CENTRE DE RECHERCHE SUR LES STRATÉGIES ÉCONOMIQUES

On compensation schemes for data sharing within the european REACH legislation

SYLVAIN BÉAL, MARC DESCHAMPS

February 2014

Working paper No. 2014 – 01

CRESE
30, avenue de l'Observatoire
25009 Besançon
France
http://crese.univ-fcomte.fr/

UNIVERSITÉ DE FRANCHE-COMTÉ

On compensation schemes for data sharing within the european REACH legislation*

Sylvain Béal‡, Marc Deschamps§,

February 7, 2014

Abstract

Article 30 of Regulation (EC) No 1907/2006 concerns the sharing of data between users of a chemical substance. We study this bargaining problem by means of a special class of games in coalitional form called data games (Dehez and Tellone [10]). For such problems, compensation schemes specify how the data owners should be compensated by the agents in needs of data. On the class of data games, the Core, the Nucleolus and the Shapley value provide relevant compensation schemes. We provide three comparable axiomatic characterizations of the set of all (additive) compensation schemes belonging to the Core, of the Nucleolus and of the Shapley value. The axioms reflects principles of various theories of justice.

Keywords: REACH, Data sharing problem, Core, Nucleolus, Shapley value, Pooling, Compensation, Reasonableness, Invariance to deleting non-exclusive data, Equal treatment of Equals, Invariance to enlarging the owner set, Equal concessions.

JEL Classification number: C71, D71, K32, L65.

1 Introduction

The chemical industry includes a wide range of products and activities covering upstream to downstream, base chemicals, specialty chemicals and pharmaceuticals. Nowadays, it appears very difficult to find products that do not contain chemicals in their production in most societies. For example, plastics derived from petrochemical processing bases are widely used in packaging, buildings and automobiles. The European Commission estimates that more than one hundred thousand chemicals are currently used within the European Union. In Europe, the chemical industry is a major economic sector. According to Cefic (European Chemical Industry Council), in 2012 it produced more than 21% of the chemicals sold in the world, is a net exporter, directly employs 1.2 million workers in 29,000 companies and contributes to the European economy by 673 billion euros.

As a consequence, european governments are confronted with two major challenges. The first is to ensure the health and environmental management of each chemical product as well as their combinations, especially in a context where European citizens are concerned about these issues. The second lies in the fact that the chemical industry plays a central role in European economies and is an asset for economic growth.

It is to meet these challenges simultaneously that the European institutions have decided, since December 13, 2006, the implementation of a new harmonized legislative framework in the field of chemical industry: the European Regulation (EC) No 1907/2006 REACH (Registration, Evaluation, Authorization and Restriction of Chemicals), which ultimately aims to ensure greater human and environmental safety (based on impact studies by the European Commission it will reduce the number of deaths due to cancer by a range of 2000 to 4000 per year and lead to a reduction in public health spending of up to 50 billion over thirty years), while preserving and enhancing the competitiveness of the European chemical industry. These objectives are explicitly set out in its Article 1: "This

*Financial support by the National Agency for Research (ANR) — research programs "DynaMITE" and "DAMAGE"— is gratefully acknowledged. We also thank Philippe Solal for constructive comments.

‡Corresponding author. Université de Franche-Comté, CRESE, 30 Avenue de l'Observatoire, 25009 Besançon, France. E-mail: sylvain.beal@univ-fcomte.fr. Tel: (+33)(0)3-81-66-68-26. Fax: (+33)(0)3-81-66-65-76.

§Université de Nice Sophia-Antipolis, GREDEG (CNRS UMR 7321) and BETA (CNRS 7522), France. E-mail: marc.deschamps@univ-lorraine.fr.

Regulation should ensure a high level of protection of human health and the environment as well as the free movement of substances, on their own, in preparations and in articles, while enhancing competitiveness and innovation". In other words, REACH aims to transform the European chemical industry into the world's first sustainable chemical industry. According to the European institutions, about 30,000 chemicals are directly affected by this legislation.

The REACH Regulation entered into force on June 1st, 2007 and sets out the principle that it is for manufacturers, importers and downstream users "to ensure that they manufacture, place on the market or use such substances that do not adversely affect human health or the environment. Its provisions are underpinned by the precautionary principle" (Article 1, §3). The effectiveness of this principle is guaranteed by the fact that all substances, preparations or articles which do not comply with the Regulation may not be marketed within the European Union. In practical terms, it is the ECHA (European Chemicals Agency), a European Agency established by the Regulation, which is responsible for implementing the REACH legislation. It is to ensure that all existing and future chemicals produced or imported in quantities over one ton per year in the European Union are registered, evaluated and authorized, except in cases that are expressly exempted (e.g radioactive substances).

The most innovative aspect of this Regulation lies in the fact that companies are forced to provide all existing data on the properties of the chemicals they use when registering with the ECHA. The procedure provides that, without data, companies cannot use the substances: "No data, no market", summarizes ECHA. To facilitate the exchange of toxicological and ecotoxicological information, to avoid the duplication of tests including those involving vertebrate animals, and to allow different registrants to agree on a classification and labeling of a substance, the Regulation created a transitional regime. During this period, any company wishing to declare a substance must participate in a SIEF (Substance Information Exchange Forum). Within a SIEF, members are free to decide on legal form, organization and communication modalities. The Regulation requires, however, that any member of a SIEF meet the demand for information from another member and that all members work collectively to identify and implement additional studies that may prove necessary. The purpose of SIEF, operational from January 2009 to June 2018 (when all substances will have been recorded), is to allow a joint submission for each substance with a lead registrant, other registrants needing only to refer to this registration and to provide data on the characteristics of its production or its use. In cases where such a joint submission involves too high costs for a company or where it requires it to disclose confidential information, the Regulation provides that it can voluntarily exclude the obligation to participate in the joint submission. Nevertheless, it remains a member of the SIEF and must continue to share its non-confidential data. After the transitional regime, the Regulation insists on the fact that registrants "shall make every effort to ensure that the costs of sharing the information are determined in a fair, transparent and non-discriminatory way" (Article 27, §3), and ECHA decides in case of disagreement between the parties on the necessary compensations. The question of how to realize data sharing within a SIEF and to determine the above-mentioned compensations is therefore a critical issue until June 2018.

