2017

2017-2 A Non-Parametric Approach to Testing the Axioms of the Shapely Value with Limited Data

Victor Aguiar

Roland Pongou

Jean-Baptiste Tondji

A Non-Parametric Approach to Testing the Axioms of the Shapely Value with Limited Data

by

Victor Aguiar, Roland Pongrou and Jean-Baptiste Tondji

Research Report # 2017-2 **November 2017**

Department of Economics Research Report Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

A Non-Parametric Approach to Testing the Axioms of the Shapley Value with Limited Data*

Victor Aguiar[†] Roland Pongou[‡] and Jean-Baptiste Tondji[§]

November 1, 2017

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 nonparametric 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 or a cost allocation 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.

JEL: C71, C78, D20, D30, J30.

Keywords: Shapley value, fairness violations, limited data, inequality.

*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, Greg Pavlov, Rajiv Vohra, Kareen Rozen, Roberto Serrano, and the participants of the Brown Theory Lunch for useful comments and encouragement.

[†]Department of Economics, University of Western Ontario, vaguiar@uwo.ca.

[‡]Department of Economics, University of Ottawa, roland.pongou@uottawa.ca.

[§]Department of Economics, University of Ottawa, jtond063@uottawa.ca.

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 commercial product's attributes, and causal assessment in an epidemiological or 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.

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 current 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 work 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 recent 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 limited data sets

(which is generically the type of data that most practitioners have access to).[1] 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[2] 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.[3] As is explained below, these two characterizations also provide a basis for studying the formal relationships that exists 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 positive 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

[1] In this regard, it is analogous to the Revealed Preference approach to testing consumer theory models (Afriat, 1973; Varian, 1983).

[2] *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.

[3] A referee has drawn our attention to the fact that the axiomatic characterizations of the Shapley value on the full domain of transferable-utility games in Young (1985) and Shapley (1953) is rather often attributed to Shubik (1962) since Shapley combines the axioms of additivity and null player into the carrier axiom. Pintér (2015) shows that this characterization works for the domain of monotone games, which is the class of games we analyze in this paper.

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, only upon observing one game and one pay profile. 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.[4] This is interesting because, in real-life scenarios, it is very difficult to observe more than one game; it is difficult to observe output under two different technologies, or the ways in which a pay scheme behaves under different production functions at the same level of input factors). Without the non-negativity and feasibility assumptions on the pay scheme, marginality may not have any testable implications in limited datasets.

We develop one main application for our analysis. In particular, we consider the problem of inferring unfairness from a given income distribution in a population with a known production technology. First, we extend our results to a more general environment that involves a finite set of agents who supply inputs in discrete units (e.g., zero, one, two, and so on) up to a maximum amount, and a production technology that maps each input profile to an output. This environment is more flexible than the traditional transferable-utility environment used by Shapley (1953) in that each agent may supply a different amount of his input. In this more general setup, we define the Shapley payoff function, which we compare to two popular pay schemes, namely the quasi-linear contract and the piece-work pay scheme, also known as the linear contract. Unlike the Shapley value, the linear contract is appealing because it is externality-free–that is, the payoff to an agent does not depend on the inputs supplied by the other agents. We find that, in general, these schemes violate all of the axioms that characterize the Shapley value, the exception being that the quasi-linear contract is efficient. This means that the level of inequality produced by these popular pay schemes is generally unfair, as measured by the Shapley axioms. We therefore conduct a comparative-statics analysis that reveals how the pay rate under the linear contract affects the violation of each of these axioms.

To the best of our knowledge, no other work has analyzed and quantified departures from the Shapley value. de Clippel and Rozen (2013) propose a way to test the axioms of symmetry, null player, additivity, and marginality under the assumption of efficiency. Unlike our paper, their main focus in not on quantifying the departures from the Shapley value.[5]

[4] More generally, any non-negative and feasible pay scheme satisfying the marginality axiom on the class of monotonic 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 non-negativity, provides an upper bound on the payoff, which excludes the case of constant pay schemes (except in the case where one assigns a null payoff vector in all monotone games).

[5] They suggest using a regression-based methodology and restrictions over coefficients of such regressions for testing the different axioms at the aggregate level. In particular, they assume that all heterogeneity in the sample is caused by non-systematic (additive) errors and they limit their regression analysis for linear solution concepts. Our approach can be applied

Another related approach to our work can be found in Gomez et al. (2003), who defines a measure of centrality for networks as the difference between the Shapley value and the Myerson value. The decomposition of a goodness-of-fit measure into components that correspond to the violations of axioms was first explored in Aguiar and Serrano (2017) in the context of consumer theory. We study a completely different economic environment. We add to their idea that decomposable measures of departures from classical concepts in economic theory provide a novel way of studying empirical counterparts of such concepts that usually do not conform to the theory. We also hope to complement the classical works of Shapley (1953), Shubik (1962) and Young (1985) by providing an approach to the systematic comparison of any pay profile to the Shapley value pay profile.

Our paper also contributes to the small literature that studies economic inequality using transferable-utility games. Some works on this topic include Einy and Peleg (1991) and Nembua and Wendji (2016). Einy and Peleg (1991) provide an axiomatic characterization of linear inequality measures for coalitional games, obtaining measures that are generalized Gini functions of the Shapley value. Nembua and Wendji (2016) compare linear, symmetric and efficient values in the class of weakly linear transferable-utility games (Freixas (2010)) using the Pigou-Dalton transfers principle and the Lorenz criterion. This class of values includes the Shapley value. We contribute to this literature by answering the question of when inequality can be considered "fair" or "unfair", and by quantifying the sources of any particular case of unfairness based on the axioms that characterize the Shapley value.

The rest of this paper is organized as follows. In section 2, we provide preliminary definitions and introduce the notion of a dataset in a transferable-utility environment. We also recall the two fundamental axiomatic characterizations of the Shapley value. In section 3, we propose a local non-parametric test for violations of the axioms characterizing the Shapley value by any observed pay profile, as well as a decomposable measure of such violations. In section 4, we propose an extension of the analysis to the case of full datasets. In section 5, we provide an application that illustrates the usefulness of our results. We conclude in section 6. All proofs are collected in an appendix.

2 Preliminaries

2.1 Transferable-Utility Environment and Dataset

In this section, we introduce preliminary definitions. Let N be a non-empty and finite set of players or factors[6], with $|N| = n$. A coalition is a non-empty subset C of factors: $C \subseteq N$, $C \neq \emptyset$.

A transferable-utility environment is a pair (N, f) where $f : 2^N \mapsto \mathbb{R}_+$ is a characteristic function

at the individual/subject level and is non-parametric. Therefore, it allows for unconstrained heterogeneity in fairness attitudes among individual games. We also provide a way of identifying the sources of any particular violation of fairness (according to the axioms that characterize the Shapley value).

[6] We use the words "players", "factors" and "inputs" interchangeably. The words "factors" and "inputs" are more general in the context of this paper, given the application to any type of datasets.

such that $f(\emptyset) = 0$. Many strategic and non-strategic situations can be modeled using a transferable-utility environment (e.g., transferable-utility games, firm production, voting games, risk allocation, disease production, etc.). In what remains, we fix N, so that an environment is completely defined by a characteristic function f. Without loss of generality, we will call a transferable-utility environment a game.We denote by Γ^0 the set of all games. Throughout this paper, we will be only concerned with monotone games, which we define below.

Definition 1. (**Monotone game**) A game f is said to be monotone if for any coalitions B and C such that $B \subseteq C \subseteq N$, $f(C) \geq f(B)$. We denote by Γ the set of all monotone games.

It is useful to observe that any game f can be written as an element of $\mathbb{R}_+^{2^n}$ with each component corresponding to the worth of a coalition in N.

A pay scheme, formally defined below, is a way to share the output produced by the grand coalition N among the players.

Definition 2. (**Pay scheme**) A pay scheme is a function $\boldsymbol{\theta} : \Gamma \mapsto \mathbb{R}_+^n$ that maps any monotone game f to a nonnegative real vector $\theta = \boldsymbol{\theta}(f) = (\boldsymbol{\theta}_1(f), \boldsymbol{\theta}_2(f), ..., \boldsymbol{\theta}_n(f))' \in \mathbb{R}_+^n$ such that $\sum_{i \in N} \boldsymbol{\theta}_i(f) \leq f(N)$. The vector θ is called a pay profile, and for each factor $i \in N$, $\theta_i \in \mathbb{R}_+$ is interpreted as the payoff of i out of the output $f(N)$. The set of all the pay schemes is denoted Θ.[7]

In the definition above, the condition $\sum_{i \in N} \boldsymbol{\theta}_i(f) \leq f(N)$ is a **feasibility condition** which says that the total payoff cannot exceed the worth of the grand coalition. In addition, no player can receive a negative payoff. Obviously, these are reasonable assumptions for the domain of monotone games. The *nonnegativity assumption* on a pay scheme can be justified as the impossibility to penalize productive players. Evidently, in a monotone game, the worth of the singleton coalitions is at least equal to zero, hence individual rationality will rules out negative payoffs.[8] The *feasibility assumption* is very natural from an economic point of view. The Shapley value (defined below), in particular, satisfies these two assumptions in the class of monotone games.

We now introduce the notion of a dataset and related concepts.

An **observation** is a pair (f, θ) where f is a game and $\theta \in \mathbb{R}_+^n$ is a pay profile (a distribution of the output generated by the grand coalition). Note that θ here is not necessarily a function of f but only an observed vector of $\mathbb{R}_+^n$ (such as the wage profile in a firm).

Let $\mathcal{T}$ be a non-empty indexed set of games (possibly uncountable). A **dataset** is a list of observations $\mathfrak{D} = (f^t, \theta^t)_{t \in \mathcal{T}}$. A **complete dataset** is a list of observations $\mathfrak{D} = (f^t, \theta^t)_{t \in \mathcal{T}^0}$ where $\mathcal{T}^0$ contains all

[7]Given a monotone game f, any pay profile in this paper is the realization of a possibly unobserved pay scheme; it follows that any pay profile is such that the sum of the players' payoffs is at most equal to $f(N)$.

[8]The nonnegativity assumption is called the monotonicity axiom in Weber (1988) and Weak monotonicity in Malawski (2013) for the domain of monotone games. Kalai and Samet (1987) also use the nonnegativity condition on the context of weighted Shapley values.

possible games f (i.e., there is a one-to-one function between $\mathcal{T}^0$ and Γ). A **limited dataset** is a list of observations $\mathfrak{D} = (f^t, \theta^t)_{t \in \mathcal{T}}$ where $\mathcal{T} = \{1, \cdots, T\}$ consists of a finite number of games f. Any given limited dataset is a subset of the complete dataset.

Definition 3. (**Data generating pay scheme**) We say that $\boldsymbol{\theta} : \Gamma \to \mathbb{R}^n_+$ is a data generating pay scheme if it is the unique pay scheme such that $\boldsymbol{\theta}(f^t) = \theta^t$ for any element (f^t, θ^t) of the complete dataset $\mathfrak{D} = (f^t, \theta^t)_{t \in \mathcal{T}^0}$.

We assume that the complete dataset always has a data generating pay scheme. In the context of a limited dataset, we do not have the details about how the data generating pay scheme $\boldsymbol{\theta}$ distributes the total output $f^\tau(N)$ for a game τ that is not in the dataset T (i.e, $\tau \notin \mathcal{T}$); but we know the realized pay profile θ^t for any game $t \in \mathcal{T}$ (and we know that the data generating pay scheme $\boldsymbol{\theta}$ is such that $\theta^t = \boldsymbol{\theta}(f^t)$ for all $t \in \mathcal{T}$). However, we have full information on f^t, (i.e., we know the values of $f^t(C)$ for all $C \subseteq N$). An example of an observation (f, θ) for $n = 2$, is given as follows: $f(\emptyset) = 0$, $f(\{i\}) = 1$ for $i = 1, 2$, $f(\{i, j\}) = 2$, and $\theta_i = 1$ for $i = 1, 2$. We have no information about how θ depends on f, but we know f for all coalitions.

In the analysis, we consider the case where the data generating pay scheme $\boldsymbol{\theta}$ is fixed (i.e. the solution concept is not changing and there is only one true solution concept) but unknown to the observer. In practice, we can assume that we have only one observation (f, θ) in an environment such a firm, and we only know that the data generating pay scheme $\boldsymbol{\theta}$ is such that $\boldsymbol{\theta}(f) = \theta$ (i.e., it is very difficult to observe two production functions in real life, or two alternative labor contracts for the same players).

Since we can never observe a complete dataset, because the set of all of the possible games is infinite in our environment, we will focus on the idea of extending an pay scheme observed for a limited dataset to the complete dataset.

Definition 4. (**Extension of a dataset**) An extension of the set of pay profiles observed in a limited dataset $\mathfrak{D} = (f^t, \theta^t)_{t \in \mathcal{T}}$ to the domain Γ is a pay scheme $\boldsymbol{\vartheta} : \Gamma \mapsto \mathbb{R}^n_+$ such that the restriction $\boldsymbol{\vartheta}|\mathfrak{D} : \mathfrak{D} \mapsto \mathbb{R}^n_+$ satisfies $\boldsymbol{\vartheta}(f^t) = \theta^t$ for all $t \in \mathcal{T}$, and for any $g \in \Gamma$, $\sum\limits_{i \in N} \boldsymbol{\vartheta}_i(g) \leq g(N)$ and $\boldsymbol{\vartheta}_i(g) \geq 0$. The set of all possible extensions of a dataset $\mathfrak{D} = (f^t, \theta^t)_{t \in \mathcal{T}}$ is denoted by $\boldsymbol{V}(\mathfrak{D})$.

Remark that for any limited dataset $\mathfrak{D}$, the **data generating pay scheme** belongs to $\boldsymbol{V}(\mathfrak{D})$. Of course, the observer cannot know exactly which of the extensions of a limited dataset is the data generating pay scheme. However, we can make inference about the properties of the data generating pay scheme using a limited dataset. In particular, when all the extensions of a limited dataset fail a property, then the data generating pay scheme must also fail this property. We exploit this observation in our main result (Theorem 1).

2.2 The Shapley Value

In this section, we recall the definition of the Shapley value for transferable-utility environments as well as its two fundamental axiomatic characterizations. These characterizations provide an axiomatic basis for analyzing the different ways in which an arbitrary pay scheme might violate basic principles of fairness. The following definition will be needed for the statement of these characterizations.

Definition 5. Let $i, j \in N$ be two players, and f be a game.

1. The marginal contribution of player $i \in N$ to a coalition $C \subseteq N$ such that $i \notin C$ is $f(C \cup \{i\}) - f(C)$, and is denoted by $mc(i, f, C)$.
2. Player i is a *null-player at* f if for any coalition $C \subseteq N$ such that $i \notin C$, we have $mc(i, f, C) = 0$.
3. Players i and j are said to be symmetrical or identical *at* f if for any coalition $C \subseteq N$ such that $i, j \notin C$, $mc(i, f, C) = mc(j, f, C)$.

We now define the axioms that characterize the Shapley value.

