Non-manipulable partitioning*

Conal Duddy[†] and Ashley Piggins[‡]

August 23, 2010

Abstract

Consider the following social choice problem. A group of individuals seek to partition a finite set X into two subsets. The individuals may disagree over the partition and so an aggregation rule is applied to determine a compromise outcome. We require that the group partition should not be imposed, nor should it be manipulable. We prove that the only aggregation rules satisfying these properties are dictatorships.

JEL classification: D71.

Keywords: Partitions, aggregation, non-manipulation.

1 Introduction

The potential for social choice correspondences to be manipulated is by now well understood in the literature.[1] These correspondences select a nonempty

*Financial support from the Irish Research Council for the Humanities and Social Sciences, the Spanish Ministry of Science and Innovation through Feder grant SEJ2007-67580-C02-02 and the NUI Galway Millennium Fund is gratefully acknowledged. We thank Thierry Marchant for helpful references.

†Corresponding author. Government of Ireland Scholar, J.E. Cairnes School of Business and Economics, National University of Ireland Galway, University Road, Galway, Ireland. Email: conal.duddy@gmail.com

‡J.E. Cairnes School of Business and Economics, National University of Ireland Galway, University Road, Galway, Ireland. Email: ashley.piggins@nuigalway.ie

[1] Barberà (2010) is the best survey to date.

subset of the set of social alternatives at each profile of individual preferences. In this paper we focus on a social choice problem that has received less attention in the literature.

Imagine that a group of individuals seek to partition a finite set X into two subsets. The individuals may disagree over the partition and so an aggregation rule is applied to determine a compromise outcome. The outcome of the aggregation could itself be a partition of X into two subsets, but it may be something more general than this. The only constraint we impose is that neither of these two subsets can contain all of the elements in X.

An individual's opinion of how X should be partitioned is described by a non-constant function $v_i : X \to \{0, 1\}$. All individuals assign 1 to those alternatives in X that correspond to the same part of the proposed group partition. In what follows we call this part the set of alternatives that are "collectively approved of". The other part is the set of alternatives that are "collectively disapproved of".

Here are two examples that illustrate the importance of aggregating partitions.[2] Our first example is taken from Kasher and Rubinstein (1997).

Example 1. Who is a J?

Consider the following "group identification problem". Each member of a group makes a judgment as to which members of that group have a certain property. This property could be a religious affiliation, for instance. The individuals agree that at least one member of the group has the property, but not all. The group then seeks to aggregate these judgments on who has the given property into a collective judgment. How should this collective judgment be determined?

Our second example is taken from List (2008).

[2] A literature exists on this problem, inspired by issues of classification. Objects may possess different attributes, and each attribute partitions the set of objects into equivalence classes. How should these partitions be aggregated to determine which objects are, in fact, equivalent? In this vein, impossibility theorems were discovered by Mirkin (1975), Leclerc (1984) and Fishburn and Rubinstein (1986). See also Barthélemy et al. (1986), Dimitrov et al. (2009) and Chambers and Miller (2010).

Example 2. Which worlds are possible?

In logic propositions are sometimes modelled as sets of possible worlds as opposed to sentences in a formal language. Let X denote a finite set containing at least three worlds. A proposition is a tautology if every world is possible. A proposition is a contradiction if every world is impossible. Non-contradictory and non-tautological propositions are, therefore, represented by non-empty, strict subsets of X. Imagine that each member of a group accepts a single non-contradictory and non-tautological proposition. These beliefs can be represented by a function $v_i : X \to \{0, 1\}$. To satisfy the aforementioned constraints, this function cannot be constant. The group then seeks to aggregate these individual beliefs to form a collective judgment. How should this set of worlds be determined?

2 Aggregation rules

Imagine that X contains five alternatives and that there are five individuals. We can create a 5×5 matrix where the rows represent the individuals and the columns represent the alternatives. The elements of the matrix are either 0 or 1. The element 1 in the (i, j) position of the matrix indicates that v_i assigns 1 to the alternative represented by the j^{th} column. Conversely, the element 0 indicates that v_i assigns 0 to the alternative represented by the j^{th} column. This framework goes back to Wilson (1975).[3]

Suppose that $X = \{v, w, x, y, z\}$. Consider the following profile.

