 is the impossibility result corresponding to the context $C_{asy} \cup C_{comp}$, i.e. the impossibility of aggregating tournament relations (*together* with entailed optimality judgements) into a "social" tournament relation (*together* with entailed optimality judgements) under Monotone Independence.

Comparing the present enhanced formulation of the preference aggregation problem with the classical one, one can view the latter as having implicitly taken the side in favor of premise-based aggregation, treating the actual optimality judgement as a mere fallout of the preference aggregation.

3.3 Discussion

As a response to the strengthened impossibility flavor of Theorem 2 as compared to Theorem 1, one might argue that the independent aggregation of interdependent

[7] Again, the present results are weaker due to the simultaneous aggregation of optimality judgements. For derivations of (versions of) Arrow's theorem in the context of judgement aggregation, see Nehring (2003) and Dietrich and List (2005).

premises may not be plausible. Whether this is the case will depend on the particular situation and on the nature of the interdependencies. In particular, Monotone Independence is likely to remain appealing if the interdependency among premises is consistent with non-degenerate aggregation rules among the premises themselves. For instance, in the example of a hiring decision based on a multi-criteria evaluation, the independence condition naturally arises from an *ordinal* interpretation of points or grades, such as a classification of candidates according to "excellent," "good," and "poor" teaching abilities. Here, independence can be satisfied consistently, for example by premise-wise majority voting which amounts to median voting on each criterion.

By contrast, Monotone Independence is less plausible if the criteria are scaled *cardinally.* In this case, one could base the hiring decision on, say, the mean score of a candidate (cf. Rubinstein and Fishburn (1986)). Clearly such an aggregation method violates independence among the individual grade judgements. However, the appeal of cardinal aggregation rules in the present context is immediately reduced by the observation that cardinal aggregation even on a *single* criterion will lead to instances of the discursive dilemma. For instance, suppose that all committee members agree that a candidate should be hired if and only if her uni-dimensional score is above some cut-off value. Then, it is easily possible that a candidate's mean score is above the cut-off level while at the same time a majority thinks that the candidate should not be hired.

Finally, the relevance of strategy-proofness considerations as an indirect motivation for Monotone Independence is not *per se* affected by the possible logical interdependence of premises.

4 Conclusion

In this paper, we have provided a general analysis of the possibility of aggregating judgements independently yet consistently in situations in which one or more outcome propositions depend truth-functionally on its premises. Our main advance over the existing literature was to demonstrate the existence of robust impossibilities already in the fundamental case of simple agendas with a single outcome proposition. We also showed that whether these take oligarchic or dictatorial form depends on the structure of the truth-functional dependence of the outcome on the premises, and that interdependence among the premises never opens new possibilities but further tilts the balance towards dictatorship.

Appendix A: Proofs

Evidently, Theorem 1 is a special case of Theorem 2. In order to prove the latter, we first characterize the aggregation rules on $\mathcal{D}$ satisfying conditions (S) and (MI). It follows from the analysis in Nehring and Puppe (2002, 2006a) that their common structure can be described as voting by propositions ("voting by issues") satisfying a simple combinatorial condition, the "Intersection Property," as follows.

A family of *winning coalitions* is a non-empty family $\mathcal{W}$ of subsets of the set N of all individuals satisfying $[W \in \mathcal{W} \text{ and } W' \supseteq W] \Rightarrow W' \in \mathcal{W}$. A *structure of winning coalitions* on the negation closure Z^* of an agenda Z is a mapping $p \mapsto \mathcal{W}_p$ that assigns a family of winning coalitions to each proposition $p \in Z^*$ satisfying the following condition,

$$W \in \mathcal{W}_p \Leftrightarrow (N \setminus W) \notin \mathcal{W}_{\neg p}. \qquad \text{(A.1)}$$

In words, a coalition is winning for p if and only if its complement is not winning for the negation of p. A mapping $F : \mathcal{D}^n \to \mathcal{D}$ is called *voting by issues*, or, in our context simply *propositionswise voting*, if for some structure of winning coalitions and all $p \in Z^*$,

$$p \in F(J_1, ..., J_n) \Leftrightarrow \{i : p \in J_i\} \in \mathcal{W}_p.$$

Observe that, so far, nothing guarantees that the outcome judgement $F(J_1, ..., J_n)$ propositionwise voting is consistent with the context C. The necessary and sufficient condition for consistency can be described as follows.