Axiom 1. ***(Symmetry or Equal-treatment)***
A pay scheme $\boldsymbol{\theta}$ *satisfies the symmetry or equal-treatment property if for any game* f*, and any players* i *and* j *that are symmetrical at* f*,* $\boldsymbol{\theta}_i(f) = \boldsymbol{\theta}_j(f)$*.*

Axiom 2. ***(Efficiency)***
A pay scheme $\boldsymbol{\theta}$ *is efficient if for any game* f*,* $\sum_{i \in N} \boldsymbol{\theta}_i(f) = f(N)$*.*

Axiom 3. ***(Marginality/Strong monotonicity)***
A pay scheme $\boldsymbol{\theta}$ *satisfies marginality if for any games* f *and* w*, any player* $i \in N$*,* $[mc(i, f, C) \geq mc(i, w, C); \forall C \subseteq N \setminus \{i\}] \Rightarrow [\ \boldsymbol{\theta}_i(f) \geq\ \boldsymbol{\theta}_i(w)]$*.*

Axiom 4. ***(Null player property)***
A pay scheme $\boldsymbol{\theta}$ *satisfies the axiom of null-player if for any game* f*, and any null-player* $i \in N$ *at* f*,* $\boldsymbol{\theta}_i(f) = 0$*.*

Axiom 5. ***(Additivity)***
A pay scheme $\boldsymbol{\theta}$ *is additive if for any games* f *and* w*,* $\boldsymbol{\theta}(f + w) = \boldsymbol{\theta}(f) + \boldsymbol{\theta}(w)$*.*

These axioms require little justification, with perhaps the exception of the additivity axiom. The equal-treatment axiom is a no-discrimination condition (horizontal equality) that requires that players who make the same marginal contribution in a game f receive the same pay. Efficiency requires that the output of the grand coalition be fully shared among the various contributors. It can also be thought of in terms of Pareto optimality because if a pay profile is feasible but not efficient, it cannot be Pareto optimal under very general conditions (on the players' tastes). Marginality means that a player's pay should be greater under a game that places a higher marginal value on his participation (or input). This is a very appealing property because it requires that the payoff of a player depends only on his marginal contribution given other players' inputs. The null player property requires that those who do not contribute marginally should not receive any part of the realized output. The additivity axiom means that, following an additive technological improvement, a player's payoff should only change by the extent to which the new technology augments the value of his input.[9]

Despite the appeal of these axioms, it should be noted that testing axioms that are defined using at least two games such as marginality and additivity requires that two or more characteristic functions be observed. This is not possible in a real-world setting, as we only have access to a limited dataset (e.g., we typically observe only one game, which could, for example, be the production function or technology of a firm). A distinctive feature of our work is that we are able to quantify departures of any pay scheme from these axioms in limited datasets, which also means that our analysis has testable implications.

The results set out hereunder establish the necessary and sufficient axioms that characterize the Shapley payoff function (defined by equation (1) below).

Claim 1. (Young (1985)) There exists a unique pay scheme, denoted $\boldsymbol{\varphi}$, that satisfies the efficiency, equal-treatment, and marginality axioms, and it is given, for any game f, by:

$$\boldsymbol{\varphi}_i(f) = \sum_{C \subseteq N \setminus \{i\}} \frac{|C|!(n-|C|-1)!}{n!}[f(C \cup \{i\}) - f(C)], \ for \ all \ i \in N. \tag{1}$$

Claim 2. (Shapley (1953)) The pay scheme $\boldsymbol{\varphi}$ defined by (1) is the unique pay scheme that satisfies efficiency, the null-player axiom, equal-treatment, and additivity.

These two characterization are the most popular in the literature. In the view of many scholars, they also provide the most important foundation of the Shapley value as a concept of distributive justice or fairness. Pintér (2015) shows that the characterization of Shapley using the axioms in Young (1985) is valid in the domain of monotone games (Γ), which is the class of games we consider in this paper. The axiomatization due to Shapley (1953) is also valid for monotone games.

In order to understand the Shapley payoff function, one should recall that for coalition C, $|C|$ is the size of the coalition. We assume that players enter the production process in a random order and that all of the $|S|!$ orderings of the players supplying a positive level of effort are equally likely. Then the fraction $\frac{|C|!(n-|C|-1)!}{n!}$ represents the probability that a given player i joins a coalition C (such that $i \notin C$). When a player i joins the other players who have already chosen to join the coalition, the new coalition is $C \cup \{i\}$ and the game's outcome is $f(C \cup \{i\})$; thus the marginal contribution of player i is $f(C \cup \{i\}) - f(C)$. The value φ_i is the expected marginal contribution of player i. Throughout this paper, we denote $\frac{|C|!(n-|C|-1)!}{n!}$ by ω_C for fixed N, thus the Shapley value can be written as $\boldsymbol{\varphi}_i(f) = \sum_{C \subseteq N \setminus \{i\}} \omega_C m(i, f, C)$ for all $i \in N$.

We link the two characterizations of the Shapley value by means of the lemma described hereunder, which holds for the class of monotone games, under our definition of a pay scheme (assumed to be nonnegative and feasible).

[9] This interpretation is interesting in the context of the ongoing debate on how technological improvement affects income inequality. According to the additivity axiom, if a technological improvement nullifies the value of a player's input, then that player should be laid off. This has recently been observed in stores where cashiers are replaced by electronic machines.

Lemma 1. *If a pay scheme $\boldsymbol{\theta}$ satisfies the marginality axiom, then $\boldsymbol{\theta}$ satisfies the null-player axiom.*

The proof of Lemma 1 is easy. However, its implication for a direct, though partial, test of marginality is important. In fact, the marginality axiom a priori is not easy to test because it is defined using two games. But Lemma 1 implies that, despite this fact, this axiom can be partially tested upon observing only one game and a pay profile because the null-player axiom is defined using only one characteristic function. Indeed, if we want to know whether a given pay scheme satisfies the marginality axiom under a given game, we can first determine whether it satisfies the null-player axiom. If it violates the null-player axiom, then we can safely conclude that it violates the marginality axiom thanks to Lemma 1. Again, this is true because we are only concerned with the class of monotone games and a pay scheme is by our definition nonnegative and feasible. However, as acknowledged above, notice that testing the marginality axiom via the null-player axiom only provides a partial test because if a pay scheme satisfies the null-player axiom, it does not necessarily mean that it satisfies the marginality axiom. In the next sections, we will provide a way to fully test the marginality axiom only upon observing one game.

We now prove the following and perhaps simple result, which is a corollary of 1.

Corollary 1. *If for an observation (f, θ), there exists a null-player $i \in N$ at f such that $\theta_i > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ fails the null-player axiom (i.e., all possible extensions of (f, θ) fail the null-player axiom). Moreover, $\boldsymbol{\theta}$ fails the marginality axiom.*

If $\boldsymbol{\theta}$ is the data generating pay scheme of (f, θ), then by definition $\boldsymbol{\theta}(f) = \theta$. It follows that $\boldsymbol{\theta}_i(f) = \theta_i > 0$ for the null-player $i \in N$ at f. By Lemma 1 it must therefore be that $\boldsymbol{\theta}$ fails marginality.[10] Evidently, the pay profile $\theta \in \mathbb{R}^n_+$ cannot be said to fail marginality because it is not a pay scheme, but instead is an observed distribution of the output. However, we can use a single observation (f, θ) to reject the **null hypothesis** that the data generating pay scheme $\boldsymbol{\theta}$ of the limited dataset satisfies the marginality axiom.

Some remarks are in order. (i) First, holds only for monotone games (it holds for Γ not for Γ^0) under the nonnegativity and feasibility restrictions imposed on the data generating pay scheme; and (ii) second, does not exhaust all the empirical implications of marginality for the domain of monotone games (for the complete result see Theorem 2).

We state below a similar result for the case of the symmetry and efficiency axioms.

Corollary 2.

1. *If for any observation (f, θ), there are any two players i and j that are symmetrical at f such that $\theta_i \neq \theta_j$, then the data generating pay scheme $\boldsymbol{\theta}$ fails the symmetry axiom.*

[10] Equivalently, we can establish this result using the idea of the set of extensions of a dataset, and using a proof by contradiction. First, assume that there exists an extension $\boldsymbol{\vartheta}$ of (f, θ) such that $\boldsymbol{\vartheta}$ satisfies the null-player axiom. Because $\boldsymbol{\vartheta}$ extends the dataset it must be that $\boldsymbol{\vartheta}(f)_i = \theta_i > 0$ for a null-player $i \in N$, but this is a contradiction. We conclude that every extension in $\boldsymbol{V}((f, \theta))$ has to fail the null-player axiom; moreover it fails marginality. Since the data generating pay scheme is an element of the set of extensions, the result is established.

2. If for any observation (f, θ), it is the case that $\sum_{i \in N} \theta_i < f(N)$, then the data generating pay scheme $\boldsymbol{\theta}$ fails the efficiency axiom.

3 Quantifying the Departures from the Shapley Value

Our main goal is the comparison of any pay profile θ with the Shapley payoff function $\boldsymbol{\varphi}$ in limited datasets. Denote the Euclidean norm defined in $\mathbb{R}^n$ by $|| \cdot ||$. Also denote the inner product associated with the euclidean norm by $< \cdot, \cdot >$. We have the following definition of the Shapley distance.

Definition 6. For any fixed game f, the **Shapley distance** of a pay profile $\theta \in \mathbb{R}^n_+$, denoted $||\theta - \boldsymbol{\varphi}(f)||$, is the distance between θ and the Shapley pay profile $\boldsymbol{\varphi}(f) \in \mathbb{R}^n_+$ at f.

We provide below a unique **orthogonal decomposition** of the square of the Shapley distance into terms that measure violations of the classical axioms of the Shapley value. This approach is analogous to that of Aguiar and Serrano (2017) who study departures of a demand function from rationality. Despite the similarity of the two approaches, in this paper, we are tackling a completely new question in a different environment.

Moreover we prove that in finite datasets, these terms can be used to make partial inferences about the violations of the axioms defined for complete datasets, and to make complete inference about the violations of the axioms defined for a fixed game. This is of interest because the observer usually does not have information about a pay scheme under different games, thus making it practically impossible to check the validity of the axioms that require comparisons between different games.

3.1 An Orthogonal Decomposition of the Shapley Distance with Limited Datasets

We now provide an orthogonal decomposition of the Shapley distance. Let f be a game and $\theta \in \mathbb{R}^n_+$ an observed pay profile generated by a pay scheme that may not be known (to the observer). We can always decompose it into a sum of the Shapley value at the observed game f and an error term $\theta = \boldsymbol{\varphi}(f) + e^{sh}$, by defining $e^{sh} = \theta - \boldsymbol{\varphi}(f) \in \mathbb{R}^n$. Moreover, we are going to show that the error term e^{sh} can be further decomposed uniquely into three vectors that are orthogonal to each other, with these vectors being respectively connected to the violation of symmetry (sym), efficiency (eff), and marginality (mrg). Formally, this means that we can write $e^{sh} = e^{sym} + e^{eff} + e^{mrg}$ such that the inner product of these axioms errors (roughly their correlation) is zero.

We find this orthogonal decomposition to be the result of the following procedure. First we find the closest pay scheme to θ that satisfies sym (remark that there is no conceptual issue here even though the pay scheme is a function and θ is a pay profile or a point; see below for a formalization) ; then we find the closest pay scheme to θ that satisfies eff in addition to sym; and finally we find the closest pay scheme to θ that satisfies mrg in addition to sym and eff, which is simply the Shapley value itself. The order in which we impose these constraints is the only one that we know that produces the orthogonality of the different error vectors. This decomposition is also meaningful as each component measures a quantity of economic interest that completely and effectively "isolates" one of the three conditions sym, eff and mrg.

We start by fixing a pair consisting of an observed pay profile and a game (θ, f) and consider the Shapley distance of θ at this point, which is:

$$||e^{sh}|| = ||\theta - \boldsymbol{\varphi}(f)||.$$

Let $\boldsymbol{v}^{sym}$ be the closest pay scheme to θ that satisfies symmetry (pointwise under the chosen norm) (i.e., $\boldsymbol{v}^{sym} \in argmin_{\boldsymbol{v}\in\boldsymbol{\Theta}}||\theta - \boldsymbol{v}(f)||$ s.t. $\boldsymbol{v}$ satisfies sym).[11] We prove that each entry evaluated at f is given by v_i^{sym} that corresponds to the average pay under θ among the players who are symmetrical or identical to i under f. We then establish that θ can be written uniquely as the sum of its symmetric part $v^{sym} = \boldsymbol{v}^{sym}(f)$ and a residual e^{sym} that is orthogonal to v^{sym} under the Euclidean inner product:

$$\theta = v^{sym} + e^{sym}.$$

In a similar way, let $\boldsymbol{v}^{sym,eff}$ be the pay scheme that is pointwise closest to the symmetric pay scheme v^{sym} and that satisfies efficiency (i.e. $\boldsymbol{v}^{sym,eff} \in argmin_{\boldsymbol{v}\in\boldsymbol{\Theta}}||v^{sym} - \boldsymbol{v}(f)||$ s.t. $\boldsymbol{v}$ satisfies sym and eff). We prove that $v_i^{sym,eff} = \boldsymbol{v}_i^{sym,eff}(f)$ is given by the summation of v_i^{sym} and the output wasted by θ divided by the number of players in N. Again, we show that we can write v^{sym} uniquely as:

$$v^{sym} = v^{sym,eff} + e^{eff},$$

where e^{eff} is the negative of the wasted output by θ divided by the number of players in N.

Finally, we exploit the fact that the pay scheme satisfying the axiom of marginality that is pointwise closest to the symmetric and efficient pay scheme $v^{sym,eff}$, which we denote by $\boldsymbol{v}^{sym,eff,mrg}$, must be the Shapley value because of the uniqueness established in Claim 1. Thus $v^{sym,eff,mrg} = \boldsymbol{\varphi}(f)$. We let $e^{mrg} = v^{sym,eff} - \boldsymbol{\varphi}(f)$. Notice that we can always decompose θ (pointwise) as:

$$\theta = \boldsymbol{\varphi}(f) + e^{sh},$$

because θ and $\boldsymbol{\varphi}(f)$ belong to the same vector space. With this preview in hand, we establish the main result of this section.

Theorem 1. *For any given observation (f, θ), we have the unique pointwise decomposition:*

$$\theta = \boldsymbol{\varphi}(f) + e^{sym} + e^{eff} + e^{mrg}.$$

Moreover, the distance to the Shapley pay scheme can be uniquely decomposed as:

$$||e^{sh}||^2 = ||e^{sym}||^2 + ||e^{eff}||^2 + ||e^{mrg}||^2,$$

into its symmetric, efficiency and marginality departures, such that for any $i, j \in \{sym, eff, mrg\}$, $i \neq j$, $< e^i, e^j >= 0$.

[11] Existence is easy to verify noticing that the space of symmetric pay schemes (that are also monotone) is convex and closed.