[3] A recent application of this framework is Dokow and Holzman (2010).

v	w	x	y	z
1	0	0	0	1
0	1	1	1	1
1	1	1	1	0
1	1	0	1	0
1	0	1	1	1

Table 1: A profile.

How should the group partition be determined? One possibility is that an alternative is collectively approved if and only if the number of individuals who approve of it (assign it the value 1) exceeds the number of individuals who disapprove of it (assign it the value 0). However, in our example, this implies that v, w, x, y and z are all collectively approved of. This contradicts the requirement that not every alternative be collectively approved of. Majority voting is, therefore, not a legitimate aggregation rule for this kind of aggregation problem.

However, consider the following rule. First, assign each alternative a number that is equal to the number of individuals who assign 1 to this alternative. Second, add up these numbers and divide by the cardinality of X. For obvious reasons, we call this number "the mean". The rule then says that an alternative is collectively approved if and only if its number is greater than the mean, and it is collectively disapproved if and only if its number is less than the mean. If its number equals the mean then it is neither approved nor disapproved (one possible interpretation is that the group is indifferent as to whether it should be approved or disapproved).

As we can see from the example, the mean is $\frac{17}{5}$. So v and y are approved of and w, x and z are disapproved of.

Duddy and Piggins (2010) call this rule the "Mean Rule" and provide an axiomatic characterisation of it. Unlike majority voting, the Mean Rule is a legitimate aggregation rule. Under this rule it is impossible for all alternatives to be in one part of the group partition (of course, both parts can be empty).

Although this rule can be defended on normative grounds, it is manipulable in the following sense. Consider the following profile.

v	w	x	y	z
0	0	1	1	1
0	1	0	1	1
1	0	1	1	1
1	0	1	1	0
0	1	0	1	1

Table 2: Another profile.

Under the Mean Rule, y and z are approved of and v, w and x are disapproved of. Note that the individual represented by the first row approves of x. It is not, however, approved of by the group. Imagine that this individual submits a new partition. This leads to the following profile.

v	w	x	y	z
0	0	1	0	0
0	1	0	1	1
1	0	1	1	1
1	0	1	1	0
0	1	0	1	1

Table 3: Manipulation.

Under the Mean Rule, y, z and x are approved of and v and w are disapproved of. As we can see, by changing her partition the first individual can ensure that x is approved of, other things equal. If her sincere partition is the one represented in Table 2 then we can interpret this change in the group outcome as a profitable misrepresentation. The Mean Rule is manipulable.[4]

[4]Manipulability in this sense is similar to a concept developed by Dietrich and List (2007) in the context of judgment aggregation.

The question then arises as to which aggregation rules are non-manipulable in this setting. The answer, in the spirit of the Gibbard-Satterthwaite theorem[5], is that only dictatorships are non-manipulable. This is what we prove in this paper.

3 Model

X is a finite set containing at least three alternatives. $N = \{1, ..., n\}$ with $n \geq 2$ is finite set of numbers that represent the individuals in society.

An evaluation is a function $v : X \to \{0, 1\}$ that is not constant. Each individual's proposed partition can be represented by an evaluation.

We allow the social outcome to be more general than a partition. A $[0, 1]$-evaluation is a function $w : X \to [0, 1]$ such that, for all $x, y \in X$ and each $c \in \{0, 1\}$, if $w(z) = c$ for all $z \in X - \{x, y\}$ then $w(x) = 1-c$ or $w(y) = 1-c$. This means that if a $[0, 1]$-evaluation takes a value of one (zero) at all but two of the alternatives then it takes a value of zero (one) at at least one of those remaining two alternatives.[6]

$\mathcal{V}$ denotes the set of all evaluations and $\mathcal{W}$ denotes the set of all $[0, 1]$-evaluations. Note that $\mathcal{V} \subset \mathcal{W}$.

