# A comment on P. Dehez and D. Tellone *Data games: sharing public goods with exclusion*

(published in JPET 15(2013): 654–673)

Theo S.H. Driessen* Anna Khmelnitskaya†

September 20, 2013

**Abstract**

We show that the mathematical model of the data cost game introduced in Dehez and Tellone (*JPET*, 2013) coincides with the model of the library cost game studied in Driessen, Khmelnitskaya and Sales (*TOP*, 2012) where its core, nucleolus and Shapley value were also investigated.

**Keywords:** cooperative TU game, data game, library game, core, nucleolus, Shapley value, 1-concavity

**JEL Classification Number:** C71, H41

The discussed in Dehez and Tellone [1] data sharing problem faced nowadays by the EU chemical industry definitely strongly motivates a particular interest for study of data sharing problems. In general the data sharing situation involves a finite group of agents (firms), a set of all data, data sets owned by individual agents, and a vector of costs for reproducing each data. Each agent owns a subset of data. No a priori restrictions are imposed on the individual data sets. In particular, some data may be owned simultaneously by several agents, whereas some agents may hold no data or hold the complete data set. The only assumption, which in fact can be avoided, is that individual data sets altogether cover the all data that are needed. The main question is to determine how to compensate the agents for the data they contribute to the common pull. In Dehez and Tellone [1] it is shown that this compensation problem can be framed within a transferable utility cost sharing game, the so-called data game, defined by the cost function that specifies for each coalition the cost of acquiring the data it misses, so no cost is charged to the whole group of agents. To obtain a solution of a data game standard cost allocation rules, such as the Shapley value, the nucleolus, the core and so on, can be applied. In the paper it is shown that the core of a data game is always nonempty and has a regular simplex structure, the nucleolus coincides with the barycenter of the core, and simple expressions for computation of the Shapley value and nucleolus are introduced. We

*T.S.H. Driessen, University of Twente, Department of Applied Mathematics, P.O. Box 217, 7500 AE Enschede, The Netherlands, e-mail: t.s.h.driessen@ewi.utwente.nl

†A.B. Khmelnitskaya, Saint-Petersburg State University, Faculty of Applied Mathematics, Universitetskii prospekt 35, 198504, Petergof, Saint-Petersburg, Russia, e-mail: a.khmelnitskaya@math.utwente.nl

would like to emphasize that the modeling of the data sharing problem in terms of the cooperative data game is definitely worthwhile since it provides an important practical application of cooperative game theory. However, the mathematical model of the data game and the results concerning its core, Shapley value and nucleolus are not new and they are already discussed in the literature. In fact the data game is a particular case of another cost game, the so-called library game, introduced and studied in Driessen, Khmelnitskaya, and Sales [3]. To show that recall briefly the definitions of both games using compatible notation.

A library game that models the situation in which a university library consortium has to fix the tariff that each member should pay in charge for the joint subscription for electronic scientific journals, is defined as follows. Let $N = \{1, \ldots, n\}$ be a finite set of $n$ players (universities). $G = \{1, \ldots, m\}$ is a set of all liable goods (electronic journals) to be chosen. $D = (d_{ij})_{\substack{i \in N \\ j \in G}}$ is a demand $(n \times m)$-matrix where $d_{ij} \geq 0$ presents the number of units of $j$th journal in the historical demand of $i$th university. Let $c_j \geq 0$ be the cost per unit of $j$th journal based on the price of the paper version in the historical demand, whereas $\alpha \in [0,1]$ is the common percentage for goods that were never requested in the past; in applications usually $\alpha = 10\%$. The characteristic function $c^l$ of the library cost game on the player set $N$ is given by

$$c^l(S) = \sum_{j \in G} \Big[ \sum_{i \in S} d_{ij} \Big] c_j + \sum_{\substack{j \in G \\ \sum_{i \in S} d_{ij} = 0}} \alpha\, c_j, \qquad \text{for all } S \subseteq N.$$