The proposed decomposition of the Shapley distance that we just derived has economic meaning described hereunder:

a) $||e^{sym}||^2 = \sum_{i \in N} [\theta_i - v_i^{sym}]^2$, where for any player i, v_i^{sym} is the average payoff within the class $[i]^f$ of players who are symmetric or equivalent to i at f. This means that $||e^{sym}||^2$ is a dispersion measure within the equivalence classes of players. In other words, this quantity measures horizontal inequality, which is the inequality among players who are identical.

b) $||e^{eff}||^2 = E^2/n$, where $E = [f(N) - \sum_{i \in N} \theta_i]$ is the total waste produced by the pay profile. This means that $||e^{eff}||^2$ increases solely due to the lack of efficiency.

c) $||e^{mrg}||^2 = \sum_{i \in N} [v^{sym,eff} - \varphi(f)]^2$, where $v^{sym,eff}$ is the symmetrized and efficient pay profile that is closest to the original pay profile θ. This means that $||e^{mrg}||^2$ is a measure of departures from the marginality principle conditional on fulfilling horizontal equality and efficiency.

$||e^{sh}||^2$ is the first measure of departures from the Shapley axioms. It has the unique advantage to be a unified treatment of the three axioms in the form of a numerical and additive decomposition. Due to its non-parametric and deterministic nature, it can be applied to individual games. In this, it differs from existing tests based on regression analyses and using samples such as those proposed in de Clippel and Rozen (2013). In this regard, our approach is analogous to the revealed preference methodology used to empirically test models in consumer theory. A clear advantage of our decomposition analysis is that each component of $||e^{sh}||^2$ measures a violation of a Shapley axiom, with the main result providing a formal and unified theoretical foundation for using the three components.

In order to prove Theorem 1, we need some preliminary lemmas that are interesting in their own rights. We define first the **equivalence class of symmetric players at** f: $[i]^f = \{j \in N : i \sim^{sym} j \text{ in } f\}$, where $i \sim^{sym} j$ indicates that players i and j are symmetric players in f

Lemma 2. *The closest pay scheme satisfying symmetry to any pay profile θ is given by $\boldsymbol{v}^{sym}$, which is a pay scheme that gives the average pay of a group of symmetrical players to each of the players at f: $\boldsymbol{v}_i^{sym}(f) = \frac{1}{|[i]^f|} \sum_{j \in [i]^f} \theta_j$.*

We let $v^{sym} = \boldsymbol{v}^{sym}(f) \in \mathbb{R}_+^n$ be the pay profile generated by the closest pay scheme that satisfies symmetry. Now, we present the solution to the closest efficient pay scheme.[12] (Its proof is obvious and thus is omitted.)

Lemma 3. *The the closest pay scheme that satisfies symmetry and efficiency, to any pay profile θ and in particular to the symmetric pay profile v^{sym}, is given by $\boldsymbol{v}^{sym,eff}$, which is a pay scheme that gives*

[12] An anonymous referee accurately pointed out that the construction of $v^{sym,eff}$ applies the same principle as the principle used to construct the least square prenucleolus from the Banzhaf value (Ruiz et al. (1996)) and the efficient extension of the Myerson value in van den Brink et al. (2012). Béal et al. (2015) remark that this efficient extension is the closest efficient payoff vector to the Myerson value according to the Euclidean distance.

each player i his payoff according to v_i^{sym} plus the wasted output shared equally among all the players: $\boldsymbol{v}_i^{sym,eff}(f) = v_i^{sym} + \frac{[f(N) - \sum_{i \in N} \theta_i]}{n}$.

We let $v^{sym,eff} = \boldsymbol{v}^{sym,eff}(f) \in \mathbb{R}_+^n$ be the pay profile generated by the closest pay scheme that satisfies symmetry and efficiency. We also need to prove a mathematical lemma that will be crucial to prove our decomposition result. We first define a skew symmetric pay scheme.

Definition 7. A skew symmetric pay scheme is a pay scheme such that for an equivalence class defined by $i \sim^{sym} j$, where i and j are identical players in f, we have :

$$\sum_{j \in [i]^f} \boldsymbol{v}_j(f) = 0.$$

Notice that when there are only two players who are identical, say $i \sim^{sym} j$, we have the usual notion of skew symmetry in that $v_i = -v_j$. Moreover, for the case of a unique player k to whom no other player is identical, we have $v_k = 0$.

We also need to introduce the notion of orthogonal pay schemes. We say that two pay schemes $\boldsymbol{\theta}, \boldsymbol{\eta}$ are orthogonal if for every $f \in \Gamma$, $< \boldsymbol{\theta}(f), \boldsymbol{\eta}(f) >= 0$. Now, we are ready to prove the following property of skew symmetric pay schemes.

Lemma 4. *Any skew symmetric pay scheme is orthogonal to any symmetric pay scheme.*

This is the appropriate moment to prove Theorem 1, which is done in the appendix. To establish the decomposition, we prove that the different residuals e^{sym}, e^{eff} and e^{mrg} are orthogonal to one another. Then the main decomposition theorem follows as a consequence.

3.2 A Test of the Violations of the Shapley Axioms with Limited Data

In what follows, we provide a test of the violation of the axioms that characterize the Shapley value with limited data. Given a game f and an observed payoff profile $\theta \in \mathbb{R}^n$ (e.g., f may be the production function of a firm and θ the wage profile of that firm), we can test not only whether its underlying data generating pay scheme $\boldsymbol{\theta}$ departs from the Shapley value, but also identify the Shapley axioms that may be violated by $\boldsymbol{\theta}$. More importantly, we quantify the size of each violation. The test of the marginality axiom reveals that this axiom is much stronger for the characterization of the Shapley value than needed, which requires us to define the marginality upper bound $K(f)$. We emphasize that we use only the limited dataset (f, θ) to make inference about the behavior of the unobserved data generating pay scheme $\boldsymbol{\theta}$.

We first need the definitions below.

Definition 8. (**Marginalist pay scheme**) A pay scheme $\boldsymbol{\theta}$ is said to be marginalist if it admits a representation:

$$\boldsymbol{\theta}_i(f) = \boldsymbol{\phi}_i(\{f(C \cup \{i\}) - f(C)\}_{C \subseteq N \setminus \{i\}}), \forall i \in N,$$

for some non-decreasing function $\boldsymbol{\phi}_i : \mathbb{R}^{2^{n-1}} \mapsto \mathbb{R}$ such that $\phi_i(0) = 0$ for any game f .

Define (the closed and bounded set) $\mathcal{F}^{mrg} \subset \mathbb{R}^n$, such that $\vartheta \in \mathcal{F}^{mrg}$ if $\vartheta = (\vartheta_i)_{i\in N}$ is a pay profile (i.e., $\vartheta_i \in \mathbb{R}$, $\vartheta_i \geq 0$, $\sum_{i\in N} \vartheta_i \leq f(N)$) and is such that if $mc(i, f, C) \geq 0$ for all $C \subseteq N \setminus \{i\}$ then $\vartheta_i \geq 0$, and if $mc(i, f, C) \leq 0$ for all $C \subseteq N \setminus \{i\}$ then $\vartheta_i = 0$ (let $mc(i, f) = (mc(i, f, C))_{C\subseteq N\setminus\{i\}}$ for all $i \in N$). Define the (closed set) $\mathcal{F}^{mrg,sym} \subset \mathcal{F}^{mrg}$ of symmetric pay profiles such that $\vartheta \in \mathcal{F}^{mrg,sym}$ if $\vartheta \in \mathcal{F}^{mrg}$ and $\vartheta_i = \vartheta_j$ whenever $i \sim^{sym} j$ at f. Define the (closed set) $\mathcal{F}^{mrg,eff} \subset \mathcal{F}^{mrg}$ such that $\vartheta \in \mathcal{F}^{mrg}$ and $\sum\limits_{i\in N} \vartheta_i = f(N)$; when $n \geq 3$, this is equivalent to saying that ϑ_i can be written as a random value evaluated at f (the notion of a random value is defined below). In fact, this means that $\mathcal{F}^{mrg,eff}$ coincides with the set of random values for a fixed f. Finally, we define the (closed set) $\mathcal{F}^{mrg,sym,eff} \equiv \{\varphi(f)\}$ that corresponds to the Shapley value at f, i.e., $\vartheta_i = \varphi_i(f)$. This is the only value that satisfies marginality, symmetry, and efficiency. With this in hand, we define the following set function that depends on f and θ:

$$\mathcal{F}(\boldsymbol{\theta}(f)) = \begin{cases} \{\boldsymbol{\varphi}(f)\} & if \quad ||e^{sym}|| = 0 \text{ and } ||e^{eff}|| = 0 \\ \mathcal{F}^{mrg,eff} & if \quad ||e^{eff}|| = 0 \\ \mathcal{F}^{mrg,sym} & if \quad ||e^{sym}|| = 0 \\ \mathcal{F}^{mrg} & otherwise. \end{cases}$$

We show in the appendix that $\mathcal{F}(\boldsymbol{\theta}(f))$ is a closed and bounded set for a fixed point (θ, f). We now define the marginality upper bound.

Definition 9. (**Marginality upper bound**) The marginality upper bound of any pay scheme θ is the following non-negative constant:

$$K(f) =$$

$$\max_{\vartheta\in\mathcal{F}(\boldsymbol{\theta}(f))} \sum_{i\in N} \left\{ [\frac{f(N)}{n} - \boldsymbol{\varphi}_i(f)] - [\frac{1}{n}\sum_{k\in N} \vartheta_k - \frac{1}{|[i]^f|}\sum_{j\in[i]^f} \vartheta_j] \right\}^2 . \tag{2}$$

The intuition behind $K(f)$ is that it is a critical value of $||e^{mrg}||$ under the null hypothesis that the data generating pay scheme $\boldsymbol{\theta}$ is consistent with marginality. In an analogous way to how critical values are derived for statistical hypothesis tests, in our deterministic framework, we test the null hypothesis of marginality and compute $||e^{mrg}||$ in the worst (upper bound) possible case. This critical value allows us to avoid providing false positives, than can happen when $||e^{mrg}|| = 0$ while at the same time, it is impossible that the data generating pay scheme $\boldsymbol{\theta}$ satisfies marginality.

In practice, this upper bound is easily computed for small to moderate size of players, such as in a laboratory. For large number of players, if we assume efficiency and at least three players $n \geq 3$, then we can pin down a random value which helps reduce the computational complexity of the problem. We denote by $R(N)$ the set of all possible linear orderings defined on N, and $\gamma \in \Delta(R(N))$ the simplex of probabilities defined over it. Let f be a game and $r \in R(N)$ be a given order of players. We define by

$C(r^i) = \{j \in N\backslash\{i\} | j \ r \ i\}$ the set of players that precede i in the order r, and we denote $mc(i, f, r) = f(C(r^i) \cup \{i\}) - f(C(r^i))$ the marginal contribution of a player $i \in N$ to the coalition $C(r^i)$.

Definition 10. (**Random value**) A pay scheme $\boldsymbol{\theta}$ is a random value if it admits a representation:

$$\boldsymbol{\theta}_i(f) = \sum_{r \in R(N)} \gamma(r) mc(i, f, r),$$

for any f.[13]

Under the efficiency axiom, the marginality axiom implies that $\boldsymbol{\theta}$ is a random value (this is established in Theorem 2 in Khmelnitskaya (1999)). The definition below introduces the "marginality upper bound" of an efficient pay scheme.

Definition 11. (**Marginality upper bound with efficiency**) The marginality upper bound of any efficient pay scheme θ is the following nonnegative constant:

$$K^{eff}(f) = \max_{\gamma \in \Delta(R(N))} \sum_{i \in N} \left\{ \sum_{r \in R(N)} \left\{ \gamma(r) - \frac{1}{n!} \right\} \frac{1}{|[i]^f|} \sum_{j \in [i]^f} mc(j, f, r) \right\}^2 .$$

The marginality upper bound is the square of the maximum possible distance from the set of *symmetrized* random values at f to the corresponding Shapley value. Recall that the Shapley value is a random value with the following uniform distribution:

$$\boldsymbol{\varphi}_i(f) = \frac{1}{n!} \sum_{r \in R(N)} mc(i, f, r).$$

This quantity can be computed under limited datasets because it only requires the knowledge of f but it is independent of $\boldsymbol{\theta}$.

The importance of this quantity is established next.

Theorem 2. *For a given observation (f, θ):*

(i) If $||e^{sh}|| > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ fails either symmetry, efficiency, or marginality;

(ii) If $||e^{sym}|| > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ fails symmetry;

(iii) If $||e^{eff}|| > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ fails efficiency; and,

(iv) If $||e^{mrg}|| > \sqrt{K(f)}$, then the data generating pay scheme $\boldsymbol{\theta}$ fails marginality, where $K(f)$ is the marginality upper bound.

(v) If $n \geq 3$, with θ generated by a pay scheme $\boldsymbol{\theta}$ that satisfies efficiency, it follows that: if $||e^{mrg}|| > \sqrt{K^{eff}(f)}$, then the data generating pay scheme $\boldsymbol{\theta}$ fails marginality, where $K^{eff}(f)$ is the marginality upper bound with efficiency.

[13] See Weber (1988) for a treatment of random values.

(*vi*) *If* $||e^{sym}|| = 0$, $||e^{eff}|| = 0$, *and if* $||e^{mrg}|| > 0$, *where* $K(f) = 0$, *then the data generating pay scheme* $\boldsymbol{\theta}$ *fails marginality.*

Notice that if θ is generated by a pay scheme $\boldsymbol{\theta}$ that is efficient and symmetric, we establish that if $||e^{mrg}|| > 0$, then marginality fails. Theorem 2 also deals with the case where $\boldsymbol{\theta}$ is efficient but fails the symmetry axiom, and with the case where $\boldsymbol{\theta}$ fails both symmetry and efficiency simultaneously. It establishes that if $||e^{mrg}||$ is larger than a non-zero constant, we can conclude that marginality fails with certainty, even in limited datasets.[14] While it suffices for the norm of the residuals associated with the symmetry and the efficiency axioms to be strictly positive for these axioms to be violated, this is not the case for the marginality axiom. In fact, this axiom is much stronger for the characterization of the Shapley value than needed, and as we will show later, it "correlates" with the symmetry axiom.[15]

Example 1. Let $N = \{1, 2, 3\}$, if (f, θ) is such that $f(C) = \alpha$ for all $C \subseteq N$ such that $1 \in C$ and $f(C) = 0$ for all $C \subseteq N$ such that $1 \notin C$, and $\theta_i = \frac{\alpha}{3}$ for all $i \in N$, then $||e^{sym}|| = 0$, $||e^{eff}|| = 0$ and $||e^{mrg}|| = \sqrt{2(\frac{\alpha}{3})^2} > 0$. Clearly, the data generating pay scheme $\boldsymbol{\theta}$ fails the null-player axiom and therefore marginality. The marginality distance $||e^{mrg}|| > 0$ quantifies the intensity of this violation of marginality as a function of the excess pay to the two null players 2 and 3. This example illustrates that Theorem 2 contains as a special case an an implication of Lemma 1 for testing the marginality axiom using only one observation.

For cases where $\boldsymbol{\theta}$ fails symmetry or efficiency, the *marginality upper bound* provides the maximum value the distance $||e^{mrg}||$ can take if $\boldsymbol{\theta}$ is a marginalist pay scheme. The bound is obtained by direct computation. In the case where θ is such that $||e^{mrg}|| > \sqrt{K(f)}$, it is impossible that $\boldsymbol{\theta}$ is a marginalist value, thus we can safely conclude, with one observation, that the data generating pay scheme fails marginality. An analogous reasoning applies for the marginality upper bound under efficiency, which exploits the fact that if $\boldsymbol{\theta}$ satisfies marginality and efficiency, it has to be a random value.