A profile is an n-tuple of evaluations, with $\mathcal{V}^n$ being the set of all profiles.

An aggregation rule (more simply, a rule) is a function $f : \mathcal{V}^n \to \mathcal{W}$.

Given profiles $V, V' \in \mathcal{V}^n$ we write $V = (v_1, \ldots, v_n)$, $V' = (v'_1, \ldots, v'_n)$ and so on. And let $f(V) = w$, $f(V') = w'$ and so on.

Define a binary relation $\trianglerighteq$ on $[0, 1]$ as follows. The relation $\trianglerighteq$ is equal to the union of $\{(1, 1), (0, 0)\}$ and $(0, 1] \times [0, 1)$. Let $\triangleq$ be the symmetric part of $\trianglerighteq$. Clearly $\trianglerighteq$ is a weak ordering on $[0, 1]$, with $1 \trianglerighteq \alpha \triangleq \beta \trianglerighteq 0$ for all $\alpha, \beta \in (0, 1)$.

[5]Gibbard (1973) and Satterthwaite (1975).

[6]The fact that the co-domain of w is $[0, 1]$ allows for an interpretation of $w(x)$ in terms of fuzzy set theory. Then $w(x)$ is the "degree" to which x is collectively approved of. This interpretation is entirely optional, however.

We now present several properties that aggregation rules may satisfy.

Non-manipulable (NM). For all $x \in X$, all $i \in N$ and all $V, V' \in \mathcal{V}^n$, if V and V' differ at the ith component only then $v_i(x) = 1$ implies $w(x) \trianglerighteq w'(x)$, while $v_i(x) = 0$ implies $w'(x) \trianglerighteq w(x)$.

The non-manipulation condition says the following. Imagine that individual i unilaterally changes her partition. Then, at the new profile, the following must be true. Consider those alternatives that this individual approved of at the original profile. If any of them were collectively disapproved of at this profile then their social value cannot now increase at the new profile. Similarly, if any of them received a social value less than 1 at the original profile, then they cannot now be collectively approved of at the new profile.

Next, consider those alternatives that this individual disapproved of at the original profile. If any of them were collectively approved of at this profile then their social value cannot now decrease at the new profile. Similarly, if any of them received a social value greater than 0 at the original profile, then they cannot now be collectively disapproved of at the new profile.

This is a natural condition of non-manipulation in this setting and it is consistent with our earlier example.

The following two conditions are standard.

Non-imposition (NI). For every $x \in X$, there exists $V \in \mathcal{V}^n$ such that $w(x) = 0$.

Dictatorial. There exists $i \in N$ such that, for all $V \in \mathcal{V}^n$, $w = v_i$.

4 Theorem

Our theorem states the following.

Theorem. *A non-manipulable rule satisfies non-imposition if and only if it is dictatorial.*

We write $v_N(x) = a$ if $v_i(x) = a$ for all $i \in N$. Similarly, by $v_N(x) \geq v'_N(x)$ we mean $v_i(x) \geq v'_i(x)$ for all $i \in N$, by $v_N(x) \neq v'_N(x)$ we mean $v_i(x) \neq v'_i(x)$ for all $i \in N$ and so forth. Given a subset S of X, we write $v(S) = a$ if $v(x) = a$ for all $x \in S$.

The following condition features in Lemma 1.

Monotone independence (MI). For all $x \in X$ and all $V, V' \in \mathcal{V}^n$, if $v_N(x) \geq v'_N(x)$ then $w(x) \trianglerighteq w'(x)$.

This condition says the following. Take any two profiles V, V' and any alternative x. If every individual who approves of x at V' still approves of x at V, and every individual who disapproves of x at V' either still disapproves of x at V or now approves of x at V, then (i) if the social value of x was greater than 0 at V' then it cannot now fall to 0 at V, and (ii) if x was collectively approved of at V' then its social value cannot now fall below 1 at V.