The library game $c^l$ is in fact a sum of an additive game and multiplied by $\alpha$ nonadditive game $\bar{c}^l$ given by

$$\bar{c}^l(S) = \sum_{\substack{j \in G \\ \sum_{i \in S} d_{ij} = 0}} c_j, \qquad \text{for all } S \subseteq N.$$

A data sharing situation is determined by a finite set $N = \{1, \ldots, n\}$ of $n$ players, a set of all data (public goods) $G = \{1, \ldots, m\}$, a collection of sets $G_i \subseteq G$, $i \in N$, that specify the data held by each player, and a vector of costs $c_j$ of reproducing the data $j \in G$. It is also assumed that $\cup_{i \in N} G_i = G$. Then the characteristic function $c^d$ of the data cost game on the player set $N$ is given by

$$c^d(S) = \sum_{j \in G \setminus G_S} c_j, \qquad \text{for all } S \subseteq N.$$

where $G_S = \cup_{i \in S} G_i$ for all $S \subseteq N$.

**Proposition 1** *Assume that in both games, a data game $c^d$ and a library game $c^l$, the sets of players $N$ and of goods $G$[1] and the vectors of costs $c_j$, $j \in G$, are the same. Moreover, assume that in the library game $c^l$ the demand matrix $D = (d_{ij})_{\substack{i \in N \\ j \in G}}$ relates to individual data sets $G_i \subseteq G$, $i \in N$, in the data game $c^d$ as: $d_{ij} = 0$ iff $G_i \not\ni j$. Then the data game $c^d$ coincides with the nonadditive component $\bar{c}^l$ of the library game $c^l$.*

[1] In case of a library situation the set of goods $G$ is interpreted as the set of all journals, while in case of a data situation $G$ is the set of all data.

*Proof.* To prove the coincidence of both games $c^d$ and $\bar{c}^l$, it suffices to show that for any coalition $S \subseteq N$, the set $\{j \in G \mid \sum_{i \in S} d_{ij} = 0\}$ in the library situation coincides with the set $G \setminus G_S$ in the data situation. Clearly, for any $j \in G$, the coalitional constraint $\sum_{i \in S} d_{ij} = 0$ is equivalent to individual constraints $d_{ij} = 0$ for all $i \in S$, which by hypothesis concerning $d_{ij}$ mean that $j \notin G_i$ for all $i \in S$, which is the same as $j \notin G_S$, i.e., $j \in G \setminus G_S$. ∎

In Driessen et al. [3] it is proved that a library game, and in particular its nonadditive component, belongs to the class of 1-concave games introduced and studied in Driessen and Tijs [4] and Driessen [2], and as a corollary to that and known properties of the solutions of 1-concave games it is shown that the core of a library game is always nonempty and has a regular simplex structure, the nucleolus is linear and coincides with the $\tau$-value and the barycenter of the core. Due to Proposition 1 these results extend straightforwardly to data games as well, which covers the results of sections 4 and 5 in Dehez and Tellone [1]. Remark also that the expressions for the nucleolus and Shapley value of the data game given in Dehez and Tellone [1] by formulas (10) and (12) can be found in adapted notation in Driessen et al. [3] on page 590.

## References


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

[2] Driessen, T.S.H. (1985), Properties of 1-convex $n$-person games, *OR Spektrum, 7*, 19–26.

[3] Driessen, T.S.H., A.B. Khmelnitskaya, and J. Sales (2012), 1-concave basis for TU games and the library game, *TOP, 20*, 578–591, published online 01 September 2010; the first version of the paper was published online in 2005 as Memorandum 1777, Department of Applied Mathematics, University of Twente, The Netherlands.

[4] Driessen, T.S.H., S.H. Tijs (1983), The $\tau$-value, the nucleolus and the core for a subclass of games, *Methods of Operations Research, 46*, 395–406.