Due to the importance of the marginality axiom in the literature, we would like to provide an additional analysis of the violation of this axiom by exploiting its relationship with the null player and the additivity axioms.

[14]The converse implication is not true in general. To see this, consider the null game $w(C) = 0$ for all coalition $C \subseteq N$. Even if the data generating pay scheme fails symmetry or efficiency or marginality due to the nonnegativity and feasibility of the pay scheme, every player must receive a zero payoff. Then $||e^{sh}|| = 0$ for this observation $(w, 0)$. This is common with empirical tests with limited data (Afriat (1973); Varian (1983)), and the fact that for certain observations it may not be possible to detect the violations of the axioms is called lack of power. This is not a defect of the test but rather an artifact of limited datasets.

[15]For this reason, in a former version of the paper, we provide a new axiomatic characterization of the Shapley value that uses a new marginality axiom, and this allows us to conclude whether this new axiom is violated if $||e^{mrg}|| > 0$. The formal results are available upon request.

3.3 Additional Decomposition of the Shapley Distance: The Case of the Null-Player and Additivity Axioms

We can further decompose the residual term e^{mrg} into the null-player and additivity axioms to detect violations of marginality without computing the marginality upper bound. For this reason, we exploit the second characterization of the Shapley value which uses the axioms of additivity and null player (in addition to efficiency and symmetry). In addition, we use the result stating that marginality implies the null-player axiom. Our decomposition proceeds by first imposing the null-player and the additivity axioms.

Denote by $\boldsymbol{v}^{sym,eff,null}$ the closest pointwise pay scheme approximation of a symmetric and efficient pay profile $v^{sym,eff} = \boldsymbol{v}^{sym,eff}(f)$ satisfying the null-player axiom. We have the following result.

Lemma 5. *The closest pay scheme satisfying the symmetry, efficiency, and null-player axiom to any pay profile θ and in particular to the symmetric pay profile $v^{sym,eff}$ is given by $\boldsymbol{v}^{sym,eff,null}$, where $\boldsymbol{v}^{sym,eff,null}(f)$ is the sum of $v^{sym,eff}$ and the vector e_i^{null} that extracts all of the payoffs from the null players and shares it equally among all of the remaining players, while giving zero to the null players. Formally:*

$$e_i^{null} = -\frac{\sum_{k\in\mathcal{N}^f}[v_k^{sym,eff}]}{[n - |\mathcal{N}^f|]} \ \ for \ i \in N\backslash\mathcal{N}^f,$$

and

$$e_k^{null} = -v_k^{sym,eff} \ \ for \ k \in \mathcal{N}^f,$$

with $\mathcal{N}^f$ being the set of null players in f. Thus,

(i) $\boldsymbol{v}_i^{sym,eff,null}(f) = v_i^{sym,eff} + \frac{\sum_{k\in\mathcal{N}}[v_k^{sym,eff}]}{[n-|\mathcal{N}^f|]}$ *for all $i \in N\backslash\mathcal{N}^f$; and,*

(ii) $\boldsymbol{v}_k^{sym,eff,null}(f) = 0$ *for all $k \in \mathcal{N}^f$.*

We also notice that $\boldsymbol{v}^{sym,eff,null,add}$ is the Shapley value $\boldsymbol{\varphi}$, because of Claim 2, and we define the residual $e^{add} = \boldsymbol{v}^{sym,eff,null}(f) - \boldsymbol{\varphi}(f)$.

We establish this second decomposition theorem which allows us to test and quantify the violation of marginality through the departure from the null-player axiom.

Theorem 3. *For any given observation (θ, f), we have the unique pointwise decomposition:*

$$\theta = \boldsymbol{\varphi}(f) + e^{sym} + e^{eff} + e^{null} + e^{add}.$$

Moreover, the distance to the Shapley pay scheme can be uniquely decomposed as:

$$||e^{sh}||^2 = ||e^{sym}||^2 + ||e^{eff}||^2 + ||e^{add}||^2 + ||e^{null}||^2 + 2\langle e^{add}, e^{null}\rangle$$

into its symmetry, efficiency, null-player and additivity departures (with $||e^{mrg}||^2 = ||e^{add}||^2 + ||e^{null}||^2 + 2\langle e^{add}, e^{null}\rangle$, and $\langle e^{add}, e^{null}\rangle \neq 0$ in general). Moreover, (*i*) *If* $||e^{null}|| > 0$ *and* θ *is generated by a pay scheme* $\boldsymbol{\theta}$ *that satisfies efficiency, then the pay scheme* $\boldsymbol{\theta}$ *fails the null-player and marginality axioms;*

(*ii*) *If* $||e^{add}|| > 0$ *and* θ *is generated by a pay scheme* $\boldsymbol{\theta}$ *that satisfies symmetry, efficiency and the null-player property, then* $\boldsymbol{\theta}$ *fails additivity.*

(*iii*) *If* θ *is generated by a pay scheme* $\boldsymbol{\theta}$ *that also satisfies the null player axiom and* $||e^{add}|| > \sqrt{K(f)}$*, then* $\boldsymbol{\theta}$ *fails marginality and additivity.*

The theorems above establish tractable and easy ways to understand measures of departures from the properties of the Shapley value. More importantly, they work for limited datasets, which is a realistic situation, in the sense that the observer may not observe the behavior of a pay scheme $\boldsymbol{\theta}$ over all possible technologies.

4 Converse Implications

In this section, we provide converse implications of the results shown in Theorems 1, 2, and 3. These decomposition results are provided for limited datasets (i.e., for a given observation (θ, f)). We first consider a situation of limited datasets where we observe more than one game. In this case, we can generalize the Shapley distance to a summation that takes into account all the observations of a dataset. More specifically, we have the following:

$$||\theta - \boldsymbol{\varphi}(f)||_T^2 = \sum_{t \in \mathcal{T}} ||\theta^t - \boldsymbol{\varphi}(f^t)||^2,$$

where $\mathcal{T} = \{1, \cdots, T\}$ is an index set of observations.

We define $e^{j,t} : \mathcal{T} \mapsto \mathbb{R}^n$ for $j \in \{sh, sym, eff, mrg, null, add\}$ pointwise; $e^{j,t}$ is defined in terms of the pair (f^t, θ^t), following the prequel definitions. Abusing notation we denote $e^j(\theta^t, f^t) = e^{j,t}$. We state the following remark.

Remark 1. For a given dataset $(\theta^t, f^t)_{t \in \mathcal{T}}$, if $||e^j(\theta^t, f^t)|| > 0$ for a fixed observation (θ^t, f^t), then $||e^j||_T > 0$ for $j \in \{sh, sym, eff, mrg, null, add\}$. Moreover, Theorems 1, 2, and 3 hold for the extended data, replacing $|| \cdot ||$ by $|| \cdot ||_T$.

We are going to complement the results of Theorems 2, 3, deriving partial converse results of the "moreover statements" using the idea of extensions to full datasets. We have the following result.

Theorem 4. *For any finite set of observations* $(\theta^t, f^t)_{t \in \mathcal{T}}$*:*

(*i*) *If* $||e^{sh}||_T^2 = \sum_{t \in \mathcal{T}} ||\theta^t - \boldsymbol{\varphi}(f^t)||^2 = 0$*, then there is an extension* $\boldsymbol{\vartheta}$ *of* $(f^t, \theta^t)_{t \in \mathcal{T}}$ *to* Γ *that corresponds exactly to the Shapley payoff function (i.e.,* $\boldsymbol{\vartheta} = \boldsymbol{\varphi}$*).*

(*ii*) *If* $||e^{sym}||_T^2 = \sum_{t \in \mathcal{T}} ||e^{sym,t}||^2 = 0$*, then there is an extension* $\boldsymbol{\vartheta}$ *of* $(f^t, \theta^t)_{t \in \mathcal{T}}$ *to* Γ *that satisfies symmetry for each game* $f \in \Gamma$*.*

(iii) *If* $||e^{eff}||_T^2 = \sum_{t \in \mathcal{T}} ||e^{eff,t}||^2 = 0$*, then there is an extension* $\boldsymbol{\vartheta}$ *of* $(f^t, \theta^t)_{t \in \mathcal{T}}$ *to* Γ *that satisfies efficiency for each game* $f \in \Gamma$*.*

(iv) *If* $||e^{sym}||_T = 0$*,* $||e^{eff}||_T = 0$*, and* $||e^{null}||_T^2 = \sum_{t \in \mathcal{T}} ||e^{null,t}||^2 = 0$*, then there is an extension* $\boldsymbol{\vartheta}$ *of* $(f^t, \theta^t)_{t \in \mathcal{T}}$ *to* Γ *that satisfies efficiency and the null-player axioms for each game* $f \in \Gamma$*.*

We cannot obtain corresponding converse results for $||e^{mrg}||_T = 0$ and $||e^{add}||_T = 0$. It is easy to find counter-examples where these are zero and there is no extension that satisfies marginality or additivity (e.g., think of the value $\boldsymbol{\theta}_i(f) = \boldsymbol{\varphi}_i(f) - (-1)^i c$ for $i = 1, 2$, where $1 \sim^{sym} 2$ and $c > 0$ is a sufficiently small constant, with only one observation,).

Notice that we cannot make inference about the data generating pay scheme because with a limited dataset, even if pointwise we observe a pay profile that is numerically equivalent to the Shapley pay profile, not all possible extensions of the dataset to the domain of monotone games are equivalent to the Shapley value. However, we can ensure that at least one extension is equivalent to the Shapley value. Similarly to the statistical hypothesis testing framework, when we do not reject the null hypothesis of the consistency of the underlying data generating pay scheme with a given axiom, we cannot claim that the pay scheme will satisfy the axiom for other unobserved games (out-of-sample).

5 Application: Measuring Unfairness

In this section, we show one application of our analysis. The application is to inequality, and it answers the question of when income inequality can be considered unfair. This application is important because the Shapley value is also viewed as a way to compensate workers (Shapley and Shubik, 1967). In particular, we show how a well-known pay scheme induces a wage profile that violates the axioms characterizing the Shapley value.

We begin by generalizing the framework of a transferable-utility environment to an environment where agents have more than two options.[16] By generalizing the Shapley value to this class of environments, our work is related to recent studies including Freixas (2005), Hsiao and Raghavan (1993), Courtin et al. (2016), and Pongou and Tondji (2017). However, we have a different scope, which is to test the axioms of the Shapley value. To our knowledge, no previous study has analyzed this topic. The different options can be the numbers of worked hours a worker can supply (e.g., 0 hours, 1 hour, two hours, and so on up to a maximum number of hours). A **production environment** is modeled as a list $\mathcal{F} = (N, L, F)$ where $N = \{1, 2, ..., n\}$ is a non-empty finite set of workers of cardinality n; $L = \{0, 1, 2, ..., l\}$ is a non-empty finite

[16]This environment generalizes well-known classes of games including simple games (see, e.g., Shapley (1962), Peleg et al. (2008), and Laruelle and Valenciano (2008)), and voting games with abstention (see, e.g., Felsenthal and Machover (1997), Tchantcho et al. (2008), Freixas and Zwicker (2009), Guemmegne and Pongou (2014)). This environment also generalizes the class of bi-cooperative games introduced by Bilbao (2012). Pongou et al. (2017) use this environment to study ladder tournaments in hierarchical organizations.

set of hours of labor or effort levels that a worker can supply, where 0 denotes a situation of inaction; and F is a production function that maps each action profile $x = (x_1, ..., x_n) \in L^n$ to a real number output $F(x)$. The function F can also be interpreted as the aggregate profit or cost function. Interpreting it as the profit function might be useful in that it would be viewed as incorporating the production and the cost functions. Regardless of the interpretation adopted, we assume that $F(0, 0, ..., 0) = 0$, which means that if all the workers are inactive, there is no output.

Let $\mathcal{F} = (N, L, F)$ be a production environment and $S \in 2^N$ a set of workers. We denote by $L^{|S|}$ the set of the possible vectors of effort levels for the workers in S. An element $x \in L^{|S|}$ can be written as $x = (x_1, ..., x_s)$, where $s = |S|$ is the number of workers in S and where every $x_i \in L$ is the effort level supplied by the i^{th} worker in S.

We denote by e_i the i^{th} unit vector $(0, 0, ..., 0, 1, 0, ..., 0)$, where all the entries are zero except the i^{th} component which is one. We will also use the symbols $\trianglelefteq$ and $\triangleleft$, which we define as explained hereunder. Let $\overline{x}, x \in L^n$ be two effort profiles. We write $x \trianglelefteq \overline{x}$ to mean that $x_i \neq \overline{x}_i \Rightarrow x_i = 0$, and we write $x \triangleleft \overline{x}$ to mean that $x \trianglelefteq \overline{x}$ and $x \neq \overline{x}$. For example, $(1, 7, 5, 0, ..., 0) \triangleleft (1, 7, 5, 1, 5, 0, ..., 0)$. We denote by $|x| = |\{i \in N : x_i > 0\}|$ the number of workers who are not inactive at x. We maintain the assumption of monotonicity in the production function environment. The analogous *monotonicity property* for the production function says that $F(x) \leq F(y)$ whenever $x \trianglelefteq y$.

A pay scheme for a production environment $\mathcal{F}$ is a way to redistribute the output among the workers. Let $\mathcal{F} = (N, L, F)$ be a production environment. A pay scheme for the production $\mathcal{F}$ maps any effort profile $\overline{x} \in L^n$ to a non-null payoff profile $\boldsymbol{\theta}^F(\overline{x}) = (\boldsymbol{\theta}^F_1(\overline{x}), \boldsymbol{\theta}^F_2(\overline{x}), ..., \boldsymbol{\theta}^F_n(\overline{x}))$, where for all $i \in N$, $\boldsymbol{\theta}^F_i(\overline{x}) \in \mathbb{R}$ is interpreted as the payoff earned by i out of the output $F(\overline{x})$. In the production environment, an **observation** is a triple $(\overline{x}, F, \boldsymbol{\theta}^F(\overline{x}))$ where $\theta = \boldsymbol{\theta}^F(\overline{x})$ is an observed pay profile for any production function F and for any effort profile $\overline{x}$.

The corresponding Shapley value for the environment F, denoted by $\boldsymbol{\varphi}^F$, is given by:

$$\boldsymbol{\varphi}^F_i(\overline{x}) = \sum_{x \triangleleft\ \overline{x},\ x_i=0} \frac{(|x|)!(|\overline{x}| - |x| - 1)!}{(|\overline{x}|)!}[F(x + \overline{x}_i e_i) - F(x)],\ for\ all\ i \in N. \tag{3}$$

For simplicity, we may sometimes write $\boldsymbol{\varphi}(x, \overline{x})$ for $\frac{(|x|)!(|\overline{x}|-|x|-1)!}{(|\overline{x}|)!}$, and the marginal contribution $F(x + \overline{x}_i e_i) - F(x)$ by $mc(i, F, x, \overline{x})$.