A *critical family* is a minimal subset $Q \subseteq Z^*$ of propositions such that $Q \cup C$ is inconsistent. The sets $\{p, \neg p\}$ are called *trivial* critical families. As an example, the non-trivial critical families of the negation closure of $\{a, b, a \wedge b\}$ with an empty context are $\{\neg a, a \wedge b\}$, $\{\neg b, a \wedge b\}$ and $\{a, b, \neg(a \wedge b)\}$. Note that, e.g., $\{\neg a, \neg b, a \wedge b\}$ is also an inconsistent set of propositions but it is not minimal. A structure of winning coalitions satisfies the *Intersection Property* if for any critical family $\{p_1, ..., p_l\} \subseteq Z^*$, and any selection $W_j \in \mathcal{W}_{p_j}$,

$$\bigcap_{j=1}^{l} W_j \neq \emptyset.$$

The following result is an immediate consequence of Nehring and Puppe (2006a, Theorem 3), resp. Theorem 2 in Nehring and Puppe (2002).

Theorem 3 *An aggregation rule $F : \mathcal{D}^n \to \mathcal{D}$ satisfies (S) and (MI) if and only if it is propositionwise voting satisfying the Intersection Property.*

Using (A.1) and the fact that families of winning coalitions are closed under taking supersets, we obtain

$$\mathcal{W}_{\neg p} = \{W \subseteq N : W \cap W' \neq \emptyset \text{ for all } W' \in \mathcal{W}_p\}. \qquad \text{(A.2)}$$

The following *conditional entailment relation* plays a central role. We write $p \geq^0 q$ if there exist $q_1, ..., q_k$ such that i) $\{p, q_1, ..., q_k\} \cup C$ is consistent, ii) $\{\neg q, q_1, ..., q_k\} \cup C$ is consistent, and iii) $\{p, \neg q, q_1, ..., q_k\} \cup C$ is inconsistent. Observe that

$$p \geq^0 q \Leftrightarrow [p \neq \neg q \text{ and there exists a critical family containing } p \text{ and } \neg q]. \qquad \text{(A.3)}$$

By $\geq$ we denote the transitive closure of $\geq^0$, and by $\equiv$ the symmetric part of $\geq$. Note that $\geq$ is "negation adapted" in the sense that $p \geq q \Leftrightarrow \neg q \geq \neg p$.

Lemma A.1 *Suppose that a structure of winning coalitions satisfies the Intersection Property. Then,* $p \geq q \Rightarrow \mathcal{W}_p \subseteq \mathcal{W}_q$.

Proof of Lemma A.1 Evidently, it suffices to show that $p \geq^0 q \Rightarrow \mathcal{W}_p \subseteq \mathcal{W}_q$. Thus, let $\{p, \neg q, ...\}$ be a critical family. Since the set N of all individuals is always winning, the Intersection Property applied to this critical family implies that any $W \in \mathcal{W}_p$ intersects any $W' \in \mathcal{W}_{\neg q}$. By (A.1), we thus have $W \in \mathcal{W}_q$, hence $\mathcal{W}_p \subseteq \mathcal{W}_q$.

The following lemma is a restatement of Lemma 1 in Nehring and Puppe (2005).

Lemma A.2 *Suppose that a structure of winning coalitions satisfies the Intersection Property, and assume that* p, q, r *are contained in some critical family. If* $\mathcal{W}_{\neg p} \subseteq \mathcal{W}_q$, *then* $\{i\} \in \mathcal{W}_{\neg r}$, *for some* $i \in N$.

For the proof of Theorem 2, we need the following three further auxiliary results.

