Mathematical Social Sciences 79 (2016) 6–10

Contents lists available at ScienceDirect

Mathematical Social Sciences

journal homepage: www.elsevier.com/locate/econbase

A theorem on aggregating classifications[☆]

François Maniquet[a], Philippe Mongin[b,*]

[a] CORE, Université Catholique de Louvain, Belgium
[b] CNRS & HEC Paris, France

ARTICLE INFO

Article history:
Received 23 October 2014
Received in revised form
8 October 2015
Accepted 8 October 2015
Available online 19 October 2015

ABSTRACT

Suppose that a group of individuals must classify objects into three or more categories, and does so by aggregating the individual classifications. We show that if the classifications, both individual and collective, are required to put at least one object in each category, then no aggregation rule can satisfy a unanimity and an independence condition without being dictatorial. This impossibility theorem extends a result that Kasher and Rubinstein (1997) proved for two categories and complements another that Dokow and Holzman (2010) obtained for three or more categories under the condition that classifications put at most one object in each category. The paper discusses an interpretation of its result both in terms of Kasher and Rubinstein's group identification problem and in terms of Dokow and Holzman's task assignment problem.

1. Introduction

While preference aggregation still looms large in the agenda of social choice theory, there is a small, but growing body of literature on the aggregation of *classifications*. The general scheme is that the members of a group each propose dividing a given set of objects into categories, and that a collective division of the set results from these individual proposals by respecting various conditions of association, which are partly reminiscent of those usually defined for preference aggregation. In one version of this scheme, which appears to date back to Mirkin (1975), the individuals and the collective can partition the set in any possible ways. (See Chambers and Miller, 2011 and Dimitrov et al., 2012 for recent developments; the latter paper also surveys the field.) In another version, which can be traced to Kasher and Rubinstein (1997), there is a given list of designated categories in which the objects must be fitted. This version has been explored, both by Kasher and Rubinstein and followers, in the particular case where the objects to be classified are the very individuals who propose the classifications. As a typical application, some countries legally divide their citizens according to racial, ethnic or religious criteria. Since the citizens themselves have opinions on how this should be done, one may investigate how the legal division should reflect these opinions. Put in axiomatic form, this has come to be called the *group identification problem*. (See, among others, Samet and Schmeidler, 2003; Dimitrov et al., 2007 and Miller, 2008.)

The present paper investigates the aggregation of classifications with designated categories, and hence belongs to the second branch of analysis, but does not pursue the group identification problem specifically. Rather, it proves an impossibility theorem for this second branch at large. In a nutshell, if there are $p \geq 3$ categories in the list and $m \geq p$ objects to be classified in these categories, and if moreover both individual and collective classifications satisfy the *surjectivity* (ontoness) restriction that each category is filled with at least one object, then the collective classifications are dictatorial if they satisfy a unanimity and an independence condition. The unanimity condition says that if the individuals in a profile agree on how to classify an object, the aggregate for this profile endorses the agreed on classification. The independence condition says that if there are two profiles and each individual classifies an object identically in both of them, the corresponding two aggregates also classify the object identically. These two conditions are reminiscent of familiar ones in preference aggregation, but have a *unary* form, which leads to a distinctive analytical treatment.

Kasher and Rubinstein (1997, Theorem 2), have stated this theorem for the special case $m \geq p = 2$. They relate it to the group identification problem, but their proof is in fact independent

☆ The authors thank Franz Dietrich, Ron Holzman, Marcus Pivato, the referees and the associate editor for useful comments on previous versions.

* Correspondence to: GREGHEC, 1 rue de la Libération, F-78350 Jouy-en-Josas, France. Tel.: +33 139677253; fax: +33 139677076.
E-mail address: mongin@greg-hec.com (P. Mongin).

http://dx.doi.org/10.1016/j.mathsocsci.2015.10.001
0165-4896/

of this context.[1] Thus, our work can be seen as an extension of theirs. To free the impossibility result from the limitation to two categories is a non-trivial step, as will appear from the proof and technical comments below. Although Kasher and Rubinstein have not emphasized this point, it is essential to the impossibility that the categories, whatever their number $p \geq 2$, are never left empty by either the individual or the collective classifications. This is the surjectivity restriction mentioned above.

If one is ultimately interested in the group identification problem, this is a natural restriction to consider. One may expect lawmakers to confer legal status on a social category only if they believe it to be applicable at least to some citizens, and in a country where democratic principles hold, one may further expect that the categories have been agreed on between the lawmakers and the citizens prior to being used in practice. Accordingly, citizens would no more than lawmakers leave any category unfilled, even though they would no doubt disagree on its precise extension. The group identification literature alludes to political examples that seem to warrant this analysis. Kasher and Rubinstein (1997) implicitly draw inspiration from the legal religious denominations in Israel, and Miller (2008) explicitly mentions the racial divisions recognized by the US Census. If Israeli or US citizens were asked to classify a significant sample of their respective populations, they would be very unlikely to leave any of the available categories vacuous, except perhaps for strategic purposes that we will not consider in this paper. Concerning the group identification problem, our view is that the most troublesome idealizing assumption is not surjectivity, but the very form of the poll, which requires each citizen to classify any other, whereas most political examples only involve self-designation.