In this article, we use games in coalitional form, a classical model to study bargaining, to tackle this question. The coalitional function specifies for each coalition the total cost of acquiring the missing data available in the SIEF should its members cooperate. Such games in coalitional form are called data games and have been introduced in Dehez and Tellone [10]. Since all members of the SIEF have the obligation to cooperate, it is essential to find compensation vectors specifying how the data owners should be compensated by the other SIEF members. A compensation scheme assigns a compensation vector to any data game.

There exists infinitely many compensation schemes, several of which may be considered as intuitively appealing. In order to distinguish among them, we adopt the axiomatic approach and proceed in two steps.[1] Firstly, we list desirable properties for a compensation scheme on the class of data games. These properties conveys principles such as efficiency, equity and fairness. Secondly, various combinations of the properties are tried out so as to eventually retain a unique compensation scheme or a pertinent family of compensation schemes. We believe that the axiomatic method is of particu-

[1] We refer to Thomson [26] for a detailed survey of the axiomatic method.

lar interest in the field of Law and Economics. Judges and arbitrators often adhere to philosophical principles which reflects their political ideologies and serve as a basis for making their decisions. For instance, Kornhauser [17], Spiller [25] and Fleurbaey and Roemer [15] propose an axiomatic approach to the courts' decision making process.

In data sharing problems, the determination of fair compensation schemes also relies on philosophical principles which interpret the ethical vision of the bargainers within a SIEF. On these aspects, theories of justice offer an abundant literature along the lines of Rawls [22], Dworkin [11], Young [29] and Fleurbaey and Maniquet [14], among others. These scholars underline that compensation schemes aim at restoring equality or diminishing individual inequalities. The formal statements of our axioms borrow to a great extent these theories as exemplified in section 2. While ECHA provides a detailed guide on data sharing[2], this guide only rests on very specific examples instead of offering a general and operational approach. Our axiomatic analysis remedies this weakness by formulating properties and investigating compensations schemes for all data sharing problems.

We study compensations schemes stemming from the main three solution concepts used in games in coalitional form. The Core (Gillies [16]) collects those compensation vectors which no coalition of SIEF members can improve upon in the Pareto sense. The Core of data games is always nonempty, but can contain many compensation vectors. It is therefore interesting to extract specific compensation schemes from the Core. The Nucleolus (Schmeidler [23]) is probably the most-known such Core-selector. Roughly speaking, the Nucleolus incorporates Rawls [22] vision of distributive justice in order to be of the greatest benefit to the least-advantaged coalition of SIEF members. Besides Core compensation vectors, there exist other compensation schemes incorporating fair considerations, such as the Shapley value (Shapley [24]). The Shapley value measures, for each SIEF member, his average contribution to all coalitions he can belong to. This article offers three comparable characterizations of the Core-selectors, of the Nucleolus and of the Shapley value in the sense that the combinations of axioms in these characterizations only differ with respect to one or two axioms. These results are outlined in detail in the next section.

Other articles dealing with the data sharing within the REACH legislation via games in coalitional form are Dehez and Tellone [10], which formulates the Core, the Nucleolus and the Shapley value on the class of data games, and Béal *et al.* [4], which provides an alternative axiomatic characterization of the Shapley value and characterizations of equal surplus division values Games in coalitional form have also been employed extensively to tackle problems in Law and Economics but in the very different perspective of measuring voting power. We refer the reader to Bindseil and Hantke [6], Fedeli and Forte [12] and Algaba, Bilbao and Fernández [3] for the distribution of voting power within the various institutions of the European Union, and to Braham and Steffen [7], Leech and Manjón [18] and Casajus, Labrenz and Hiller [9] for studies of voting power in corporate finance.

We are also aware of some limits of our model. The simplicity of the data games comes at a cost, for instance because no reference is made to the underlying problem of reaching a consensus on the cost parameters. Moreover, the model is silent on situations in which the composition of the SIEF members evolves over time. As a consequence, even if our model can be adapted to take these features into account, it should be considered as a first step towards more sophisticated analysis. It is also true that we neglect the economic consequences of other facets of the REACH legislation, on which it is surprising to note the tenuousness of academic work, in particular by economists.[3]

The rest of the article is organized as follows. An informal presentation of the results is given in section 2. The data sharing problems modeling the bargaining within a SIEF is introduced in section 3. This section also presents the class of data games. Section 4 is devoted to the axiomatic study and contains the main results. Concluding remarks are provided in section 5. The logical independence of the axioms used in our results is demonstrated in the Appendix.

[2] `http://echa.europa.eu/documents/10162/13631/guidance_on_data_sharing_en.pdf`

[3] An attempt is made by Canton and Allen [8]. The other few exceptions are Wolf and Delgado [28], Logomasini [19], Ackerman, Stanton and Massey [1], Ackerman *et al.* [2], Lorenz, Lebreton and van Wassenhove [20], Von Holleben, Rosenfeld and Scheidmann [27] and Bergkamp [5] but lack of theoretical foundations.

2 Outline of the results

The first principle that we impose is the unquestionable **compensation** principle, which stipulates that the compensations chosen by the SIEF members (often called agents henceforth) add up to zero. This principle is a restatement for data games of the axiom of (Pareto) efficiency for general games in coalitional form. Our second principle is called **pooling** and possesses a natural bargaining interpretation within the REACH legislation: the outcome of multi-data negotiations does not depend on the agenda chosen by the negotiators. Whether data are discussed separately or all at once does not affect the result of the negotiation. This principle is a variation for data games of the axiom of additivity for general games in coalitional form that Shapley himself supports as follows: "the third axiom ("law of aggregation") states that when two independent games are combined, their values must be added player by player" (Shapley [24], page 309). All the compensation schemes mentioned in the introduction satisfy these first two basic principles. In order to distinguish among them, we invoke several other principles.