Lemma A.3 *Suppose that* p *is truth-functionally determined given* C *by* $a_1, ..., a_m$ *and that the set* $\{a_1, ..., a_m\}$ *is minimal with that property. Then, for all* $j = 1, ..., m$, $p \equiv^0 a_j$ *or* $\neg p \equiv^0 a_j$.

Proof of Lemma A.3 Consider any fixed $j \in \{1, ..., m\}$, and denote by l_{-j} a generic selection of a_l or $\neg a_l$ for each $l \neq j$. By minimality of $\{a_1, ..., a_m\}$, there exists l_{-j} such that $\{a_j, l_{-j}, p\} \cup C$ and $\{\neg a_j, l_{-j}, \neg p\} \cup C$ are consistent, or $\{a_j, l_{-j}, \neg p\} \cup C$ and $\{\neg a_j, l_{-j}, p\} \cup C$ are consistent. Indeed, if for all l_{-j} for which $\{a_j, l_{-j}, p\} \cup C$ is consistent, $\{\neg a_j, l_{-j}, \neg p\} \cup C$ would be inconsistent, we would obtain that also $\{\neg a_j, l_{-j}, p\} \cup C$ is consistent; but then a_j would be a truth-functionally redundant premise.

Thus, assume without loss of generality that $\{a_j, l_{-j}, p\} \cup C$ and $\{\neg a_j, l_{-j}, \neg p\} \cup C$ are consistent, and hence that $\{a_j, l_{-j}, \neg p\} \cup C$ and $\{\neg a_j, l_{-j}, p\} \cup C$ are inconsistent. Since $\{\neg a_j, l_{-j}, p\} \cup C$ is inconsistent, $\{\neg a_j, l_{-j}, p\}$ contains a critical family. But since both $\{a_j, l_{-j}, p\} \cup C$ and $\{\neg a_j, l_{-j}, \neg p\} \cup C$ are consistent, this critical family must contain both p and $\neg a_j$, which implies $p \geq^0 a_j$. By a completely symmetric argument, the inconsistency of $\{a_j, l_{-j}, \neg p\} \cup C$ implies $a_j \geq^0 p$.

Lemma A.4 *Let* $\{a_1, ..., a_m; p_1, ..., p_k\}$ *be an irreducible truth-functional agenda. The sign of the propositions and their respective negations can be chosen such that* $p_h \equiv a_j$ *for all* $j \in \{1, ..., m\}$ *and all* $h \in \{1, ..., k\}$.

Proof of Lemma A.4 The proof proceeds by induction. For all p_h, let $A_h := \{a_1^h, ..., a_{m_h}^h\}$ denote a minimal set of premises by which p_h is truth-functionally determined given C, and let $B_h := A_1 \cup ... \cup A_h$. By Lemma A.3, we can label the premises in A_1 such that $p_1 \equiv a$ for all $a \in A_1$. Now consider p_l, and assume that we have labeled all conclusions $p_1, ..., p_{l-1}$ and all premises in B_{l-1} as desired. By irreducibility, we have $A_l \cap B_{l-1} \neq \emptyset$; there are now two cases. First, suppose that there exists $a \in A_l \cap B_{l-1}$ such that $p_l \equiv a$. Then in fact, $p_l \equiv a$ for all $a \in B_{l-1}$ by transitivity of the relation $\equiv$, and by Lemma A.3, one can also label all premises $a \in A_l \setminus B_{l-1}$ such that $p_l \equiv a$. Next suppose that, for all $a \in A_l \cap B_{l-1}$, one has $\neg p_l \equiv a$. Then replace p_l by $\neg p_l$, and, again by Lemma A.3, label all premises a in $A_l \cap B_{l-1}$ such that $\neg p_l \equiv a$. In both cases, we thus obtain $p_l \equiv a$ for all $a \in B_l$ by transitivity.