If one is not particularly interested in this problem, one may turn to more direct cases of aggregating classifications for which surjectivity appears to be appropriate. Consider a panel of astronomers who meet to classify distant celestial bodies into stars, exoplanets, brown dwarfs and other less identifiable objects. Each astronomer proposes his own classification, and the chair tries to turn these individual data into an authoritative classification. The classification is well-established on prior grounds, so if the set of celestial bodies under consideration is large enough, neither the individual astronomers nor the chair will leave any of the four categories empty.[2] This is of course a theoretical example, but it is worth noting that the status of celestial bodies is currently discussed at a collective level, with aggregative steps – typically votes – being sometimes taken (for an intriguing account of the discussions surrounding Pluto, see Marschall and Maran, 2009).

We will provide further motivation for surjectivity while interpreting our framework in terms of a *collective task assignment problem*, as in Dokow and Holzman (2010). Having this interpretation in view, these authors investigate the same problem of aggregating classifications as ours, but make the opposite assumption that there are $m \geq 3$ objects and $p \geq m$ positions. They show that if individual and collective classifications satisfy the *injectivity* restriction that each category is filled with at most one object, then the collective classifications are dictatorial if they satisfy unanimity (in a reinforced version) and independence. The two impossibility results complement each other very naturally. Dokow and Holzman's actually belongs to an abstract theory of nonbinary evaluations, which they develop for its own sake, and we had borrowed this powerful apparatus to carry out our first proof (Maniquet and Mongin, 2014). For ease of exposition, we have shifted here to the language and ultrafilter proof technique of standard social choice theory, but the interested reader may consult this earlier version, which also discusses the connections between social choice theory and the recently developed judgment aggregation theory.

[1] This two-category case is a corollary to an impossibility theorem proved by Rubinstein and Fishburn (1986, Theorem 3).

[2] Notice that surjectivity here follows as a fact of the situation, and not on normative grounds. A referee alerted us to this distinction.

2. The formal setup and the theorem

There are a set $N = \{1, \ldots, n\}$ of individuals, a set $X = \{1, \ldots, m\}$ of objects, and a set $P = \{1, \ldots, p\}$ of positions (or categories), with $p \geq 3$. The individuals classify the objects by putting each of them in a position. Formally, classifications are mappings $X \to P$. By assumption, there are at least as many objects as positions, and each classification assigns at least one object to each position. Formally, $m \geq p$, and the set of classifications is the *surjectivity* (ontoness) *domain*:

$$\mathcal{C} = \{k : X \to P \mid \forall r \in P, \exists x \in X : k(x) = r\}.$$

An *aggregation function* associates a social classification with any profile of individual classifications:

$$F : \mathcal{C}^n \to \mathcal{C}, (c_1, \ldots, c_n) \mapsto F(c_1, \ldots, c_n).$$

We abridge $F(c_1, \ldots, c_n)$, $F(c'_1, \ldots, c'_n)$,..., as c, c'.... The definition of F encapsulates a universal domain condition. We introduce three more conditions axiomatically. *Independence* requires that if an object occupies the same positions in two profiles of individual classifications, x occupies the same position in the associated social classifications.

Condition 1. Independence*: For all $(c_1, \ldots, c_n)$, $(c'_1, \ldots, c'_n) \in \mathcal{C}^n$ and all $x \in X$, if for all $i \in N$, $c_i(x) = c'_i(x)$, then $c(x) = c'(x)$.*

Unanimity requires that if all individual classifications in a profile give an object the same position, the social classification give it that position.

Condition 2. Unanimity*: For all $(c_1, \ldots, c_n) \in \mathcal{C}^n$, all $x \in X$, all $r \in P$, if for all $i \in N$, $c_i(x) = r$, then $c(x) = r$.*

The last condition states that one individual imposes his classification to society.

Condition 3. Dictatorship*: There is $j \in N$ such that for all $(c_1, \ldots, c_n) \in \mathcal{C}^n$, $c = c_j$.*

Independence and *Unanimity* are reminiscent of Independence of Irrelevant Alternatives and the Pareto conditions in Arrovian social choice theory. They can be defended normatively by roughly parallel arguments—*Independence* being connected with computational ease and nonmanipulability, and *Unanimity* with the individuals' sovereignty. Dictatorship is meant to be as undesirable here as it is there. Notice however that the present conditions are *unary*, i.e., bear on one object at the time, as suits a classification aggregation problem, whereas the Arrovian conditions are *binary*, as suits a preference aggregation problem.

Theorem 1. *If an aggregation function F satisfies* Independence *and* Unanimity, *it satisfies* Dictatorship.

The proof of the theorem consists in showing that the set of decisive subsets of N is an ultrafilter. In the present context, a subset of N is *decisive* if for every profile, every object and every position, when the profile is such that all individuals in this subset agree to put the given object in the given position, then society endorses this agreement. This notion of a decisive subset appears only as Definition 4 in the course of the proof. We first introduce weaker variant notions of decisiveness that are graded in logical strength, i.e., Definitions 1–3, exploring their properties in