The principle of **reasonableness** has been introduced by Milnor [21] and requires that any agent's compensation lies between its worst and best (marginal) contributions to the coalitions he can belong to. These contributions are interpreted by Milnor [21] as the minimal and maximal compensations that an agent can legitimately claim. The principle of **invariance to deleting non-exclusive data** states that a compensation scheme should not be affected if a data held by more than one agent is removed and has a basic non-cooperative economic justification. Assume that a data is not exclusive in the sense that it is possessed by at least two agents. Then these agents somehow compete against each other in order to attract monetary compensations from the agents lacking the data. The situation is therefore similar to a Bertrand oligopoly.[4] The data owners set prices at which the agents lacking the data can buy it. The data owner with the lowest price gets the full demand, *i.e.* it sells to all agents lacking the data. Thus, prices are null at equilibrium since there is a zero marginal cost here. If one adheres to this interpretation, then it seems natural to require an absence of compensation for any non-exclusive data. Proposition 2 states that if a compensation scheme satisfies the pooling principle, then it is a Core-selector if and only if it further satisfies the principles of compensation, reasonableness and invariance to deleting non-exclusive data.

In Proposition 3, the Nucleolus is selected among the set of all Core-selectors. It is sufficient to replace in Proposition 2 the principle of reasonableness by the principle of **equal concessions**. The latter principle only deals with data sharing problems in which every data is exclusive to some agent. The best contribution of an agent to the coalitions he can belong may be interpreted as his utopia claim (especially because most of time the sum of the corresponding compensations is unrealistic, *i.e.* much less than zero). With this view in mind, equal concessions imposes that all agents abandon exactly the same amount from their utopia claims. The principle behind equal concessions has the flavor of the equity based on tangible claims in Young [29], and one of its consequences is to compensate the lack of individual internal resources (*i.e.* the lack of data here) as suggested by Dworkin [11].

Finally, Proposition 4 characterizes the Shapley value by replacing invariance to deleting non-exclusive data in Proposition 3 by invariance to enlarging the owner set and adding equal treatment of equals. The principle of **invariance to enlarging the owner set** stipulates that the price necessary to acquire a data for an agent lacking it should not depend on how many agents initially own it. Invariance to enlarging the owner set and invariance to deleting non-exclusive data belong to the larger category of axioms of consistency, which require invariance of a solution if specific modifications are brought to the problem. **Equal treatment of equals** is the basic principle which imposes to treat equally two agents owning the same data. Since Aristotle, this principle appears under different forms in most of the theories of justice (see Fleurbaey [13] for instance).

[4]The Regulation reminds that data owners agreeing on the price of their data would be placed under the condemnation of the European Union competition law.

3 Data sharing problems and data games

Let $N = \{1, \dots, n\}$ is a finite set of at least three agents involved in a SIEF. We shall use subscripts i and j to refer to agents. A coalition of agents is denoted by S and we refer to 2^N as the set of all coalitions on N. A **data sharing problem** on N is described by a pair (D, O), where

- D is a nonempty finite set of data. Data in D are relevant to the study of the chemical substance associated to the SIEF. Subscripts k and h are used to refer to data. Data $k \in D$ is characterized by a cost $c_k \in \mathbb{R}_+$, which is interpreted as the cost of replicating the data.
- $O : D \longrightarrow N$ is an ownership function assigning to each data k the nonempty set O_k of agents owning the data, and $o_k := |O_k|$. In other words, function O specifies the distribution of property rights within the SIEF. Note that every data is owned by some agent.

We denote by $\mathcal{P}(N)$ the set of all data sharing problems on N. For a given data sharing problem $(D, O) \in \mathcal{P}(N)$, the set of data held by an agent $i \in N$ is denoted by $D_i(O) = \{k \in D : i \in O_k\}$. For each nonempty coalition of agents $S \in 2^N$, $D_S(O) = \cup_{i \in S} D_i(O)$ stands for the set of all data held by the members of S. We shall keep the notations $D_i(O)$ instead of $D_{\{i\}}(O)$ and D instead of $D_N(O)$. For each $i \in N$, $D_i^E(O)$ is the set of data in (D, O) that i exclusively holds, *i.e.* the set, possibly empty, of data held by i and by no other agent in $N \backslash \{i\}$. Thus $D_i^E(O) = D \backslash D_{N \backslash \{i\}}(O)$ or equivalently $D_i^E(O) = \{k \in D : O_k = \{i\}\}$. Define $D^E(O) = \cup_{i \in N} D_i^E(O)$ as the set of exclusive data in (D, O).

To each data sharing problem $(D, O) \in \mathcal{P}(N)$, it is useful to associate a **data sharing game in coalitional form** or simply a **data game** $C_{(D,O)}$ on N, where the characteristic function $C_{(D,O)} : 2^N \longrightarrow \mathbb{R}_+$ assigns to each coalition $S \in 2^N$ the total cost $C_{(D,O)}(S) \in \mathbb{R}_+$ of the data in D that the members of S do not hold. Formally, for each $S \in 2^N$, $S \neq \emptyset$:

$$C_{(D,O)}(S) = \sum_{k \in D \backslash D_S(O)} c_k,$$

and by convention $C_{(D,O)}(\emptyset) = 0$. Observe that $C_{(D,O)}(N) = 0$. We denote by $\mathcal{C}(N)$ the set of all data games that can be constructed from $\mathcal{P}(N)$, *i.e.* $\mathcal{C}(N) = \{C_{(D,O)} : (D, O) \in \mathcal{P}(N)\}$. Each data k generates **elementary data games** $C_{(\{k\},O)} \in \mathcal{C}(N)$.

A **compensation vector** is a vector $x \in \mathbb{R}^n$ which assigns a compensation $x_i \in \mathbb{R}$ to each $i \in N$. If $x_i > 0$, then agent i has to pay a compensation and if $x_i < 0$, then agent i receives a compensation. For any coalition $S \in 2^N$ with at least two members, we shall use the shorter notation $x(S)$ instead of $\sum_{i \in S} x_i$. A **compensation solution** is an application which assigns to each data game a set of compensation vectors. In this article, we are rather interested in compensation schemes, *i.e.* compensation solutions assigning a unique compensation vector to each data game. Formally, a **compensation scheme** is an application f that assigns to each $C_{(D,O)} \in \mathcal{C}(N)$ a unique compensation vector $f(C_{(D,O)}) \in \mathbb{R}^n_+$.