Lemma A.5 *Let $\{a_1, ..., a_m; p_1, ..., p_k\}$ be an irreducible truth-functional agenda with $p_h \equiv a_j$ for all h, j. For each $h \in \{1, ..., k\}$, either there exists a critical family containing p_h and at least two elements of $\{\neg a_1^h, ..., \neg a_{m_h}^h\}$, or there exists a critical family containing $\neg p_h$ and at least two elements of $\{a_1^h, ..., a_{m_h}^h\}$.*

Proof of Lemma A.5 By assumption, for each $j \in \{1, ..., m_h\}$, there exists a critical family containing p_h and $\neg a_j^h$. Suppose that all these have cardinality two, i.e. are given by the collection $\{p_h, \neg a_j^h\}_{j=1,...,m_h}$. Then, we obtain $p_h =_C \bigwedge_{j=1,...,m_h} a_j^h$, which implies that $\{\neg p_h, a_1^h, ..., a_{m_h}^h\}$ is a critical family.

Proof of Theorem 2 Let $Z = \{a_1, ..., a_m, p_1, ..., p_k\}$ be an irreducible truth-functional agenda. By Lemma A.4, we may assume without loss of generality that the labeling of premises and conclusions is such that $p_h \equiv a_j$ for all h, j. First, we show that all aggregation rules on Z satisfying (S) and (MI) must be oligarchic. By Theorem 3, all such rules can be described as voting by propositions. By Lemmas A.1 and A.4, we obtain $\mathcal{W}_{p_h} = \mathcal{W}_{a_j}$ for all h, j, hence by (A.2) also $\mathcal{W}_{\neg p_h} = \mathcal{W}_{\neg a_j}$ for all h, j. By Lemma A.5, there exists a critical family containing p_1 and $\neg a_j, \neg a_l$ for some $j \neq l$, or there exists a critical family containing $\neg p_1$ and a_j, a_l for some $j \neq l$; Without loss of generality, assume the latter. Applying Lemma A.2 to $a_j, \neg p_1, a_l$, we obtain $\{i\} \in \mathcal{W}_{\neg a_l}$ for some i. Since the winning coalitions for all $\neg p_h$ and all $\neg a_j$ are identical, we thus obtain $\{i\} \in \mathcal{W}_{\neg p_h} = \mathcal{W}_{\neg a_j}$ for all h, j. Let M be the set of all voters i such that $\{i\} \in \mathcal{W}_{\neg p_h} = \mathcal{W}_{\neg a_j}$. Using (A.1) and (A.2), it is easily verified that the structure of winning coalitions is given by

$$\mathcal{W}_{\neg p_h} = \mathcal{W}_{\neg a_j} = \{W \subseteq N : W \ni i \text{ for some } i \in M\} \text{ and} \tag{A.4}$$
$$\mathcal{W}_{p_h} = \mathcal{W}_{a_j} = \{W \subseteq N : W \supseteq M\}, \tag{A.5}$$

for all h, j. Evidently, this structure of winning coalitions describes an oligarchic rule with oligarchy M and default $J_0 := \{\neg a_1, ..., \neg a_m, \neg p_1, ..., \neg p_k\}$; indeed, any member of M is alone winning for each $\neg a_j$ and for each $\neg p_h$, and collectively they are winning for each a_j and for each p_h.

We now show that Z admits non-dictatorial rules, i.e. oligarchic rules with $\#M \geq 2$, if and only if Z is independently conjunctive. Using the Intersection Property, it is easily seen that an oligarchic rule with $\#M \geq 2$ and default J_0 is consistent if and only if no critical family meets J_0 more than once. Indeed, if i and j are members of the oligarchy, $\{i\}$ and $\{j\}$ are disjoint coalitions either of which is winning for each proposition in J_0 by (A.4). Thus, by the Intersection Property, consistency of the oligarchic rules requires that no critical family meets J_0 more than once. Conversely, under that condition the Intersection Property is always satisfied by (A.4) and (A.5).