The lemma below shows that, for a fixed level of efforts $\overline{x}$, all the information given by the production environment can be equivalently expressed using a transferable-utility game. The following lemma makes it possible to generalize all our results to the new production environment.

Lemma 6. *For any set of players N, any fixed effort profile $\overline{x}$, and any production function F, there is a transferable-utility game $(N, G^F_{\overline{x}})$ such that $G^F_{\overline{x}}(S) = F(\overline{x}_S)$ where $\overline{x}_S$ is defined as $\overline{x}_{S,i} = 0$ for all $i \in N \backslash S$ and $\overline{x}_{S,i} = \overline{x}_i$ for all $i \in S$.*

The Shapley value $\boldsymbol{\varphi}^F$ for the production function F can be equivalently defined for any fixed effort level $\overline{x} \in L^n$ as the Shapley value of the corresponding transferable-utility game $G^F_{\overline{x}}$: $\boldsymbol{\varphi}^F(\overline{x}) \equiv \boldsymbol{\varphi}^{TU}(N, G^F_{\overline{x}})$

where $\boldsymbol{\varphi}^{TU}(N, G^F_{\overline{x}})$ is the Shapley value of the game $(N, G^F_{\overline{x}})$. We now apply these notions to examine two well-known pay schemes, namely the quasi-linear contract and the linear contract. Similarly, we can use this lemma to require the axioms of symmetry, efficiency and marginality to the TU-game representation of the production environment for a fixed level of effort. We define these axioms in our firm environment in the appendix.

5.1 The Quasi-Linear Contract

Our main application is to the quasi-linear pay scheme. For simplicity, we assume two players (e.g., an employer and an employee), with each choosing his effort level from a set that contains two levels. The quasi-linearity of the pay scheme means that one player is paid a rate on the amount of input he contributes and that the other player receives the residual output. So clearly, this pay scheme is efficient.

Our objective is to measure the divergence of this pay scheme from the Shapley value and to identify the sources of this divergence. This methodology can be viewed as a way to measure unfairness in the income distribution. The example set out below shows that the quasi-linear pay scheme violates the equal-treatment and marginality axioms, and hence the additivity axiom.

Example 2. Consider a production environment $\mathcal{F} = (N, L, F)$ where $N = \{1, 2\}$ is the set of players, $L = \{0, 1\}$ is the set of effort levels, and F is the (monotone) production function defined as follows:

$$F(x) = \begin{cases} 1 & if \quad x \neq (0,0) \\ 0 & if \quad x = (0,0) \end{cases} \tag{4}$$

Consider the quasi-linear pay scheme $\boldsymbol{Qlc}$ defined as follows:

$$\boldsymbol{Qlc}_1(x) = \frac{3}{4}x_1 \ and \ \boldsymbol{Qlc}_2(x) = F(x) - \frac{3}{4}x_1, \ for \ each \ x \in L^2.$$

For each $x \in L^2$, we have $\boldsymbol{Qlc}_1(x) + \boldsymbol{Qlc}_2(x) = F(x)$, which means that $\boldsymbol{Qlc}$ is efficient.

We now show that $\boldsymbol{Qlc}$ does not satisfy the equal-treatment property. Consider the labor supply $\overline{x} = (1, 1)$. The only vector x such that $x < \overline{x}$ with $x_1 = x_2 = 0$ is $x = (0, 0)$; moreover we have $mc(1, F, x, \overline{x}) = mc(2, F, x, \overline{x}) = 1$, which shows that the two players are identical at $\overline{x} = (1, 1)$. However, $\boldsymbol{Qlc}_1(1, 1) \neq \boldsymbol{Qlc}_2(1, 1)$, and therefore, $\boldsymbol{Qlc}$ does not satisfy the equal-treatment property.

In order to quantify the violations of the properties that characterize the Shapley value, let us first derive the Shapley payoff of each player at each vector $\overline{x}$. The Shapley payoff profile at each $\overline{x}$ is given by the following matrices: $\boldsymbol{\varphi}^F(\overline{X}) = \begin{pmatrix} (0,0) & (0,1) \\ (1,0) & (\frac{1}{2}, \frac{1}{2}) \end{pmatrix}$, where $\overline{X} = \begin{pmatrix} (0,0) & (0,1) \\ (1,0) & (1,1) \end{pmatrix}$ is the matrix that contains all of the possible vectors of effort levels, with the first component of each cell denoting the effort level of player 1, and the second component denoting the effort level of player 2.

The quasi-linear payoff profile is given by: $\boldsymbol{Qlc}(\overline{X}) = \begin{pmatrix} (0,0) & (0,1) \\ (\frac{3}{4}, \frac{1}{4}) & (\frac{3}{4}, \frac{1}{4}) \end{pmatrix}$.

Using the difference between the two matrices, $\boldsymbol{\varphi}^F(\overline{X}) - \boldsymbol{Qlc}(\overline{X}) = \begin{pmatrix} (0,0) & (0,0) \\ (\frac{1}{4}, -\frac{1}{4}) & (-\frac{1}{4}, \frac{1}{4}) \end{pmatrix}$, we can compute the Shapley distance $\|\boldsymbol{\varphi}^F - \boldsymbol{Qlc}\|^2 = \begin{pmatrix} 0 & 0 \\ \frac{2}{16} & \frac{2}{16} \end{pmatrix}$.

Note that Theorem 1 applies for each fixed effort level, equivalently for each entry of the matrix $\overline{X}$. For a fixed effort $\overline{X}$, using Lemma 6, we define the game $(N, G^F_{\overline{x}})$ (for short $G^F_{\overline{x}}$). In addition we let $\theta^{\overline{x}} = \boldsymbol{Qlc}(\overline{x})$. The limited dataset is $(G^F_{\overline{x}}, \theta^{\overline{x}})$. Even if we study four different efforts, we are thinking of a situation where the observer can only analyze one effort level at a time. Thus our exercise is interested in studying each effort as a single case.

We now determine how the amount by which the violation of each property characterizing the Shapley value contributes to the total violation of 1/4. We know that:

$$\boldsymbol{Qlc}(\overline{x}) = \boldsymbol{\varphi^F}(\overline{x}) + e^{sym} + e^{eff} + e^{mrg}.$$

1. Let $e^{sym} = Qlc - v^{sym}$. For all $x \neq (1,1)$, $\boldsymbol{Qlc}(x) = \boldsymbol{v}^{sym}(x)$. For $x = (1,1)$, we have $\boldsymbol{v}_1^{sym}(x) = \boldsymbol{v}_2^{sym}(x) = \frac{1}{2}[\boldsymbol{Qlc}_1(x) + \boldsymbol{Qlc}_2(x)] = \frac{1}{2}$. Hence, the distribution $v^{sym} = \begin{pmatrix} (0,0) & (0,1) \\ (\frac{3}{4}, \frac{1}{4}) & (\frac{1}{2}, \frac{1}{2}) \end{pmatrix}$. It follows that e^{sym} can be represented by the matrix : $\begin{pmatrix} (0,0) & (0,0) \\ (0,0) & (\frac{1}{4}, -\frac{1}{4}) \end{pmatrix}$, leading to $\|e^{sym}\|^2 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{2}{16} \end{pmatrix}$.

2. Let $e^{eff} = v^{sym} - v^{sym,eff}$. We have $\boldsymbol{v}_i^{sym,eff}(x) = \boldsymbol{v}_i^{sym}(x) + \frac{F(x) - \sum_i \boldsymbol{Qlc}_i(x_i)}{2} = v_i^{sym}(x)$.

   It follows that e^{eff} can be represented by a null matrix and that $\|e^{eff}\|^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

3. Let $e^{mrg} = v^{sym,eff} - \varphi^F = \begin{pmatrix} (0,0) & (0,0) \\ (-\frac{1}{4}, \frac{1}{4}) & (0,0) \end{pmatrix}$, so $\|e^{mrg}\|^2 = \begin{pmatrix} 0 & 0 \\ \frac{2}{16} & 0 \end{pmatrix}$.

We can see that even when the total Shapley distance is the same for some effort levels, the decomposition may be completely different. The efforts $(1,1)$ and $(1,0)$ have a different decomposition: (i) the former effort is associated with an unfair quasilinear contract that is explained by a failure of symmetry (or horizontal equality); and (ii) the latter effort is associated with an unfair quasilinear contract that is explained by a failure of marginality (see below for the formal argument). In both cases the unfairness level has a very different explanation.

Now we compute the marginality bound for each effort vector.

1. For the effort vector $(1,1)$, the random value $\theta_1 = \gamma(r_{12})1 + \gamma(r_{2,1})0$ and $\theta_2 = \gamma(r_{1,2})0 + \gamma(r_{2,1})1$; the Shapley payoff profile is $(\frac{1}{2}, \frac{1}{2})$, and the symmetrized random value is $v^{sym} = \gamma(r_{12})\frac{1}{2} + \gamma(r_{2,1})\frac{1}{2}$. The bound is therefore $K^F((1,1)) = \max_{\gamma \in \Delta(R(N))} \{2(\gamma(r_{12})\frac{1}{2} + \gamma(r_{2,1})\frac{1}{2} - \frac{1}{2})^2\} = 0$, which implies that, if $||e^{mrg,F}(1,1)|| > 0$, then marginality is violated.

2. For the effort vector $(1,0)$, the random value $\theta_1 = \gamma(r_{12})1 + \gamma(r_{2,1})1$ and $\theta_2 = \gamma(r_{1,2})0 + \gamma(r_{2,1})0$; the Shapley payoff profile is $(1,0)$, and the symmetrized random value is the same. The marginality bound is given by $K^F((1,0)) = \max_{\gamma \in \Delta(R(N))} \{(0)^2 + (0)^2\} = 0$. At this effort, we reject the null hypothesis that the ***Qlc*** pay scheme fails marginality.

3. For the effort vector $(0,1)$, the bound is $K^F((0,1)) = 0$.

Finally, we can aggregate the different efforts into a consolidated Shapley distance as in section 4. In that case, we let $\|e^{sym}\|_T^2 = \|e^{mrg}\|_T^2 = \frac{2}{16}$ observe that $\|e^{mrg}\|_T^2 + \|e^{eff}\|_T^2 + \|e^{sym}\|_T^2 = \frac{1}{4}$. We conclude that 50% of the unfairness of the quasi-linear pay scheme in this example is explained by the violation of the equal-treatment property, and that the other 50% is explained by the violation of the marginality property. It is important to note that, notwithstanding the fact that the statement of the marginality axiom requires that all of the production functions be known, in this example, we were able to quantify the violation of this axiom knowing only one effort level (and a production function). This again shows the empirical relevance of our approach.

5.2 Additional Applications

We provide three additional applications of our Shapley distance decomposition in the appendix. In the second application, we analyze an arbitrary linear pay scheme in which each worker's pay is a linear function of his effort level, and study the effect of increasing the pay rate of a worker on the violation of the Shapley axioms. Our third application provides an axiomatic test of ordinary least squares (OLS), which is an estimation method of the unknown parameters in a linear regression model. We find that OLS violates all of the axioms of the Shapley value (for decomposing the goodness-of-fit following Huettner et al. (2012)), and may not be a good method for quantifying the relative importance of explanatory variables in a linear regression. Finally, we study intra-firm bargaining focusing on the effects of bargaining power on firm unfairness. For a particular example, we find that firm bargaining power monotonically increases the violations of symmetry and marginality.

6 Conclusion

In this study, we have provided a local measure of the departures of an arbitrary pay profile from the Shapley value in limited datasets. The local measure permits one to draw inference about violations of the classical axioms that characterize this value (efficiency, equal treatment of identical players, and marginality). Our measure is decomposable into these axioms. Our findings have testable implications for the different ways in which a pay scheme may violate basic properties of fairness. We provide an application to pay schemes widely used in real-life situations.

Theorem 2 shows that testing the marginality axiom in limited datasets can be difficult. In fact, we have shown that a positive distance between the symmetrized and efficient payoff and the Shapley value may not guarantee that the data generating pay scheme fails the marginality axiom. This is in part due to the fact that the traditional axioms that characterize the Shapley value, even when logically independent, are not totally independent in a more subtle sense. To see that the marginality axiom and the symmetry axiom are not completely independent, we can observe that marginality implies the null-player axiom (in the domain of monotone games). This requires that the payoff has to be equal fo all null players (and it has to be equal to zero). On the other hand, symmetry requires that all the null players have the same payoff. It then is clear that marginality and symmetry have overlapping consequences. This interdependence is one reason why testing marginality is difficult in limited datasets. To the extent that an observer or a practitioner is interested in an easy test of the violation of each axiom, this interdependence between the classical axioms can be viewed as a weakness. For this reason, a new axiomatic characterization of the Shapley value may be desirable. In results not shown here, we provide such an axiomatic characterization, also obtaining an easier test of each axiom. A distinctive feature of the current paper, however, is in providing a test of the null hypothesis for the violation of the classical marginality axiom.

References

Afriat, S. N. (1973). On a system of inequalities in demand analysis: an extension of the classical method. *International economic review*, 14(2), 460–472.

Aguiar, V. H. and Serrano, R. (2017). Slutsky matrix norms: The size, classification, and comparative statics of bounded rationality. *Journal of Economic Theory*, 172(Supplement C), 163 – 201.

Béal, S., Casajus, A., and Huettner, F. (2015). Efficient extensions of the myerson value. *Social Choice and Welfare*, 45(4), 819–827.

Bilbao, J. M. (2012). *Cooperative games on combinatorial structures*, volume 26. Springer Science and Business Media.

Bolton, G. E., Chatterjee, K., and McGinn, K. L. (2003). How communication links influence coalition bargaining: a laboratory investigation. *Management Science*, 49(5), 583–598.

Brügemann, B., Gautier, P., and Menzio, G. (2015). Bargaining, intra firm and Shapley values. *Mimeo.*

Courtin, S., Nganmeni, Z., and Tchantcho, B. (2016). The Shapley–Shubik power index for dichotomous multi-type games. *Theory and Decision*, 81(3), 413–426.

de Clippel, G. and Rozen, K. (2013). Fairness through the lens of cooperative game theory: An experimental approach. *Mimeo.*

Einy, E. and Peleg, B. (1991). Linear measures of inequality for cooperative games. *Journal of Economic Theory*, 53(2), 328–344. cited By 4.

Felsenthal, D. S. and Machover, M. (1997). Ternary voting games. *International journal of game theory*, 26(3), 335–351.

Freixas, J. (2005). The Shapley–Shubik power index for games with several levels of approval in the input and output. *Decision Support Systems*, 39(2), 185–195.

Freixas, J. (2010). On ordinal equivalence of the Shapley and Banzhaf values for cooperative games. *International Journal of Game Theory*, 39(4), 513–527.

Freixas, J. and Zwicker, W. S. (2009). Anonymous yes–no voting with abstention and multiple levels of approval. *Games and Economic Behavior*, 67(2), 428–444.

Gomez, D., González-Arangüena, E., Manuel, C., Owen, G., del Pozo, M., and Tejada, J. (2003). Centrality and power in social networks: a game theoretic approach. *Mathematical Social Sciences*, 46(1), 27–54.