The most well-known compensation scheme is the **Shapley value** (Shapley [24]), which assigns to each data game $C_{(D,O)} \in \mathcal{C}(N)$, the compensation vector

$$Sh_i(C_{(D,O)}) = \sum_{S \in 2^N : S \ni i} \frac{(n - |S|)!(|S| - 1)!}{n!} \big(C_{(D,O)}(S) - C_{(D,O)}(S \backslash \{i\})\big), \quad i \in N.$$

The **Core** (Gillies [16]) is the most-used compensation solution that assigns to each data game $C_{(D,O)} \in \mathcal{C}(N)$ the set of compensation vectors

$$\text{Core}(C_{(D,O)}) = \{x \in \mathbb{R}^n : x(N) = 0 \text{ and } x(S) \leq C_{(D,O)}(S), \forall S \subset N\}.$$

The Core of any data game is nonempty since it always contains the n-dimensional null vector $\mathbf{0} := (0, \dots, 0)$. Nonetheless, the Core often contains many compensation vectors. Therefore, it is interesting to look for a **Core-selector**, *i.e.* a compensation scheme f which selects a Core element $f(C_{(D,O)})$ in every data game $C_{(D,O)} \in \mathcal{C}(N)$. The most famous Core-selector is the **Nucleolus** (Schmeidler [23]), denoted by η, which selects the unique Core compensation vector x that lexicographically minimizes the excesses $x(S) - C_{(D,O)}(S)$ of all coalitions. On the class of data games, the Shapley value, the Core and the Nucleolus have a specific structure as summarized by the following result.

Proposition 1 (Dehez and Tellone [10]) *Consider any data game $C_{(D,O)} \in \mathcal{C}(N)$. Then $Core(C_{(D,O)}) = \sum_{k\in M} Core(C_{(\{k\},O)})$. Moreover,*

- *For each $k \in D$ such that $o_k \geq 2$, it holds that $Core(C_{(\{k\},O)}) = \{\mathbf{0}\}$,*
- *For each $k \in D$ such that $O_k = \{i\}$ for some $i \in N$, it holds that*

$$Core(C_{(\{k\},O)}) = conv(\{z^j(k)\}_{j\in N}),$$

where $z^i(k) = \mathbf{0}$, and $z^j_i(k) = -c_k$, $z^j_j(k) = c_k$ and $z^j_r(k) = 0$ for each $r \in N\backslash\{i,j\}$.

Furthermore, $\eta(C_{(D,O)}) = \mathbf{0}$ if $D^E(O) = \emptyset$ and

$$\eta_i(C_{(D,O)}) = \sum_{k\in D^E(O)} \frac{c_k}{n} - \sum_{k\in D^E_i(O)} c_k, \quad i \in N$$

if $D^E(O) \neq \emptyset$. Finally,

$$Sh_i(C_{(D,O)}) = \sum_{k\in D} \frac{c_k}{n} - \sum_{k\in D_i(O)} \frac{c_k}{o_k}, \quad i \in N.$$

This proposition reveals that the Core and the Nucleolus possess an additive property: they can be computed by looking at elementary data games. The formulas of the Nucleolus and the Shapley value are also close to each other. The Shapley value contains two parts. In the first part, each agent contributes to a communal fund through the same fraction $1/n$ of the total cost of data. In the second part, the accumulated fund is redistributed data by data at an equal rate to each owner. It is easy to figure out that the Nucleolus implements the same principle, but only for the exclusive data.

4 Axiomatic studies

Since data games are build from data sharing problem, it makes sense to formulate axioms depending not only on the structure of data games but also on the primitive of the associated data sharing problems. Our approach consists of mixing these two types of axioms. More specifically, we invoke the following axioms for a compensation scheme f on $\mathcal{C}(N)$, which have already been informally discussed in section 2.

Compensation. $\forall C_{(D,O)} \in \mathcal{C}(N)$, $f(C_{(D,O)})(N) = 0$.

Pooling. $\forall C_{(D,O)} \in \mathcal{C}(N)$ and D', D'' such that $D' \cup D'' = D$ and $D' \cap D'' = \emptyset$, $f(C_{(D,O)}) = f(C_{(D',O)}) + f(C_{(D'',O)})$.

Equal treatment of equals. $\forall C_{(D,O)} \in \mathcal{C}(N)$ and all $i, j \in N$ such that $D_i(O) = D_j(O)$, $f_i(C_{(D,O)}) = f_j(C_{(D,O)})$.

Two of the next axioms rely on the following notation. For a data game $C_{(D,O)} \in \mathcal{C}(N)$ and an agent $i \in N$, we denote by

$$b^{\max}_i(C_{(D,O)}) = \max_{S\in 2^N : i\in S} \Big(C_{(D,O)}(S) - C_{(D,O)}(S\backslash\{i\}) \Big)$$

and

$$b_i^{\min}(C_{(D,O)}) = \min_{S\in 2^N : i\in S} \Big(C_{(D,O)}(S) - C_{(D,O)}(S\backslash\{i\}) \Big)$$

agent i's maximal (or worst) and minimal (or best) marginal contributions to the coalitions he belongs to, respectively.

Reasonableness. $\forall C_{(D,O)} \in \mathcal{C}(N)$, $\forall i \in N$, $b_i^{\min}(C_{(D,O)}) \leq f_i(C_{(D,O)}) \leq b_i^{\max}(C_{(D,O)})$.

Reasonableness prevents an agent to pay more than the total cost of the data he needs, and to be paid more than the total cost of his data.

Equal concessions. $\forall C_{(D,O)} \in \mathcal{C}(N)$ such that $D^E(O) = D$, $\forall i, j \in N$, $f_i(C_{(D,O)}) - b_i^{\min}(C_{(D,O)}) = f_j(C_{(D,O)}) - b_j^{\min}(C_{(D,O)})$.