Lemma 1. *Every non-manipulable rule satisfies monotone independence.*

Proof. Let f be a rule that is NM. Assume, by way of contradiction, that f is not MI. Therefore, $\exists x \in X$ and $\exists V, V' \in \mathcal{V}^n$ with $v_N(x) \geq v'_N(x)$ but $w(x) \ntrianglerighteq w'(x)$. Consider the following sequence of profiles, beginning at $V^{(0)} = V$ and ending at $V^{(n)} = V'$.

$$
\begin{aligned}
V^{(0)} &= (v_1, \ldots, v_n), \\
V^{(1)} &= (v'_1, v_2, \ldots, v_n), \\
V^{(2)} &= (v'_1, v'_2, v_3, \ldots, v_n), \\
&\cdots \\
V^{(n)} &= (v'_1, \ldots, v'_n).
\end{aligned}
$$

Since $\trianglerighteq$ is transitive and $f(V^{(0)})(x) \ntrianglerighteq f(V^{(n)})(x)$, there must exist $j \in N$ such that $f(V^{(j-1)})(x) \ntrianglerighteq f(V^{(j)})(x)$. Let $V^{(j-1)} = (v_1^\alpha, \ldots, v_n^\alpha)$ and $V^{(j)} = (v_1^\gamma, \ldots, v_n^\gamma)$.

We have, by construction, $v_N^\alpha(x) \geq v_N^\gamma(x)$ and so $v_j^\alpha(x) \geq v_j^\gamma(x)$. Therefore, we have $v_j^\gamma(x) = 0$ or $v_j^\alpha(x) = 1$. So we have two profiles $(v_1^\alpha, \ldots, v_n^\alpha)$ and $(v_1^\gamma, \ldots, v_n^\gamma)$ that differ at the jth component only and we know that either (i) $v_j^\gamma(x) = 0$ and $w^\alpha(x) \not\trianglerighteq w^\gamma(x)$ or (ii) $v_j^\alpha(x) = 1$ and $w^\alpha(x) \not\trianglerighteq w^\gamma(x)$. NM is violated in either case. Loosely speaking, if (i) is true then NM is violated in the move from $(v_1^\gamma, \ldots, v_n^\gamma)$ to $(v_1^\alpha, \ldots, v_n^\alpha)$ because individual j has successfully lowered the social value assigned to x. If (ii) is true then NM is violated in the move from $(v_1^\alpha, \ldots, v_n^\alpha)$ to $(v_1^\gamma, \ldots, v_n^\gamma)$ because individual j has successfully raised the social value assigned to x. □

The following condition features in Lemma 2.

Unanimity. For all $x \in X$ and all $V \in \mathcal{V}^n$, $v_N(x) = 0$ implies $w(x) = 0$, and $v_N(x) = 1$ implies $w(x) = 1$.

Lemma 2. *If a non-manipulable rule satisfies non-imposition then it satisfies unanimity.*

Proof. Let f be a rule that is NM and NI. Take any $x \in X$. NI implies that there exists $V \in \mathcal{V}^n$ with $w(x) = 0$. Consider a profile $V' \in \mathcal{V}^n$ with $v'_N(x) = 0$. We know by Lemma 1 that since f is NM it is also MI. Since $v_N(x) \geq v'_N(x)$, MI implies that $w(x) \trianglerighteq w'(x)$. Hence, $w'(x) = 0$.

Consider a profile $V'' \in \mathcal{V}^n$ with $v''_N(x) = 1$. Let $V^* \in \mathcal{V}^n$ be the profile with $v^*_N(x) = 1$ and $v^*_N(X - \{x\}) = 0$. We have seen that $w^*(X - \{x\}) = 0$. Therefore, recalling the definition of a $[0,1]$-evaluation, we have $w^*(x) = 1$. Since $v''_N(x) \geq v^*_N(x)$, MI implies that $w''(x) \trianglerighteq w^*(x)$. Hence, $w''(x) = 1$. □

The following condition features in Lemma 3.