Guemmegne, J. T. and Pongou, R. (2014). A policy-based rationalization of collective rules: Dimensionality, specialized houses, and decentralized authority. *Journal of Mathematical Economics*, 52, 182–193.

Hsiao, C.-R. and Raghavan, T. (1993). Shapley value for multichoice cooperative games I. *Games and economic behavior*, 5(2), 240–256.

Huettner, F., Sunder, M., et al. (2012). Axiomatic arguments for decomposing goodness of fit according to shapley and owen values. *Electronic Journal of Statistics*, 6, 1239–1250.

Kalai, E. and Samet, D. (1987). On weighted shapley values. *International Journal of Game Theory*, 16(3), 205–222.

Kalisch, G. K., Milnor, J. W., Nash, J., and Nering, E. (1954). Some Experimental n-Person Games. In *Decision Processes, edited by RM Thrall, CH Coonmbs, and RL Davis.* (pp. 301–327). New York.

Khmelnitskaya, A. B. (1999). Marginalist and efficient values for TU games. *Mathematical Social Sciences*, 38(1), 45–54.

Konow, J. (2003). Which is the fairest one of all? A positive analysis of justice theories. *Journal of Economic Literature*, 41(4), 1188–1239.

Laruelle, A. and Valenciano, F. (2008). *Voting and collective decision-making: bargaining and power.* Cambridge University Press.

Malawski, M. (2013). "procedural" values for cooperative games. *International Journal of Game Theory*, 42(1), 305–324.

Nash, J. F., Nagel, R., Ockenfels, A., and Selten, R. (2012). The agencies method for coalition formation in experimental games. *Proceedings of the National Academy of Sciences*, 109(50), 20358–20363.

Nembua, C. C. and Wendji, C. M. (2016). Ordinal equivalence of values, pigou–dalton transfers and inequality in tu-games. *Games and Economic Behavior*, 99, 117 – 133.

Peleg, B. et al. (2008). Game theoretic analysis of voting in committees. *Cambridge Books*.

Pintér, M. (2015). Young's axiomatization of the shapley value: a new proof. *Annals of Operations Research*, 235(1), 665–673.

Pongou, R., Tchantcho, B., and Tedjeugang, N. (2017). Properties of ladder tournaments. *European Journal of Operational Research*.

Pongou, R. and Tondji, J.-B. (2017). Valuing inputs under supply uncertainty: The bayesian shapley value. *Games and Economic Behavior*.

Roth, A. E. (1988). *The Shapley value: essays in honor of Lloyd S. Shapley.* Cambridge University Press.

Ruiz, L. M., Valenciano, F., and Zarzuelo, J. M. (1996). The least square prenucleolus and the least square nucleolus. two values for tu games based on the excess vector. *International Journal of Game Theory*, 25(1), 113–134.

Shapley, L. (1953). A Value for n-person Games. *Contributions to the Theory of Games*, 2, 307–317.

Shapley, L. S. (1962). Simple games: an outline of the descriptive theory. *Behavioral Science*, 7(1), 59–66.

Shapley, L. S. and Shubik, M. (1967). Ownership and the production function. *The Quarterly Journal of Economics*, (pp. 88–111).

Shubik, M. (1962). Incentives, decentralized control, the assignment of joint costs and internal pricing. *Management science*, 8(3), 325–343.

Stole, L. A. and Zwiebel, J. (1996). Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies*, 63(3), 375–410.

Tchantcho, B., Lambo, L. D., Pongou, R., and Engoulou, B. M. (2008). Voters' power in voting games with abstention: Influence relation and ordinal equivalence of power theories. *Games and Economic Behavior*, 64(1), 335–350.

van den Brink, R., Khmelnitskaya, A., and van der Laan, G. (2012). An efficient and fair solution for communication graph games. *Economics letters*, 117(3), 786–789.

Varian, H. R. (1983). Non-parametric Tests of Consumer Behaviour. *The Review of Economic Studies*, 50(1), 99–110.

Weber, R. J. (1988). Probabilistic values for games. *The Shapley Value. Essays in Honor of Lloyd S. Shapley*, (pp. 101–119).

Yaari, M. E. (1981). Rawls, Edgeworth, Shapley, Nash: Theories of distributive justice re-examined. *Journal of Economic Theory*, 24(1), 1–39.

Young, H. P. (1985). Monotonic solutions of cooperative games. *International Journal of Game Theory*, 14(2), 65–72.

7 Appendix

Proof of Lemma 1. Let f be a TU game, $i \in N$ be a null-player and $\boldsymbol{\theta}_i$ be a pay scheme satisfying the marginality axiom. Consider a coalition $C \subseteq N \backslash \{i\}$; then $mc(i, f, C) = 0$ since i is a null player. Let w be a null game (which means that the worth of any coalition is 0) (we know that $w \in \Gamma$ because w is a monotone game), then $mc(i, w, C) = 0$ for all $C \subseteq N \backslash \{i\}$ by construction. It follows that if $mc(i, f, C) = mc(i, w, C)$; then $\boldsymbol{\theta}_i(f) = \boldsymbol{\theta}_i(w)$ since $\boldsymbol{\theta}$ satisfies the marginality axiom. Now we prove that given that $w(C) \equiv 0$, for all $C \subseteq N$ we have $\boldsymbol{\theta}_i(w) = 0$ for all $i \in N$. In fact, given that, by definition 2 of a pay scheme, $\boldsymbol{\theta}_i(w) \geq 0$ is nonnegative for all $i \in N$ and is feasible, such that $\sum_{i \in N} \boldsymbol{\theta}_i(w) \leq w(N) = 0$, we have $\boldsymbol{\theta}_i(w) = 0$. Therefore, $\boldsymbol{\theta}_i(f) = 0$, and we conclude that $\boldsymbol{\theta}_i$ satisfies the null-player property.

Proof of Lemma 2. We fix f and omit it from the notation when possible. It should be clear that this optimization is pointwise. We want to solve $min_{\boldsymbol{v} \in \Theta} ||\boldsymbol{v}(f) - \theta||^2$ subject to $\boldsymbol{v}$ satisfying the equal-treatment property. This problem pointwise solution can be obtained by solving $min_{v \in \mathbb{R}^n} ||v - \theta||^2$ subject to v satisfying that $v_i = v_j$ for all $i, j \in N$ that are symmetrical at f. To define the problem in a tractable way, we denote the equivalence relation $i \sim^{sym} j$, when players i and j are identical or symmetrical in f. Notice that all players are identical to themselves. With this, we recall that we defined for any player i, the equivalence class $[i]^f = \{j \in N | j \sim^{sym} i\}$. We notice that imposing the restriction $(v_i - v_j) = 0$ for $i, j \in [i]^f$ is equivalent to the "normalized" restriction $\frac{1}{|[i]^f|}(v_i - v_j) = 0$, where $|[i]^f| \geq 1$ is the cardinality of the equivalence class.

Formally, solving the problem of interest can be formulated as:

$$min_{v \in \mathbb{R}^n} \frac{1}{2} \sum_{i \in N} (v_i - \theta_i)^2 + \sum_{i \in N} \lambda_i \sum_{j \in [i]^f} (v_i - v_j).$$

The first-order conditions (which are necessary and sufficient) are:

$$v_i - \theta_i + \lambda_i - \sum_{j \in [i]^f} \lambda_j = 0, \; for \; all \; i \in N;$$

$$(v_i - v_j) = 0 \; for \; all \; i \in N \; and \; all \; j \in [i]^f.$$

Because $v_i = v_j$ for all $i, j \in [i]^f$ we can call $v_{[i]^f} = v_i$ (without the index). Thus:

$$\sum_{j \in [i]^f} v_j = |[i]^f| v_{[i]^f},$$

and

$$\lambda_i - \lambda_j = \theta_i - \theta_j.$$

Adding up the last expression leads to:

$$|[i]^f| \lambda_i - \sum_{j \in [i]^f} \lambda_j = |[i]^f| \theta_i - \sum_{j \in [i]^f} \theta_j.$$

Then, solving for λ_i gives us:

$$\lambda_i = \theta_i - \frac{1}{|[i]^f|}\sum_{j\in[i]^f}\theta_j + \sum_{j\in[i]^f}\lambda_j.$$

We replace the last expression for λ_i in the first order conditions for θ_i: $v_i = \theta_i - \lambda_i + \sum_{j\in[i]^f}\lambda_j = 0,\ for\ all\ i \in N$; and this implies that $v_{[i]^f} = \frac{1}{|[i]^f|}\sum_{j\in[i]^f}\theta_j$.

Since, the problem has a unique solution, it follows that this is the unique solution. We conclude that the optimal solution is:

$$v_i^{sym} = v_{[i]^f} = \frac{1}{|[i]^f|}\sum_{j\in[i]^f}\theta_j.$$

This means that, the optimal solution is the average payoff of the equivalence class induced by the equivalence relation $\sim^{sym}$ of identical players.

Proof of Lemma 4. Take v^{sym} to be any symmetric pay scheme and v to be a skew-symmetric pay scheme. Notice that $< v^{sym}, v >= \sum_{i\in N} v_i^{sym} v_i$ and notice furthermore that, for singletons equivalence classes $[i]^f$ for the identical players equivalence relation (that is "unique players"), the skew symmetric pay scheme must have zero payoff, i.e $v_i = 0$ for all "unique" players. For non-unique players who are identical, say $i \sim^{sym} j$, we have that $v_i^{sym} = v_j^{sym}$ and $v_i = -v_j$, which makes $v_i^{sym}v_i + v_j^{sym}v_j = 0$. More general cases take the equivalence class $[i]^f$ and notice that:

$$< v^{sym}, v >= \sum_{j\in[i]^f} v_j^{sym} v_j = \overline{v}^{sym}\sum_{j\in[i]^f} v_j = 0,$$

where $\overline{v}^{sym}$ is a scalar value equal to the symmetric payoff given to any member of the equivalence class $[i]^f$, and because $\sum_{j\in[i]^f} v_j = 0$ by definition of skew symmetric payoff. This implies that $< v^{sym}, v >= 0$. Notice that this proof can be directly extended for the case of several equivalence classes.

Proof of Theorem 1.

Let $\varphi = \boldsymbol{\varphi}(f)$. First, we have to prove that $\theta = \varphi + e^{sym} + e^{eff} + e^{mrg}$. This is simple from the lemmas that derive the approximations and residuals v^{sym}, e^{sym} and $v^{sym,eff}, e^{eff}$; and because we notice that, $e^{mrg} = v^{sym,eff} - \varphi$ and $e^{sh} = \theta - \varphi$ leading to:

$$e^{mrg} = e^{sh} - e^{sym} - e^{eff}.$$

Now, it is necessary to obtain the decomposition. Notice that:

$$\begin{aligned} ||e^{sh}||^2 &= ||e^{sym} + e^{eff} + e^{mrg}||^2 = ||e^{sym}||^2 + ||e^{eff}||^2 + ||e^{mrg}||^2 \\ &+ 2 < e^{sym}, e^{eff} > +2 < e^{sym}, e^{mrg} > +2 < e^{eff}, e^{mrg} > . \end{aligned}$$

The proof amounts to checking that the residuals e^{sym}, e^{eff} and e^{mrg} are pairwise orthogonal.

- First we prove that $< e^{sym}, e^{eff} >= 0$.

  Notice that $e_i^{eff} = -E/n$, where $E = f(N) - \sum_{i \in N} \theta_i$ is the wasted output, and so it is a symmetric pay scheme. We also know that: $e_i^{sym} = \theta_i - \frac{1}{|[i]^f|} \sum_{i \in [i]^f} \theta_i$, where $[i]^f = \{j \in N : j \sim^{sym} i\}$ is the set of symmetric players at f, and realized effort x. Thus $\sum_{i \in [i]^f} e_i^{sym} = 0$ always, hence making e^{sym} a skew-symmetric pay scheme. By Lemma 4, we conclude that $< e^{sym}, e^{eff} >= 0$.

- Second we prove that $< e^{sym}, e^{mrg} >= 0$.

  We use the identity $e^{mrg} = \theta - \varphi - e^{sym} - e^{eff}$ and the properties of the inner product to write:

  $$< e^{mrg}, e^{sym} >=< \theta - \varphi, e^{sym} > + < -e^{sym}, e^{sym} > + < -e^{eff}, e^{sym} > .$$

  Here we notice that the third component is zero by the first step. Now, we have:

  $$< e^{mrg}, e^{sym} >=< \theta - \varphi, e^{sym} > + < -e^{sym}, e^{sym} > .$$

  Notice furthermore that the Shapley pay scheme φ either fulfills the equal-treatment property or is a symmetric pay scheme. Then,

  $$< \theta - \varphi, e^{sym} >=< \theta, e^{sym} > + < -\varphi, e^{sym} >=< \theta, e^{sym} > .$$

  Notice also that the payoff can be decomposed into its symmetric pay scheme projection and the skew symmetric residual $\theta = v^{sym} + e^{sym}$, such that:

  $$< \theta, e^{sym} >=< v^{sym}, e^{sym} > + < e^{sym}, e^{sym} >=< e^{sym}, e^{sym} > .$$

  Therefore:

  $$< e^{mrg}, e^{sym} >=< e^{sym}, e^{sym} > - < e^{sym}, e^{sym} >= 0.$$

- The third and final step consists of checking $< e^{eff}, e^{mrg} >= 0$.

  First we apply the bilinearity of the inner product to expand:

  $$< e^{mrg}, e^{eff} >=< v^{sym,eff} - \varphi, e^{eff} >= -\frac{E}{n} < \mathbf{1}, v^{sym,eff} - \varphi >= 0,$$

  where $E = f(N) - \sum_{i \in N} \theta_i$. Observe that $< \mathbf{1}, v^{sym,eff} - \varphi >= \sum_{i \in N} v_i^{sym,eff} - \sum_{i \in N} \varphi_i = 0$, because $v^{sym,eff}$ and φ are efficient.

Proof of Theorem 2.

The moreover part of the statement follows from:

(i) If the data generating pay scheme $\boldsymbol{\theta}$ satisfies the equal-treatment, efficiency, and marginality axioms, then by Claim 1, we conclude that $\boldsymbol{\theta}(f) = \theta = \boldsymbol{\varphi}(f)$ is the Shapley value at f. Thus, $||e^{sh}|| = 0$. The moreover statement (i) in the theorem follows from the contrapositive of the previous result (i.e., if $||e^{sh}|| > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ is lacking equal-treatment, efficiency or marginality).

(ii) If the data generating pay scheme $\boldsymbol{\theta}$ satisfies the equal-treatment axiom, then it follows that $v^{sym} = \theta$ (with $\theta = \boldsymbol{\theta}(f)$); thus, $||e^{sym}|| = 0$. The moreover statement (ii) in the theorem follows from the contrapositive of the previous result (i.e., if $||e^{sym}|| > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ is not symmetric).

(iii) Observe that if the data generating pay scheme $\boldsymbol{\theta}$ is efficient, then $f(N) - \sum_{i \in N} \theta_i = 0$ (with $\theta = \boldsymbol{\theta}(f)$). Thus:

$$||e^{eff}||^2 = \frac{[f(N) - \sum_{i \in N} \theta_i]^2}{n} = 0.$$

Hence, the moreover statement follows from the contrapositive of the previous result (i.e., if $||e^{eff}|| > 0$, then the data generating pay scheme $\boldsymbol{\theta}$ is not efficient).