Invariance to deleting non-exclusive data. $\forall C_{(D,O)} \in \mathcal{C}(N)$ such that $|D| \geq 2$, $\forall k \in D\backslash D^E(O)$, $f(C_{(D,O)}) = f(C_{(D\backslash\{k\},O)})$.

So far, all the axioms have been stated for data games associated with the same ownership function. The last axiom that we invoke relies on interrelated ownership functions.

Invariance to enlarging the owner set. $\forall C_{(D,O)}, C_{(D,O')} \in \mathcal{C}(N)$ such that $O_h = O'_h$ for all $h \in D\backslash\{k\}$ and $O_k \subset O'_k$, and all $i \in N\backslash O'_k$, $f_i(C_{(D,O)}) = f_i(C_{(D,O')})$.

This formal definition allows a more detailed description of the axiom than the one given in section 2. Invariance to enlarging the owner set considers two data games based on data sharing problems that only differ with respect to the owner set of data k. More specifically, the owner set in the second problem in a strict superset of the owner set in the first problem. Invariance to enlarging the owner set requires that this enlargement of the owner set has no influence on the compensation paid or received by an agent who does not hold the corresponding data in both problems. In a sense, the axioms says that the price necessary to acquire such data should not depend on how many agents initially own it.

Compensation, Pooling, Equal treatment of equals and Invariance to deleting non-exclusive data are already invoked in Béal *et al.* [4], Reasonableness is due to Milnor [21], and Equal concessions and Invariance to enlarging the owner set are newly introduced. We now have the material to study the consequences of various combinations of the above-defined axioms. As a first result, we show that if a compensation scheme satisfies Pooling, then it is a Core-selector if and only if it also satisfies Compensation, Reasonableness and Invariance to deleting non-exclusive data.

Proposition 2 *Consider any compensation scheme f on $\mathcal{C}(N)$ satisfying Pooling. Then f is a Core-selector if and only it further satisfies Compensation, Reasonableness and Invariance to deleting non-exclusive data.*

Proof. Let f be any compensation scheme on $\mathcal{C}(N)$ satisfying Pooling. The proof is split in two steps. Firstly, we show that if f is a Core-selector, then it satisfies Compensation, Reasonableness and Invariance to deleting non-exclusive data. By definition, the sum of all compensations on any element of the Core is equal to zero so that f trivially satisfies Compensation. It is known that the Core satisfies Reasonableness (stated for all compensations vector of a compensation solution) on the class of all games in coalitional form, which implies that any of its element also satisfies Reasonableness on $\mathcal{C}(N)$, and in turn that f satisfies Reasonableness on $\mathcal{C}(N)$. Finally, from Proposition 1, for each $k \in D$ such that $o_k \geq 2$, it holds that $\text{Core}(C_{(\{k\},O)}) = \{\mathbf{0}\}$. Thus, since f is a Core-selector, it must be that $f(C_{(\{k\},O)}) = \mathbf{0}$ for such a data game. By Pooling, for each $C_{(D,O)} \in \mathcal{C}(N)$ and each $k \in D$

such that $o_k \geq 2$, we get $f(C_{(D,O)}) = f(C_{(D\setminus\{k\},O)}) + f(C_{(\{k\},O)}) = f(C_{(D\setminus\{k\},O)}) + \mathbf{0} = f(C_{(D\setminus\{k\},O)})$, proving that f satisfies Invariance to deleting non-exclusive data.

Secondly, we show if a compensation scheme f on $\mathcal{C}(N)$ satisfies Compensation, Reasonableness and Invariance to deleting non-exclusive data (in addition to Pooling), then it is a Core-selector. Pick any $C_{(D,O)} \in \mathcal{C}(N)$. By Pooling and the expression of the Core given in Proposition 1, it is enough to show that $f(C_{(\{k\},O)}) \in \text{Core}(C_{(\{k\},O)})$ for each $k \in D$. We consider two cases. As a first case, suppose that $o_k \geq 2$. Consider any other data h and the data game $C_{(\{h,k\},O')}$ such that $O'_k = O_k$. On the one hand, by pooling, we get $f(C_{(\{h,k\},O')}) = f(C_{(\{h\},O')}) + f(C_{(\{k\},O')}) = f(C_{(\{h\},O')}) + f(C_{(\{k\},O)})$. On the other hand, applying Invariance to deleting non-exclusive data to k yields that $f(C_{(\{h,k\},O')}) = f(C_{(\{h\},O')})$. Combining these two observations implies that $f(C_{(\{k\},O)}) = \mathbf{0}$, and it turns out that $\mathbf{0}$ is the unique Core element according to Proposition 1. As a second case, suppose that $o_k = 1$, and denote by $i \in N$ the unique owner of data k. Choose any $j \in N\setminus\{i\}$, which is possible since $n \geq 2$. It is easy to see that $b_i^{\min}(\{k\}, O) = C_{(\{k\},O)}(N) - C_{(\{k\},O)}(N\setminus\{i\}) = -c_k$ and that $b_i^{\max}(\{k\}, O) = C_{(\{k\},O)}(\{i\}) - C_{(\{k\},O)}(\emptyset) = 0$. By Reasonableness, we obtain $f_i(C_{(\{k\},O)}) \in [-c_k, 0]$. Similarly, observe that $b_j^{\min}(\{k\}, O) = C_{(\{k\},O)}(N) - C_{(\{k\},O)}(N\setminus\{j\}) = 0$ and that $b_j^{\max}(\{k\}, O) = C_{(\{k\},O)}(\{j\}) - C_{(\{k\},O)}(\emptyset) = c_k$. Another application of Reasonableness yields that $f_j(C_{(\{k\},O)}) \in [0, c_k]$ for each $j \in N\setminus\{i\}$. Together with Compensation, it is easy to conclude that $f(C_{(\{k\},O)}) \in \text{conv}(\{z^j(k)\}_{j\in N})$ where $z^i(k) = (0, \ldots, 0)$, and $z_i^j(k) = -c_k$, $z_j^j(k) = c_k$ and $z_r^j(k) = 0$ for each $r \in N\setminus\{i,j\}$ as in Proposition 1. As a consequence, $f(C_{(\{k\},O)}) \in \text{Core}(C_{(\{k\},O)})$, as desired. ■

Replacing Reasonableness in Proposition 2 by Equal concessions singles out the Nucleolus among all the Core-selectors.

