Games and Economic Behavior 111 (2018) 41–63

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

A non-parametric approach to testing the axioms of the Shapley value with limited data ☆

Victor H. Aguiar [a,*], Roland Pongou [b], Jean-Baptiste Tondji [c]

[a] Department of Economics, University of Western Ontario, Canada
[b] Department of Economics, University of Ottawa, Canada
[c] Department of Economics and Finance, University of Texas Rio Grande Valley, United States

ARTICLE INFO

Article history:
Received 28 April 2017
Available online 18 June 2018

JEL classification:
C71
C78
D20
D30
J30

Keywords:
Shapley value
Fairness violations
Limited data
Inequality
Shapley distance

ABSTRACT

The unique properties of the Shapley value–efficiency, equal treatment of identical input factors, and marginality–have made it an appealing solution concept in various classes of problems. It is however recognized that the pay schemes utilized in many real-life situations generally depart from this value. We propose a non-parametric approach to testing the empirical content of this concept with limited datasets. We introduce the *Shapley distance*, which, for a fixed monotone transferable-utility game, measures the distance of an arbitrary pay profile to the Shapley pay profile, and show that it is additively decomposable into the violations of the classical Shapley axioms. The analysis has several applications. In particular, it can be used to assess the extent to which an income distribution can be considered fair or unfair, and whether any particular case of unfairness is due to the violation of one or a combination of the Shapley axioms.

1. Introduction

In an environment in which output is produced through the combination of several inputs, Shapley (1953) provides an axiomatic solution to the problem of valuing the contribution of each input. The unique properties of the Shapley value–efficiency, equal treatment of identical input factors (symmetry), and marginality–have made it an appealing solution concept in various classes of problems including wage determination, cost allocation, centrality measurement in networks, quantification of the importance of a product's attributes, and causal assessment in a statistical context. Yet, despite the acknowledged theoretical appeal of this concept, it is recognized that the pay schemes utilized in most real-life environments depart from it. In this paper, we provide a way to measure such departures in limited datasets. We introduce the *Shapley distance*, which, for a fixed monotone transferable-utility game (or production function), measures the distance of an arbitrary pay profile to the Shapley pay profile, and show that it is additively decomposable into the violations of its classical axioms.

☆ We thank an associate editor and two anonymous referees for very constructive suggestions that have improved the paper. We are also grateful to Geoffroy de Clippel, Jim Davies, Greg Pavlov, Rajiv Vohra, Kareen Rozen, Roberto Serrano, and the participants of the Brown Theory Lunch for useful comments and encouragement.

* Corresponding author.
E-mail addresses: vaguiar@uwo.ca (V. Aguiar), roland.pongou@uottawa.ca (R. Pongou), tondji.jean@gmail.com (J.-B. Tondji).

https://doi.org/10.1016/j.geb.2018.06.003
0899-8256/

The theoretical analysis that we propose in this paper is important for at least three reasons, as explained hereunder:

1. To the extent that the axioms characterizing the Shapley value make it a desirable concept of fairness (or distributive justice), as is generally acknowledged in the literature (Yaari, 1981; Roth, 1988), our Shapley distance is a measure of unfairness. It can be used, for instance, to determine the extent to which a given income distribution under a known production technology is unfair. Furthermore, our decomposition of this distance determines whether unfairness, if it is at all present, is due to a violation of horizontal equality (i.e., equal pay for equal work), to a lack of fair compensation for marginal efforts, or simply to output waste. In this sense, from a methodological and axiomatic point of view, our analysis can be regarded as contributing to the theory of distributive justice (Konow, 2003), and it can be applied to inform the recurrent debate around the fairness or unfairness of income inequality in most modern societies.
2. If we consider a laboratory experiment in which subjects give their opinions on how the output of a collaborative project should be shared among the different contributors, we might be interested in whether the average opinion is consistent with the Shapley value, and we might quantify the source of any discrepancy. Such an analysis might shed light on which axioms of the Shapley value are less robust from an empirical point of view. Indeed, we can think of the dataset in de Clippel and Rozen (2013) where we envision practitioners applying our methodology. We also hope that the small but growing literature on testing solution concepts in transferable-utility games can benefit from our methodological contribution. Important contributions in this area are Kalisch et al. (1954), Bolton et al. (2003), and Nash et al. (2012).
3. Our methodology also provides the first non-parametric test of the Shapley axioms in a limited dataset (which is generically the type of data that most practitioners have access to). In addition, our analysis is, to our knowledge, the first non-parametric approach to measuring and decomposing departures of any observed pay profile from the predictions of the Shapley axioms.