$(iv), (v)$

(a) First we prove that if the data generating process $\boldsymbol{\theta}$ satisfies marginality this implies that $\boldsymbol{\theta}_i(f) = \boldsymbol{\phi}_i(\{mc(i, f, C)\}_{C \subseteq N \backslash \{i\}}), \forall i \in N$ for some monotone non-decreasing mapping $\boldsymbol{\phi}_i : \mathbb{R}_+^{2^{n-1}} \mapsto \mathbb{R}_+$, i.e., it is a marginalist value for all $i \in N$, where $\boldsymbol{\theta}_i : \mathbb{R}_+^{2^{n-1}} \mapsto \mathbb{R}_+$. Notice that marginality implies that for any two games f and g and for $i \in N$, such that $mc(i, f, C) = mc(i, g, C)$ $\forall C \subseteq N \backslash \{i\}$, then $\boldsymbol{\theta}_i(f) = \boldsymbol{\theta}_i(g)$; thus $\boldsymbol{\theta}_i(f)$ is a marginalist value (the monotone non-decreasing part follows from the property that if $mc(i, f, C) \geq mc(i, g, C)$ for all $C \subseteq N \backslash \{i\}$ then marginality requires that $\boldsymbol{\theta}_i(f) \geq \boldsymbol{\theta}_i(g)$ this is equivalent to saying that $\boldsymbol{\phi}_i$ is monotone non-decreasing.

(b) Second, we prove that marginality implies that $\boldsymbol{\theta}_i(f)$ is a monotone value. A monotone value is such that for any two characteristic functions f and g such that $g(S) \geq f(S)$ when $i \in S$ and $g(S) = f(S)$ when $i \notin S$, for all $S \subseteq N$; then $\boldsymbol{\theta}_i(g) \geq \boldsymbol{\theta}_i(f)$. If for two characteristic functions f and g, it holds that $g(S) \geq f(S)$ when $i \in S$ and $g(S) = f(S)$ when $i \notin S$, then $mc(i, g, C) \geq mc(i, f, C)$, for all $C \subseteq N$ such that $i \notin C$. Therefore, by marginality property, we conclude that $\boldsymbol{\theta}_i(g) \geq \boldsymbol{\theta}_i(f)$ implying that $\boldsymbol{\theta}_i$ is a monotone value.

(c) Third we recall that if the data generating pay scheme $\boldsymbol{\theta}$ satisfies marginality then it satisfies the null player property by Lemma 1.

By items (a), (b) and (c) and by Theorem 2 in Khmelnitskaya (1999), together with an efficiency of $\boldsymbol{\theta}$ and the assumption of $n \geq 3$, we conclude that $\boldsymbol{\theta}$ is a random value:

$$\boldsymbol{\theta}_i(f) = \sum_{r \in R(N)} \gamma(r) mc(i, f, r) \ for \ \gamma \in \Delta(R(N)).$$

- Now, we prove that, if the data generating pay scheme $\boldsymbol{\theta}$ is efficient and satisfies marginality, then $||e^{mrg}|| \leq \sqrt{K^{eff}(f)}$. First recall that (with $\theta = \boldsymbol{\theta}(f)$):

$$v_i^{sym} = \frac{1}{|[i]^f|} \sum_{j \in [i]^f} \theta_j,$$

for $[i]^f = \{j \in N : j \sim^{sym} i\}$ for the equivalence relation of symmetric players in f for all $i \in N$, now from the fact that it is a random value:

$$v_i^{sym} = \sum_{r \in R(N)} \gamma(r) \sum_{j \in [i]^f} mc(j, f, r).$$

By efficiency of the data generating pay scheme $\boldsymbol{\theta}$, we have the equation that $v_i^{sym,eff} = v_i^{sym}$. Now $||e^{mrg}||^2 = ||v^{sym,eff} - \varphi||^2$. Recall that the Shapley value can be written as a random value:

$$\boldsymbol{\varphi}_i(f) = \frac{1}{n!} \sum_{r \in R(N)} \gamma(r) mc(i, f, r)$$

with uniform probability. Now, the Shapley is symmetric. Thus:

$$\frac{1}{|[i]^f|} \sum_{j \in [i]^f} \boldsymbol{\varphi}_j(f) = \boldsymbol{\varphi}_i(f).$$

We then notice that:

$$\boldsymbol{\varphi}_i(f) = \frac{1}{n!} \sum_{r \in R(N)} \frac{1}{|[i]^f|} \sum_{j \in [i]^f} mc(j, f, r).$$

Given this latter equation, we conclude that:

$$||e^{mrg}||^2 = \sum_{i \in N} \left\{ \sum_{r \in R(N)} [\gamma(r) - \frac{1}{n!}] \frac{1}{|[i]^f|} \sum_{j \in [i]^f} mc(j, f, r) \right\}^2.$$

By definition of $K^{eff}(f) = max_{\rho(r) \in \Delta(R(N))} \sum_{i \in N} \sum_{r \in R(N)} ([\rho(r) - \frac{1}{n!}] \frac{1}{|[i]^f|} \sum_{j \in [i]^f} mc(j, f, r))^2$, it follows that $||e^{mrg}||^2 \leq K^{eff}(f)$ if the data generating pay scheme $\boldsymbol{\theta}$ is efficient and marginal.

By the contrapositive if $||e^{mrg}|| > \sqrt{K^{eff}(f)}$ and if $\boldsymbol{\theta}$ is efficient, then marginality must fail.

- Finally, we prove that if the data generating pay scheme $\boldsymbol{\theta}$ satisfies marginality, hence $\boldsymbol{\theta}$ is a marginalist value, then $||e^{mrg}|| \leq \sqrt{K(f)}$.

- First recall that:

$$v_i^{sym} = \frac{1}{|[i]^f|} \sum_{j \in [i]^f} \boldsymbol{\theta}_j(f),$$

for $[i]^f = \{j \in N : j \sim^{sym} i\}$ for the equivalence relation of symmetric players in f for all $i \in N$, now from the fact that it is a marginalist value:

$$v_i^{sym} = \frac{1}{|[i]^f|} \sum_{j \in [i]^f} \boldsymbol{\phi}_i(\{mc(j, f, C)\}_{C \subseteq N \setminus \{i\}}).$$

Second we recall that:

$$v_i^{sym,eff} = v_i^{sym} + \frac{1}{n}[f(N) - \sum_{i \in N} \boldsymbol{\phi}_i(\{mc(j,f,C)\}_{C \subseteq N \setminus \{i\}})].$$

Now $||e^{mrg}||^2 = ||v^{sym,eff} - \varphi||^2$. Recall that the Shapley value can be written as:

$$\boldsymbol{\varphi}_i(f) = \sum_{C \subseteq N \setminus \{i\}} \omega_C mc(i,f,C)$$

with ω_C a weight. Now, the Shapley is symmetric. Thus:

$$\frac{1}{|[i]^f|} \sum_{j \in [i]^f} \boldsymbol{\varphi}_j(f) = \boldsymbol{\varphi}_i(f).$$

Given this latter equation, we conclude that :

$$||e^{mrg}(\varphi)||^2 = \sum_{i \in N} \left\{ \begin{array}{l} \frac{1}{|[i]|} \sum_{j \in [i]} [\boldsymbol{\phi}_j(\{mc(j,f,C)\}_{C \subseteq N \setminus \{j\}}) - \sum_{C \subseteq N \setminus \{i\}} \omega_C mc(i,f,C)] \\ + \frac{1}{n}[f(N) - \sum_{i \in N} \boldsymbol{\phi}_i(\{mc(i,f,C)\}_{C \subseteq N \setminus \{i\}})] \end{array} \right\}^2.$$

By definition of $K(f) = max_{\vartheta \in \mathcal{F}} ||e^{mrg}(\vartheta)||^2$, it follows that $||e^{mrg}||^2 \leq K(f)$ if $\boldsymbol{\theta}$ is marginal. We have to prove that $K(f) = max_{\vartheta \in \mathcal{F}} ||e^{mrg}(\vartheta)||^2$ is well-defined. First, we notice that $||e^{mrg}(\vartheta)||^2$ is continuous in its argument, the reason is simple take a sequence $(\vartheta_i^k)_k$ such that $\vartheta^k = (\vartheta_i^k)_{i \in N} \in \mathcal{F}$. We notice that if we have a limit ϑ such that $\vartheta^k \to \vartheta$ then:

$$\sum_{i \in N} \left\{ \frac{1}{|[i]^f|} \sum_{j \in [i]^f} [\vartheta_j^k - \sum_{C \subseteq N \setminus \{j\}} \omega_C mc(j,f,C)] + \frac{1}{n}[f(N) - \sum_{k \in N} \vartheta_k^k)] \right\}^2 \to$$

$$\sum_{i \in N} \left\{ \frac{1}{|[i]^f|} \sum_{j \in [i]^f} [\vartheta_j - \sum_{C \subseteq N \setminus \{j\}} \omega_C mc(j,f,C)] + \frac{1}{n}[f(N) - \sum_{k \in N} \vartheta_k] \right\}^2,$$

by the properties of the limit operator and the fact that $||e^{mrg}(\vartheta)||^2$ is a composition of sums, multiplications and exponentiation, we have $||e^{mrg}(\vartheta^k)||^2 \to ||e^{mrg}(\vartheta)||^2$. Notice that given that $\theta, f, \mathcal{F}(\theta(f))$ the set is closed and bounded, note that $\mathcal{F}^{mrg}$ is closed and bounded when endowed with the uniform norm $sup_i|\vartheta_i|$. The sets $\mathcal{F}^{mrg,sym}, \mathcal{F}^{mrg,eff}, \{\varphi(f)\}$ are subsets of $\mathcal{F}^{mrg}$ thus bounded and also closed. We conclude by Weierstrass theorem $K(f) = max_{\vartheta \in \mathcal{F}(\theta(f))} ||e^{mrg}(\vartheta)||^2$ is well-defined or it exists.

By the contrapositive if $||e^{mrg}|| > \sqrt{K(f)}$, then marginality must fail.

(vi) If θ is efficient and symmetric, then $v^{sym,eff} = \theta$; therefore, by (i) and (ii), we have $||e^{sym}|| = ||e^{eff}|| = 0$ and $||e^{sh}|| = ||e^{mrg}|| > 0$ implying that $\theta \neq \varphi(f)$ is not the Shapley value. Thus, it does not satisfy marginality by the uniqueness result of Claim 1.

Proof of Lemma 5. The solution $v^{sym,eff,null}$ is obtained by solving the following optimization problem:

$$v^{sym,eff,null} = arg \min_{v \in \mathbb{R}^N} ||v^{sym,eff} - v|| \ subject \ to \ v_k = 0 \ for \ k \in \mathcal{N}^f \ and \ such \ that \sum_{i \in N} v_i = f(N).$$

This optimization problem is equivalent to what follows:

$$\min_{v \in \mathbb{R}^N} \left\{ \frac{1}{2} \sum_{i \in N} (v_i^{sym,eff} - v_i)^2 + \sum_{k \in \mathcal{N}^f} \lambda_k v_k + \nu (\sum_{i \in N} v_i - f(N)) \right\}.$$

The first-order conditions are given by:

$$v_i = v_i^{sym,eff} - \nu \ for \ i \in N \backslash \mathcal{N}^f;$$

$$v_k = v_k^{sym,eff} - \lambda_k - \nu \ for \ k \in \mathcal{N}^f;$$

$$\lambda_k = v_k^{sym,eff} - \nu \ for \ k \in \mathcal{N}^f.$$

Replacing these conditions into the constraints, we have:

$$\sum_{i \in N - \mathcal{N}^f} [v_i^{sym,eff}] - [n - |\mathcal{N}^f|]\nu = f(N),$$

$$f(x) - \sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}] - [n - |\mathcal{N}^f|]v = f(N),$$

$$\nu = -\frac{\sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}]}{[n - |\mathcal{N}^f|]},$$

$$v_i^{sym,eff,null} = v_i^{sym,eff} + \frac{\sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}]}{[n - |\mathcal{N}^f|]} \ for \ i \in N \backslash \mathcal{N}^f,$$

$$v_k^{sym,eff,null} = 0 \ for \ k \in \mathcal{N}^f.$$

Proof of Theorem 3. First, we notice that, pointwise,:

$$\theta = \varphi(f) + e^{sym} + e^{eff} + e^{null} + e^{add}.$$

In fact, due to Theorem 2, we have:

$$\begin{aligned}
\theta &= v^{sym} + e^{sym} \\
&= v^{sym,eff} + e^{eff} + e^{sym} \\
&= v^{sym,eff,null} + e^{null} + e^{eff} + e^{sym} = v^{sym,eff,null,add} + e^{add} + e^{null} + e^{eff} + e^{sym} \\
&= \varphi + e^{add} + e^{null} + e^{eff} + e^{sym}.
\end{aligned}$$

By definition, we have:

$$\begin{aligned}
||\theta - \varphi||^2 &= ||e^{sym} + e^{eff} + e^{null} + e^{add}||^2 \\
&= ||e||^2 + 2 < e^{sym}, e^{eff} > +2 < e^{sym}, e^{null} > +2 < e^{sym}, e^{add} > \\
&+ \; 2 < e^{eff}, e^{null} > +2 < e^{eff}, e^{add} > +2 < e^{null}, e^{add} > .
\end{aligned}$$

Now, we study which residuals are orthogonal among each other.

1. We already know that $< e^{sym}, e^{eff} >= 0$ by Theorem 1.

2. Notice that $< e^{sym}, e^{null} >= 0$, because e^{null} is symmetric and e^{sym} is skew symmetric by Lemma 4.

3. We also know that :

$$< e^{eff}, e^{null} >= \frac{E}{n} < 1, e^{null} >= 0, \; with \; E = -[f(N) - \sum_{i \in N} \theta_i],$$

because $< 1, e^{null} >= 0$. In fact, by definition: $e^{null} = v^{sym,eff} - v^{sym,eff,null}$, or entry-wise :

$$e_i^{null} = -\frac{\sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}]}{||[N - \mathcal{N}^f]||} = -\frac{f(N) - \sum_{i \in N - \mathcal{N}^f} [v_i^{sym,eff}]}{||[N - \mathcal{N}^f]||} \; for \; i \in N \backslash \mathcal{N}^f;$$

and $e_k^{null} = v_k^{sym,eff} \; for \; k \in \mathcal{N}^f$. Therefore,

$$\begin{aligned}
- < \mathbf{1}, e^{null} > &= -\sum e_i^{null} \\
&= f(N) - \sum_{i \in N - \mathcal{N}^f} [v_i^{sym,eff}] - \sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}] \\
&= f(N) - \sum_{k \in N} [v_k^{sym,eff}] \\
&= f(N) - f(N), \; \text{since } v^{sym,eff} \text{ is efficient} \\
&= 0.
\end{aligned}$$

4. The additivity error is given by $e^{add} = v^{sym,eff,null} - \varphi$. We have $< e^{add}, e^{sym} >= 0$, because e^{add} is symmetric and e^{sym} is skew symmetric by Lemma 4.