Proposition 3 *The Nucleolus is the unique compensation scheme on $\mathcal{C}(N)$ that satisfies Compensation, Pooling, Invariance to deleting non-exclusive data and Equal concessions.*

Proof. As a Core-selector, the Nucleolus satisfies Compensation and Invariance to deleting non-exclusive data. It also satisfies Pooling by its expression in Proposition 1. It remains to show that it satisfies equal concessions. So pick any $C_{(D,O)}$ such that $D^E(O) = D$. By Proposition 1, this means that

$$\eta_i(C_{(D,O)}) = \sum_{k\in D} \frac{c_k}{n} - \sum_{k\in D_i(O)} c_k, \quad i \in N. \tag{1}$$

Since D only contains data held by a unique agent, we have, for each $i \in N$, that

$$b_i^{\min}(C_{(D,O)}) = C_{(D,O)}(N) - C_{(D,O)}(N\setminus\{i\}) = -\sum_{k\in D_i(O)} c_k. \tag{2}$$

Subtracting (2) to (1) yields

$$\sum_{k\in D} \frac{c_k}{n}.$$

Since this expression does not depend on agent $i \in N$, we get $\eta_i(C_{(D,O)}) - b_i^{\min}(C_{(D,O)}) = \eta_j(C_{(D,O)}) - b_j^{\min}(C_{(D,O)})$ for all $i, j \in N$, as desired.

Next, consider any compensation scheme f on $\mathcal{C}(N)$ that satisfies Compensation, Pooling, Invariance to deleting non-exclusive data and Equal concessions. Pick any $C_{(D,O)} \in \mathcal{C}(N)$, and any $k \in D$. From the proof of Proposition 2, we already know that Pooling and Invariance to deleting non-exclusive data imply that $f(C_{(\{k\},O)}) = \mathbf{0} = \eta(C_{(\{k\},O)})$ if $o_k \geq 2$. So suppose that $o_k = 1$ and denote by $i \in N$ the unique owner of data k. In the elementary data game $(\{k\}, O)$, recall from the proof of Proposition 2 that $b_i^{\min}(\{k\}, O) = -c_k$ and that $b_j^{\min}(\{k\}, O) = 0$. Thus, Equal concessions rewrites

$$f_i(\{k\}, O) + c_k = f_j(\{k\}, O) \tag{3}$$

for each $j \in N \backslash \{i\}$. Summing on all $j \in N$ and using Compensation, we obtain

$$0 = f((\{k\}, O))(N) = f_i(\{k\}, O) + (n-1)\big(f_i(\{k\}, O) + c_k\big)$$

so that $f_i(\{k\}, O) = c_k/n - c_k = \eta_i(\{k\}, O)$. By (3), conclude that $f_j(\{k\}, O) = c_k/n = \eta_j(\{k\}, O)$ for each $j \in N \backslash \{i\}$ as well. By Pooling, $f(C_{(D,O)}) = \eta(C_{(D,O)})$, which completes the proof. ■

Replacing Invariance to deleting non-exclusive data in Proposition 3 by Invariance to enlarging the owner set and adding Equal treatment of equals yield a characterization of the Shapley value. It is known that the Nucleolus also satisfies Equal treatment of equals. As such Propositions 3 and 4 provide comparable characterizations of the Nucleolus and the Shapley value that only differ with respect to one axiom.

Proposition 4 *The Shapley value is the unique compensation scheme on $\mathcal{C}(N)$ that satisfies Compensation, Pooling, Equal treatment of equals, Invariance to enlarging the owner set and Equal concessions.*

Proof. It is obvious that the Shapley value satisfies Compensation, Pooling and Equal treatment of equals. To see that it also satisfies Invariance to enlarging the owner set, consider any pair of data games $C_{(D,O)}$ and $C_{(D,O')} \in \mathcal{C}(N)$ such that $O_h = O'_h$ for all $h \in D \backslash \{k\}$ and $O_k \subset O'_k \subset N$ for some $k \in D$. Recall that $O_k \neq \emptyset$ by assumption. Choose any $i \in N \backslash O'_k$. Observe that $D_i(O') = D_i(O)$ and $O_h = O'_h$ for each $h \in D_i(O)$. As a consequence, the expression of the Shapley value given in Proposition 1 immediately yields

$$Sh_i(C_{(D,O')}) = \sum_{h \in D} \frac{c_h}{n} - \sum_{h \in D_i(O')} \frac{c_h}{o'_h} = \sum_{h \in D} \frac{c_h}{n} - \sum_{h \in D_i(O)} \frac{c_h}{o_h} = Sh_i(C_{(D,O)}).$$

Since the Nucleolus satisfies Equal concessions and $\eta(C_{(D,O)}) = Sh(C_{(D,O)})$ whenever $D^E(O) = D$, the Shapley value satisfies Equal concessions as well.

Next, consider any compensation scheme f on $\mathcal{C}(N)$ that satisfies Compensation, Pooling, Equal treatment of equals, Invariance to enlarging the owner set and Equal concessions. Pick any $C_{(D,O)} \in \mathcal{C}(N)$, and any $k \in D$. Suppose that $o_k = 1$ and once again denote by $i \in N$ the unique owner of data k. In the elementary data game $(\{k\}, O)$, from the proof of Proposition 3 and the fact that $\{k\}^E(O) = \{k\}$, we know that the combination of Compensation and Equal concessions imply that $f(C_{(\{k\},O)}) = \eta(C_{(\{k\},O)}) = Sh(C_{(\{k\},O)})$. In particular, for each $j \in N \backslash \{i\}$, we have $f_j(C_{(\{k\},O)}) = c_k/n$. Now, consider any $k \in D$ such that $o_k \geq 2$. If $o_k = n$, then Equal treatment of equals and Compensation yield $f(C_{(\{k\},O)}) = \mathbf{0}$. If $2 \leq o_k < n$, we can choose an agent $i \in O_k$ and $j \in N \backslash O_k$. Then, construct the elementary data game $(\{k\}, O')$ such that $O'_k = \{i\}$. Invariance to enlarging the owner set and Equal concessions can be applied to data games $(\{k\}, O)$ and $(\{k\}, O')$, and to agent j because $j \in N \backslash O'_k$. Thus, since $o'_k = 1$, we obtain from the previous part of the proof that $f_j(C_{(\{k\},O)}) = f_j(C_{(\{k\},O')}) = c_k/n$. Since j was arbitrary chosen in $N \backslash O_k$, this equality holds for all agents in $N \backslash O_k$. By Compensation, we get