Suppose now that all critical families meet J_0 at most once. Clearly, in this case the default is consistent with C, and the premises are logically independent given C. Moreover, by Lemma A.4 all non-trivial critical families containing p_h contain exactly one element of $\{\neg a_1^h, ..., \neg a_{m_h}^h\}$, which implies $p_h =_C \bigwedge_{j=1,...,m_h} a_j^h$. Thus, the agenda Z is independently conjunctive. Conversely, suppose that Z is independently conjunctive, and let $p_h = \wedge_{j \in I_h} a_j$ for all $h = 1, ..., k$. Clearly in this case, for no h and no j there exists a critical family containing $\neg p_h$ and $\neg a_j$; moreover, no critical family can contain more than one element of $\{\neg p_1, ..., \neg p_k\}$. Finally, since for all h, the premises of p_h are logically independent given C no critical family meets $\{\neg a_1^h, ..., \neg a_m^h\}$ more than once.

References

[1] Dietrich, F. (2004), "A Generalized Model of Judgement Aggregation," forthcoming in *Social Choice and Welfare*.

[2] Dietrich, F. (2006a), "Judgement Aggregation: (Im)Possibility Theorems," *Journal of Economic Theory* 126, 286-298.

[3] Dietrich, F. (2006b), "The Possibility of Judgement Aggregation under Subjunctive Implications," *mimeographed*.

[4] Dietrich, F. and List, C. (2004), "Strategy-Proof Judgement Aggregation," *mimeographed*.

[5] Dietrich, F. and List, C. (2005), "Arrow's Theorem in Judgement Aggregation," forthcoming in *Social Choice and Welfare*.

[6] Dietrich, F. and List, C. (2006), "Judgement Aggregation without Full Rationality," *mimeographed*.

[7] Dokow, E. and Holzman, R. (2005), "Aggregation of Binary Evaluations," *mimeographed*.

[8] Gärdenfors, P. (2006), "An Arrow-like Theorem for Voting with Logical Consequences," *Economics and Philosophy* 22, 181-190.

[9] Kornhauser, L. and Sager, L. (1986), "Unpacking the Court," *Yale Law Journal* 96, 82-117.

[10] List, C. (2003), "A Possibility Theorem on Aggregation over Multiple Interconnected Propositions," *Mathematical Social Sciences* 45, 1-13.

[11] List, C. (2004), "A Model of Path Dependence in Decisions over Multiple Propositions," *American Political Science Review* 98, 495-513.

[12] List, C. and Pettit, P. (2002), "Aggregating Sets of Judgements: An Impossibility Result," *Economics and Philosophy* 18, 89-110.

[13] List, C. and Pettit, P. (2004), "Aggregating Sets of Judgements: Two Impossibility Results Compared," *Synthese* 140, 207-235.

[14] Mongin, P. (2005), "Factoring out the Impossibility of Logical Aggregation," *mimeographed*.

[15] Nehring, K. (2003), "Arrow's Theorem as a Corollary," *Economics Letters* 80, 379-382.

[16] Nehring, K. (2006), "The Impossibility of a Paretian Rational," *mimeographed*.

[17] Nehring, K. and Puppe, C. (2002), "Strategy-Proof Social Choice on Single-Peaked Domains: Possibility, Impossibility and the Space Between," *mimeographed*.

[18] Nehring, K. and Puppe, C. (2005), "Strategy-Proof Social Choice without Dictators," *mimeographed*.

[19] Nehring, K. and Puppe, C. (2006a), "The Structure of Strategy-Proof Social Choice: General Characterization and Possibility Results on Median Spaces," forthcoming in *Journal of Economic Theory*.

[20] Nehring, K. and Puppe, C. (2006b), "Justifiable Group Choice," *mimeographed*.

[21] Pauly, M. and van Hees, M. (2006), "Logical Constraints on Judgement Aggregation," *Journal of Philosophical Logic* 35, 569-585.

[22] Pettit, P. (2001), "Deliberative Democracy and the Discursive Dilemma," *Philosophical Issues* 11, 268-299.

[23] Rubinstein, A. and Fishburn, P. (1986), "Algebraic Aggregation Theory," *Journal of Economic Theory* 38, 63-77.

[24] van Hees, M. (2004), "The Limits of Epistemic Democracy," forthcoming in *Social Choice and Welfare*.