Neutrality. For all $x, y \in X$ and all $V, V' \in \mathcal{V}^n$, (i) $v_N(x) = v'_N(y)$ implies $w(x) \triangleq w'(y)$, and (ii) $v_N(x) \neq v'_N(y)$ implies $w(x) = 0 \leftrightarrow w'(y) = 1$.

This condition says the following. If everyone's evaluation of x at V is the same as their evaluation of y at V', then the social value assigned to x at

V is in the same equivalence class as the social value assigned to y at V'. Furthermore, if everyone's evaluation of x at V is different to their evaluation of y at V', then if x is collectively disapproved at V then y must be collectively approved at V'. Moreover, in this situation, if y is collectively approved at V' then x must be collectively disapproved at V.

Lemma 3. *If a non-manipulable rule satisfies non-imposition then it satisfies neutrality.*

Proof. Let f be a rule that is NM and NI. Take any $x, y \in X$ and any $V, V' \in \mathcal{V}^n$, such that either $v_N(x) \neq v'_N(y)$ or $v_N(x) = v'_N(y)$. We consider four cases.

Case 1: $x \neq y$ and $v_N(x) \neq v'_N(y)$. Construct a profile V^* with $v^*_N(x) = v_N(x)$, $v^*_N(y) = v'_N(y)$ and $v^*_N(X - \{x, y\}) = 0$. Unanimity implies that $w^*(X - \{x, y\}) = 0$. So, recalling the definition of a $[0, 1]$-evaluation, if $w^*(x) = 0$ then $w^*(y) = 1$.

MI implies that $w(x) \triangleq w^*(x)$ and that $w'(y) \triangleq w^*(y)$. This means that $w(x) = 0 \leftrightarrow w^*(x) = 0$ and that $w'(y) = 1 \leftrightarrow w^*(y) = 1$. So, given that $w^*(x) = 0 \to w^*(y) = 1$, it must then be true that $w(x) = 0 \to w'(y) = 1$.

To see that $w(x) = 0 \leftarrow w'(y) = 1$ simply interchange everywhere 0 and 1, x and y, and the profiles V and V', in the above argument.

Case 2: $x \neq y$ and $v_N(x) = v'_N(y)$. Take any $z \in X - \{x, y\}$ and consider a profile V'' where $v''_N(z) \neq v_N(x)$. It is also true then that $v''_N(z) \neq v'_N(y)$. By the argument used in Case 1 we know that $w''(z) = 0 \leftrightarrow w(x) = 1$ and $w''(z) = 1 \leftrightarrow w(x) = 0$ and that $w''(z) = 0 \leftrightarrow w'(y) = 1$ and $w''(z) = 1 \leftrightarrow w'(y) = 0$. It follows that $w(x) = 0 \leftrightarrow w'(y) = 0$ and $w(x) = 1 \leftrightarrow w'(y) = 1$. In other words, $w(x) \triangleq w'(y)$.

Case 3: $x = y$ and $v_N(x) = v'_N(y)$. MI implies immediately that $w(x) \triangleq w'(y)$.

Case 4: $x = y$ and $v_N(x) \neq v'_N(y)$. Take any $z \in X - \{x\}$ and any profile V^{**} with $v^{**}_N(z) = v_N(x)$. We know, by Case 2, that $w(x) = 0 \leftrightarrow w^{**}(z) = 0$.

And, since $v_N^{**}(z) \neq v'_N(y)$, we know, by Case 1, that $w^{**}(z) = 0 \leftrightarrow w'(y) = 1$. It follows that $w(x) = 0 \leftrightarrow w'(y) = 1$. $\square$

Proof of the main theorem. Take any $x \in X$, and profiles V and V^* with $v_N(x) = 0$ and $v_N^*(x) = 1$. Consider the following sequence of profiles, beginning at $V^{(0)} = V$ and ending at $V^{(n)} = V^*$.

$$\begin{aligned}
V^{(0)} &= (v_1, \ldots, v_n),\\
V^{(1)} &= (v_1^*, v_2, \ldots, v_n),\\
V^{(2)} &= (v_1^*, v_2^*, v_3, \ldots, v_n),\\
&\cdots\\
V^{(n)} &= (v_1^*, \ldots, v_n^*).
\end{aligned}$$

Unanimity implies that $f(V^{(0)})(x) = 0$ and $f(V^{(n)})(x) = 1$. So there must exist $d \in N$ such that $f(V^{(d-1)})(x) = 0$ and $f(V^{(d)})(x) > 0$. Let $V^{(d-1)} = (v_1^\alpha, \ldots, v_n^\alpha)$ and $V^{(d)} = (v_1^\gamma, \ldots, v_n^\gamma)$.