$$f(C_{(\{k\},O)})(O_k) = -(n - o_k)\frac{c_k}{n}.$$

By Equal treatment of equals,

$$f_i(C_{(\{k\},O)}) = \frac{1}{o_k} \times -(n - o_k)\frac{c_k}{n} = \frac{c_k}{n} - \frac{c_k}{o_k} = Sh_i(C_{(\{k\},O)}).$$

Conclude by Pooling that $f(C_{(D,O)}) = Sh(C_{(D,O)})$ ■

5 Conclusion

The absence of available data on the types of compensation schemes that are actually used within the SIEF prevents us from concluding this article by a confrontation of our theory with the current practice. This important step is left for future research. Instead, we offer two concluding remarks by coming back to the properties mobilized throughout the article.

The set of axioms invoked in Proposition 3 and 4 are almost identical, in particular because the Nucleolus satisfies Equal treatment of equals even if this axiom is not needed in Proposition 3. Therefore, it makes sense to ask whether the separating axioms of Invariance to deleting non exclusive data and Invariance to enlarging the owner set are compatible. As a partial answer, it is possible to show that the price to have simultaneously both axioms is to rule out any form of compensation if the basic axioms of Compensation, Pooling and Equal treatment of equals are also imposed. In other words, it is possible to show that the null compensation scheme that assigns the null compensation vector $\mathbf{0}$ to any data game is the unique compensation scheme satisfying Invariance to deleting non exclusive data, Invariance to enlarging the owner set are compatible, Compensation, Pooling and Equal treatment of equals.

Such comparisons between axioms are a crucial facet of the axiomatic method. Applied to the decision problem faced by a judge or and arbitrator, this amounts to examine whether philosophical principles are not contradictory, or if a philosophical principle has stronger implications than another one. In order to have a clear overview of our results, we conclude this article by the following recap chart in which a "+" means that the compensation scheme satisfies the axiom, in which "−" has the converse meaning and in which the "⊕" symbols indicate the characterizing sets of axioms in Propositions 2, 3 and 4 respectively.[5]

	Core-selectors	Nucleolus	Shapley value
Compensation	⊕	⊕	⊕
Pooling	⊕	⊕	⊕
Equal treatment of equals	−	+	⊕
Reasonableness	⊕	+	⊕
Equal concessions	−	⊕	⊕
Invariance to deleting non exclusive data	⊕	⊕	−
Invariance to enlarging the owner set	−	−	⊕

Appendix

In order to demonstrate independence of the axioms imposed in Propositions 2, 3 and 4, we exhibit the following compensation schemes.

Proposition 2

- The Shapley value Sh on $\mathcal{C}(N)$ satisfies Pooling, Compensation, Reasonableness but not Invariance to deleting non exclusive data.
- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by

$$f_1(C_{(D,O)}) = \sum_{k \in D^E(O)} c_k \quad \text{and} \quad f_i(C_{(D,O)}) = -\frac{1}{n-1} \sum_{k \in D^E(O)} c_k, \quad i \in N \backslash \{1\},$$

[5] Recall that Proposition 2 only consider core-selectors satisfying the Pooling axiom. As such, Pooling is rather an assumption than a requirement in Proposition 2.

if $D^E(O) \neq \emptyset$ and $f(C_{(D,O)}) = \mathbf{0}$ if $D^E(O) = \emptyset$ satisfies Pooling, Compensation and Invariance to deleting non exclusive data but not Reasonableness.

- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by

$$f_i(C_{(D,O)}) = \eta_i(C_{(D,O)}) + \sum_{k \in D^E(O)} \frac{c_k}{n}, \quad i \in N,$$

satisfies Pooling, Reasonableness and Invariance to deleting non exclusive data but not Compensation.

Proposition 3

- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by $f(C_{(D,O)}) = \mathbf{0}$ satisfies Compensation, Pooling and Invariance to deleting non exclusive data but not Equal concessions.
- The Shapley value Sh satisfies Compensation, Pooling and Equal concessions but not Invariance to deleting non exclusive data.
- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by $f(C_{(D,O)}) = \eta(C_{(D,O)})$ if $D^E(O) \neq \emptyset$ and $f(C_{(D,O)}) = Sh(C_{(D,O)})$ if $D^E(O) = \emptyset$ satisfies Compensation, Equal concessions and Invariance to deleting non exclusive data but not Pooling.
- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by

$$f_i(C_{(D,O)}) = \eta_i(C_{(D,O)}) + |D^E(O)|, \quad i \in N,$$

satisfies Pooling, Equal concessions and Invariance to deleting non exclusive data but not Compensation.

Proposition 4

- The Nucleolus η satisfies Compensation, Pooling, Equal treatment of equals and Equal concessions but not Invariance to enlarging the owner set.
- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by $f(C_{(D,O)}) = \mathbf{0}$ satisfies Compensation, Pooling, Equal treatment of equals and Invariance to enlarging the owner set but not Equal concessions.
- The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by

$$f_i(C_{(D,O)}) = Sh_i(C_{(D,O)}) + |D|, \quad i \in N,$$

satisfies Pooling, Equal treatment of equals, Equal concessions, Invariance to enlarging the owner set but not Compensation.

- Consider any compensation vector $a \in \mathbb{R}^n \backslash \{\mathbf{0}\}$. The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by

$$f_i(C_{(D,O)}) = \sum_{k \in D: o_k < n} \frac{c_k}{n} - \sum_{k \in D_i(O): o_k < n} \frac{c_k}{o_k} + a_i \times |\{k \in D : o_k = n\}|, \quad i \in N,$$

satisfies Pooling, Compensation, Equal concessions and Invariance to enlarging the owner set but not Equal treatment of equals.