Two fundamental axiomatic characterizations of the Shapley value that have received wide attention in the literature guide our decomposition analysis. The first characterization, due to Young (1985), states that the Shapley value is the only pay scheme that satisfies efficiency and marginality, and that treats perfectly substitutable players or production factors identically.[1] The other characterization, which derives from the original work of Shapley (1953), states that the Shapley value is the only pay scheme that satisfies the axioms of efficiency, null player, and additivity, and that does not discriminate between identical players.[2] As is explained below, these two characterizations also provide a basis for studying the formal relationships that exist among the different violations of these appealing axioms by means of an arbitrary pay profile.

Our main contribution is to compare any arbitrary pay profile (i.e., a vector of payoffs) to the Shapley pay profile for a fixed game, when the observer has only limited data on the production environment. Given a monotone game and an arbitrary pay profile, we define the Shapley distance of this pay profile as its (Euclidean) distance to the Shapley pay profile of this game. We provide a unique *orthogonal* decomposition of this distance into nonnegative terms that measure violations of the above mentioned classical axioms of the Shapley value (Theorem 1). Importantly, we assume that an observed pay profile is generated by a pay scheme that may be unobserved (a pay scheme is a rule that maps any game into a pay profile). However, the decomposition of the Shapley distance of a pay profile can be used to make inference about the extent to which the possibly unobserved pay scheme generating this pay profile violates the axioms of the Shapley value. It is particularly interesting that this exercise shows how a violation of the marginality axiom is formally related to violations of the null player and additivity axioms, thus further highlighting the correlation or the dependence between these axioms (Theorem 3).

A clear advantage of our framework is that it makes it possible to carry out the proposed tests with only limited data, a limited dataset being a finite sequence of observations, with each observation being a pair of a game and a pay profile. In particular, we can test for departures from the main Shapley axioms using only one observation. For instance, despite the fact that the axioms of marginality and additivity are stated using two transferable-utility games, which should prevent a test of their violation if one observes only one game and one pay profile, our framework allows us to carry out such a test for the class of monotone games and nonnegative and feasible pay schemes. We partially overcome this difficulty on the class of *monotone* games by showing that under our non-negativity and feasibility assumptions on the pay scheme, marginality implies the null player property.

More generally, any nonnegative and feasible pay scheme satisfying the marginality axiom on the class of monotone games should provide a player with a payoff at least as large as the payoff obtained by a null player. This property provides a lower bound on the players' payoff, which is essential in testing the marginality axiom. Likewise, the condition that the total distributed payoff is not less than the worth of the grand coalition, combined with nonnegativity, provides an upper

[1] *Efficiency* means that the entire output of a collaborative effort is shared among the contributors, implying that no portion of it is wasted. The *marginality* axiom, due to Young (1985), states that a player should be valued more under a production technology that values his input more. This axiom is related to the *null player* and the *additivity* axioms. The null player axiom states that if a player's input never affects the output of a coalition, then that player should earn nothing. The additivity axiom states that, following an additive technological improvement, a player's payoff should only change to the extent to which the new technology augments the value of his input.

[2] A referee has drawn our attention to the fact that this axiomatic characterization of the Shapley value on the full domain of transferable-utility games is rather often attributed to Shubik (1962) since Shapley combines the axioms of additivity and null player into the carrier axiom.