Construct a profile V' with $v'_N(X - \{x, y\}) = v_N^\alpha(x)$, $v'_N(y) + v_N^\gamma(x) = 1$ and $v'_d(x) = 1$ while $v'_i(x) = 0$ for all $i \in N - \{d\}$. Neutrality implies that, since $w^\alpha(x) = 0$, we have $w'(X - \{x, y\}) = 0$. Neutrality also implies that, since $w^\gamma(x) > 0$, we have $w'(y) < 1$. Recalling the definition of a $[0, 1]$-evaluation, we see that $w'(x) = 1$. This is despite the fact that $v'_i(x) = 0$ for all $i \in N - \{d\}$.

Take a profile $V'' \in \mathcal{V}^n$ with $v''_N(x) \neq v'_N(x)$. Neutrality implies that $w''(x) = 0$. Take any profile $V^{**} \in \mathcal{V}^n$. If $v_d^{**}(x) = 1$ then, since this would mean that $v_N^{**}(x) \geq v'_N(x)$, MI implies that $w^{**}(x) = 1$. If $v_d^{**}(x) = 0$ then, since this would mean that $v_N^{**}(x) \leq v''_N(x)$, MI implies that $w''(x) = 0$. So we can see that the collective value assigned to x will always be equal to the value that individual d assigns to x. Since the rule satisfies neutrality, this must also be true of every other alternative in X. $\square$

References

[1] Barberà, S., 2010. Strategy-proof social choice, Barcelona Economics Working Paper Series.

[2] Barthélemy, J-P, Leclerc, B. and B. Monjardet, 1986. On the use of ordered sets in problems of comparison and consensus of classifications, Journal of Classification 3, 187-224.

[3] Chambers, C.P. and A.D. Miller, 2010, Rules for aggregating information, forthcoming in Social Choice and Welfare DOI 10.1007/s00355-010-0466-5.

[4] Dietrich, F. and C. List, 2007. Strategy-proof judgment aggregation, Economics and Philosophy 23, 269-300.

[5] Dimitrov, D., Marchant, T. and D. Mishra, 2009. Separability and aggregation of equivalence relations, Working paper.

[6] Dokow, E. and R. Holzman, 2010. Aggregation of binary evaluations, Journal of Economic Theory 145, 495–511.

[7] Duddy, C. and A. Piggins, 2010. Aggregating partitions, Working paper, Department of Economics, National University of Ireland Galway.

[8] Fishburn, P.C. and A. Rubinstein, 1986. Aggregation of equivalence relations, Journal of Classification 3, 61-65.

[9] Gibbard, A.F., 1973. Manipulation of voting schemes: a general result, Econometrica 41, 587-601.

[10] Kasher, A. and A. Rubinstein, 1997. On the question "who is a J?": a social choice approach, Logique et Analyse 160, 385-395.

[11] Leclerc, B., 1984. Efficient and binary consensus functions on transitively valued relations, Mathematical Social Sciences 8, 45-61.

[12] List, C., 2008. Which worlds are possible? A judgment aggregation problem, Journal of Philosophical Logic 37, 57-65.

[13] Mirkin, B.G., 1975. On the problem of reconciling partitions, in Blalock, H.M. (ed) Quantitative sociology: international perspectives on mathematical and statistical modeling, Academic Press: New York.

[14] Satterthwaite, M., 1975. Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions, Journal of Economic Theory 10, 187-217.

[15] Wilson, R. 1975. On the theory of aggregation, Journal of Economic Theory 10, 89-99.