- Consider the data game $C_{(\{h,k\},O')} \in \mathcal{C}(N)$ such that $O'_h = \{1,3\}$ and $O'_k = \{2,3\}$. The compensation scheme f on $\mathcal{C}(N)$ defined, for all $C_{(D,O)} \in \mathcal{C}(N)$ by $f(C_{(D,O)}) = Sh(C_{(D,O)})$ if $C_{(D,O)} \neq C_{(\{h,k\},O')}$ and such that $f_i(C_{(\{h,k\},O')}) = Sh_i(C_{(\{h,k\},O')})$ for each $i \in N\backslash\{1,2,3\}$, $f_1(C_{(\{h,k\},O')}) = Sh_1(C_{(\{h,k\},O')}) + c_h/2$, $f_2(C_{(\{h,k\},O')}) = Sh_2(C_{(\{h,k\},O')}) + c_k/2$ and $f_3(C_{(\{h,k\},O')}) = Sh_3(C_{(\{h,k\},O')}) - c_h/2 - c_k/2$ satisfies Compensation, Equal treatment of equals, Equal concessions and Invariance to enlarging the owner set but not Pooling.

References

[1] F. Ackerman, E. A. Stanton and R. Massey – "European chemical policy and the United States: The impacts of REACH", *Renewable Resources Journal* **25** (2007), p. 15–20.

[2] F. Ackerman, E. A. Stanton, B. Roach and A.-S. Andersson – "Implications of REACH for developing countries", *European Environment* **18** (2008), p. 16–29.

[3] E. Algaba, J. M. Bilbao and J. R. Fernández – "The distribution of power in the European Constitution", *European Journal of Operational Research* **176** (2007), p. 1752–1766.

[4] S. Béal, M. Deschamps, J.-T. Ravix and O. Sautel – "Les informations exigées par la législation reach : Analyse du partage des coûts", *Revue d'Economie Politique* **120** (2010), p. 109–132.

[5] L. Bergkamp (éd.) – *The european union REACH regulation for chemicals: Law and practice*, Oxford University Press, 2013.

[6] U. Bindseil and C. Hantke – "The power distribution in decision making among EU member states", *European Journal of Political Economy* **13** (1997), p. 71–85.

[7] M. Braham and F. Steffen – "Voting rules in insolvency law: a simple-game theoretic approach", *International Review of Law and Economics* **22** (2003), p. 421–442.

[8] J. Canton and C. Allen – "A microeconomic model to assess the economic impacts of the EU's new chemicals policy", Brussels: European Commission, DG Enterprise, 2003.

[9] A. Casajus, H. Labrenz and T. Hiller – "Majority shareholder protection by variable qualified majority rules", *European Journal of Law and Economics* **28** (2009), p. 9–18.

[10] P. Dehez and D. Tellone – "Data games: Sharing public goods with exclusion", *Journal of Public Economic Theory* **15** (2013), p. 654–673.

[11] R. Dworkin – "What is equality? part 2: Equality of resources", *Philosophy & Public Affairs* **10** (1981), p. 283–345.

[12] S. Fedeli and F. Forte – "Voting powers and the efficiency of the decision-making process in the european council of ministers", *European Journal of Law and Economics* **12** (2001), p. 5–38.

[13] M. Fleurbaey – *Fairness, responsibility, and welfare*, Oxford University Press, 2008.

[14] M. Fleurbaey and F. Maniquet – *Equality of opportunity: The economics of responsibility*, World Scientific Series in Economic Theory: Volume 1, 2012.

[15] M. Fleurbaey and J. E. Roemer – "Judicial precedent as a dynamic rationale for axiomatic bargaining theory", *Theoretical Economics* **6** (2011), p. 289–310.

[16] D. B. Gillies – "Some theorems on n-person games", *Ph.D. Dissertation*, Princeton University, Department of Mathematics, 1953.

[17] L. A. Kornhauser – "Modeling collegial courts I: Path-dependence", *International Review of Law and Economics* **12** (1992), p. 169–185.

[18] D. Leech and M. C. Manjón – "Corporate governance in Spain (with an application of the power indices approach)", *European Journal of Law and Economics* **13** (2002), p. 157–173.

[19] A. Logomasini – "Strategic approach to international chemicals management: Lack of interest belies importance", *Sustainable development law & policy* **6** (2006), p. 33–37.

[20] T. Lorenz, B. Lebreton and L. N. van Wassenhove – "The REACH directive and its impact on the european chemical industry: a critical review", INSEAD Working Paper No. 53, 2008.

[21] J. W. Milnor – "Reasonable outcomes for n-person games", Research memorandum 916, The Rand Corporation, Santa Monica, CA., 1952.

[22] J. Rawls – *A theory of justice*, Belknap Press of Harvard University Press, 1971.

[23] D. Schmeidler – "The Nucleolus of a characteristic function game", *SIAM Journal of Applied Mathematics* **17** (1969), p. 1163–1170.

[24] L. S. Shapley – "A value for n-person games", in *Contribution to the Theory of Games vol. II (H.W. Kuhn and A.W. Tucker eds). Annals of Mathematics Studies **28***, Princeton University Press, Princeton, 1953.

[25] P. T. Spiller – "Rationality, decision rules, and collegial courts", *International Review of Law and Economics* **12** (1992), p. 1986–190.

[26] W. Thomson – "On the axiomatic method and its recent applications to game theory and resource allocation", *Social Choice and Welfare* **18** (2001), p. 327–387.

[27] H. Von Holleben, A. Rosenfeld and H. Scheidmann – *REACH consortia model agreements: And other forms of data sharing according to REACH*, Lexxion, Berlin, 2008.

[28] O. Wolf and L. Delgado – "The impact of REACH on innovation in the chemical industry", European Commission, joint research center, Technical report series, EUR 20999 EN.

[29] P. H. Young – *Equity in theory and practice*, Princeton University Press, 1994.