5. We show that $< e^{add}, e^{eff} >= 0$. Indeed:

$$\begin{aligned}
< e^{add}, e^{eff} > &= < v^{sym,eff,null}, e^{eff} > - < \varphi, e^{eff} > \\
&= < v^{sym,eff,null}, \mathbf{1} > \frac{E}{n} - < \varphi, \mathbf{1} > \frac{E}{n} \\
&= \frac{f(N)E}{n} - \frac{f(N)E}{n} \\
&= 0.
\end{aligned}$$

6. The remaining term $< e^{add}, e^{null} >$ is in general non-zero. Indeed, we have what follows:

$$< \varphi, e^{null} >= -[\frac{\sum_{i \in \mathcal{N}^f} [v_i^{sym,eff}]}{||[N - \mathcal{N}^f]||}] \sum_{i \in N - \mathcal{N}^f} \varphi_i + \sum_{k \in \mathcal{N}^f} v_k^{sym,eff} \varphi_k.$$

We can decompose $< v^{sym,eff,null}, e^{null} >$ as:

$$\begin{aligned} < v^{sym,eff,null}, e^{null} > &= < v^{sym,eff}, e^{null} > + < e^{null}, e^{null} > \\ &= < v^{sym,eff}, e^{null} > + \frac{[\sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}]]^2}{[N - \mathcal{N}^f]} + \sum_{k \in \mathcal{N}^f} (v_k^{sym,eff})^2. \end{aligned}$$

In the same manner, we can rewrite $< v^{sym,eff,null}, e^{null} >$ as:

$$\begin{aligned} < v^{sym,eff,null}, e^{null} > &= -[\frac{f(N) - \sum_{i \in N - \mathcal{N}} [v_i^{sym,eff}]}{[N - \mathcal{N}^f]}] \sum_{i \in N - \mathcal{N}^f} v_i^{sym,eff} \\ &+ \sum_{k \in \mathcal{N}^f} (v_k^{sym,eff})^2 + \frac{[\sum_{k \in \mathcal{N}^f} [v_k^{sym,eff}]]^2}{[N - \mathcal{N}^f]} + \sum_{k \in \mathcal{N}^f} (v_k^{sym,eff})^2. \end{aligned}$$

Given the equation that,

$$< v^{sym,eff,null}, e^{null} >= -\frac{f(x) \sum_{i \in N - \mathcal{N}^f} v_i^{sym,eff}}{||[N - \mathcal{N}^f]||} + 2 \sum_{k \in \mathcal{N}^f} (v_k^{sym,eff})^2,$$

it follows that $< e^{add}, e^{null} > \neq 0$.

The conclusion is the following:

$$\begin{aligned} ||\theta - \varphi||^2 > &= ||e^{sym} + e^{eff} + e^{null} + e^{add}||^2 \\ &= ||e^{sym}||^2 + ||e^{eff}||^2 + ||e^{add}||^2 + ||e^{null}||^2 + 2\langle e^{add}, e^{null} \rangle, \end{aligned}$$

and $||e^{mrg}||^2 = ||e^{null}||^2 + ||e^{add}||^2 + 2 < e^{null}, e^{add} >$.

The moreover part of the statement now is established: We use $\boldsymbol{\theta}$ to denote the data generating pay scheme of (f, θ).

(i) If $\boldsymbol{\theta}$ satisfies the null-player and efficiency properties or if $\theta_k = 0$ for all $k \in \mathcal{N}^f$ null players in f, then $v_k^{sym} = 0$ for all $k \in \mathcal{N}^f$ because the null players are symmetric among each other. Under efficiency requirement, we have $v^{sym,eff} = v^{sym}$. Therefore $v^{sym,eff}$ already satisfies the null-player property. Hence $v^{sym,eff,null} = v^{sym,eff}$. If $\boldsymbol{\theta}$ satisfies the null-player property and efficiency, then $||e^{null}|| = ||v^{sym,eff} - v^{sym,eff,null}|| = 0$. Moreover if $||e^{null}|| > 0$, then $\boldsymbol{\theta}$ fails the null-player property and as a consequence marginality property must fail by Lemma 1.

(ii) If $\boldsymbol{\theta}$ is additive, symmetric, and efficient and satisfies the null-player properties, then $\theta = \varphi(f)$ is the Shapley payoff function; then $||e^{sh}|| = 0$. Moreover, if $||e^{add}|| > 0$, then $\boldsymbol{\theta}$ fails at least one of the axioms: additivity, symmetry, efficiency or the satisfaction of the null-player property. If any error component is zero and $||e^i|| = 0$ for some $i \in \{sym, eff, null\}$, then we know that $\boldsymbol{\theta}$ fails either additivity or $\{sym, eff, null\}\backslash\{i\}$. In particular, we know with certainty that, if $\boldsymbol{\theta}$ does not fail all $i \in \{sym, eff, null\}$, then $||e^{add}|| > 0$ implies that $\boldsymbol{\theta}$ fails additivity.

(iii) This statement holds because, under marginality and efficiency of $\boldsymbol{\theta}$, $e^{null} = 0$. Thus $||e^{mrg}|| = ||e^{add}||$. We conclude that, if the null-player property holds in $\boldsymbol{\theta}$ and is efficient for $n \geq 3$, then if $||e^{add}|| > K^{eff}(f)$, then we must have a violation of marginality and additivity properties.

Proof of Theorem 4.

(i) If $||e^{sh}||_T = 0$, then $\boldsymbol{\theta}(f) = \boldsymbol{\varphi}(f)$ for all $f \in \Gamma$. Thus, we build $\boldsymbol{\vartheta} = \boldsymbol{\varphi}$ which is an extension of the data.

(ii) If $||e^{sym}||_T = 0$, then $\boldsymbol{\theta}(f^t) = \boldsymbol{v}^{sym}(f^t)$; thus it is symmetric for each $t \in \mathcal{T}$. We build $\boldsymbol{\vartheta}(g) = \boldsymbol{\varphi}(g)$ for $g \in \Gamma\backslash(f^t)_{t\in\mathcal{T}}$ and $\boldsymbol{\vartheta}(f^t) = \theta^t$ for $t \in \mathcal{T}$; this is symmetric for each $f \in \Gamma$. An alternative construction that is continuous, can be build as follows: First use $(f^t, \theta^t)_{t\in\mathcal{T}}$ as nodes for interpolation so to build a continuous mapping $\boldsymbol{\eta} : \Gamma \to \mathbf{R}^n_+$ such that $\boldsymbol{\eta}(f^t) = \theta^t$ (with finite nodes this continuous mapping always exists).[17] Then define $\boldsymbol{\vartheta}(f) = \boldsymbol{\eta}^{sym}(f)$ for all $f \in \Gamma$ where $\boldsymbol{\eta}^{sym}(f)$ is the symmetrized pay scheme $\boldsymbol{\eta}$. Notice that this guarantees that $\boldsymbol{\vartheta}(f^t) = \theta^t$ for all $t \in \mathcal{T}$ and at the same time $\boldsymbol{\vartheta}$ satisfies symmetry for all $f \in \Gamma$.

(iii) If $||e^{eff}||_T = 0$, then $v^{sym,t} = v^{sym,eff,t}$ for all $t \in \mathcal{T}$ (with $v^{sym,t}$ the symmetric part of θ^t and $v^{sym,eff,t}$ is the symmetric and efficient part of $v^{sym,t}$) and $\sum\limits_{i\in N} v_i^{sym,t} = \sum\limits_{i\in N} \theta_i^t = f(N)$; thus $\boldsymbol{\theta}(f^t)$ is efficient for each $t \in \mathcal{T}$. We build $\boldsymbol{\vartheta}(g) = \boldsymbol{\varphi}(g)$ for $\in \Gamma\backslash(f^t)_{t\in\mathcal{T}}$ and $\boldsymbol{\vartheta}(f^t) = \theta^t$ for $t \in \mathcal{T}$; this is efficient for each $f \in \Gamma$. (A continuous construction can be done in an analogous way to (ii).)

(iv) If $||e^{sym}||_T = 0$, $||e^{eff}||_T = 0$, $||e^{null}||_T = 0$, then $\theta^t = v^{sym,eff,null,t}$ for all $t \in \mathcal{T}$; thus, θ^t is symmetric and, efficient and has the null-player property, since $v_k^{sym,eff,null,t} = 0$ for all $k \in \mathcal{N}^{f^t}$ null-players in f^t at observation t. We build $\boldsymbol{\vartheta}(g) = \boldsymbol{\varphi}(g)$ for $g \in \Gamma\backslash(f^t)_{t\in\mathcal{T}}$ and $\boldsymbol{\vartheta}(f^t) = \theta^t$ for $t \in \mathcal{T}$; this is symmetric and, efficient and has the null-player property for each $f \in \Gamma$.(A continuous construction can be done in an analogous way to (ii).)

Proof of Lemma 6.

Let $\mathcal{F} = (N, L, F)$ be any firm. If we fix $\overline{x} \in L^n$, we can build a transferable-utility game for the fixed effort $G^F_{\overline{x}} : 2^N \mapsto \mathbb{R}$ as follows: First define the mapping $c : L \to 2^N$ as $c(\overline{x}) = \{i \in N | x_i > 0\}$. The

[17]By continuity of $\boldsymbol{\eta}$ we mean that if we take a sequence of games (f^n) such that $f^n \in \Gamma$ for $n \geq 1$ and $f^n \Rightarrow f$, then $\boldsymbol{\eta}(f) = Lim_{n\Rightarrow\infty}\boldsymbol{\eta}(f^n)$.

mapping c takes as an input a fixed effort and maps it to a coalition or a subset of N, by including a player only if the player is providing positive effort under $\overline{x}$. Now define the characteristic function $G^F_{\overline{x}} : 2^N \to \mathbb{R}$, $\mathrm{G}^F_{\overline{x}}(S) = G^F_{\overline{x}}(c(\overline{x}_S)) = F(\overline{x}_S)$, where $\overline{x}_S$ is defined as $\overline{x}_{S,j} = 0$ for all $j \in N \backslash S$ and $\overline{x}_{S,i} = \overline{x}_i$ for all $i \in S$.

We check that $(N, G^F_{\overline{x}})$ is a game. To do this it suffices to check that $G^F_{\overline{x}}$ is a characteristic function. We observe that $G^F_{\overline{x}}(\emptyset) = F(0) = 0$ by assumption and, under the assumption of limited datasets, we observe all $\overline{x}_S$ for a given $\overline{x}$ and for any $S \subseteq N$. Thus $G^F_{\overline{x}}$ is a characteristic function. Also it is easy to verify that monotonicity of F implies monotonicity of $G^F_{\overline{x}}$. In fact, if $x \trianglelefteq y$ then $F(x) \leq F(y)$, by definition of $\trianglelefteq$ we know that if $x \trianglelefteq y$ then $c(x) \subseteq c(y)$, which means that $G^F_{\overline{x}}(c(x)) \leq G^F_{\overline{x}}(c(y))$.

7.1 Axioms for the Production Environment

Here we redefine in the language of the production environment the main axioms and concepts that we are interested in our main results.

Definition 12. Let $i, j \in N$ be two workers, $\overline{x}$ be an effort profile, and F be a production function.

1. Worker i is a *null-worker at* $(\overline{x}, F)$ if for any $x \in T^n$ such that $x < \overline{x}$ and $x_i = 0$, $mc(i, F, x, \overline{x}) = 0$.

2. Workers i and j are said to be symmetrical or identical *at* $(\overline{x}, F)$ if for all $x \in T^n$ such that $x < \overline{x}$ and $x_i = x_j = 0$, $mc(i, F, x, \overline{x}) = mc(j, F, x, \overline{x})$.

We now define the axioms.

Axiom 6. ***(Symmetry for the Production Environment)***
A pay scheme $\boldsymbol{\theta}$ satisfies symmetry if for any workers i and j that are symmetrical at $(\overline{x}, F)$, $\boldsymbol{\theta}^F_i(\overline{x}) = \boldsymbol{\theta}^f_j(\overline{x})$.

Axiom 7. ***(Efficiency for the Production Environment)***
A pay scheme $\boldsymbol{\theta}$ is efficient if at any $(\overline{x}, F)$, it must be that $\sum_{i \in N} \boldsymbol{\theta}^F_i(\overline{x}) = F(\overline{x})$.

Axiom 8. ***(Marginality for the Production Environment)***
A pay scheme θ is marginal if for any production functions F and G, any worker $i \in N$ and a given effort profile $\overline{x}$ and x such that $x < \overline{x}$ with $x_i = 0$, $[F(x + \overline{x}_i e_i) - F(x) \geq G(x + \overline{x}_i e_i) - G(x)] \Rightarrow [\boldsymbol{\theta}^F_i(\overline{x}) \geq \boldsymbol{\theta}^G_i(\overline{x})]$.

Axiom 9. ***(Null worker property for the Production Environment)***
A pay scheme $\boldsymbol{\theta}$ satisfies the property of null-worker if for any production function F, a given effort profile $\overline{x} \in T^n$, and any null-worker $i \in N$ at $(\overline{x}, F)$, $\boldsymbol{\theta}^F_i(\overline{x}) = 0$.

Axiom 10. ***(Additivity for Production Environment)***
A pay scheme $\boldsymbol{\theta}$ is additive if for any production functions F and G and any given effort profile $\overline{x} \in T^n$, $\boldsymbol{\theta}^{F+G}(\overline{x}) = \boldsymbol{\theta}^F(\overline{x}) + \boldsymbol{\theta}^G(\overline{x})$.

7.2 Additional Applications

We provide three additional applications of our Shapley distance decomposition in the appendix. In the second application, we analyze an arbitrary linear pay scheme in which each worker's pay is a linear function of his effort level, and study the effect of increasing the pay rate of a worker on the violation of the Shapley fairness axioms. Our third application provides an axiomatic test of ordinary least squares (OLS), which is an estimation method of the unknown parameters in a linear regression model. We find that OLS violates all of the axioms of the Shapley value, and that it is not a good method for quantifying the relative importance of explanatory variables in a linear regression. Finally, we study intra-firm bargaining in the spirit of Stole and Zwiebel (1996), focusing on the effects of bargaining power on firm unfairness. For a particular example, we find that firm bargaining power monotonically increases the violations of symmetry and marginality.

7.2.1 The Linear Contract: A Comparative Statics Analysis

In this second application, we analyze an arbitrary linear pay scheme in which each worker's pay is a linear function of his effort level, and study the effect of increasing the pay rate of a worker on the violation of the Shapley fairness axioms.

Consider a production environment $\mathcal{F} = (N, L, F)$ and an effort profile $\overline{x}$. The payoff of each worker i at $\overline{x}$ is $v_i^{lc} = \alpha_i \overline{x}_i$, where $\alpha_i > 0$ is the pay rate of i. The closest pay scheme that is symmetric is given by:

$$v_i^{sym} = \frac{1}{|[i]|} \sum_{j \in [i]} \alpha_j \overline{x}_j$$

for all $j \in [i]$ in the equivalence class of workers.

The closest pay scheme that is both symmetric and efficient is given by:

$$v_i^{sym,eff} = v_i^{sym} + \frac{1}{n}[F(\overline{x}) - \sum_{i \in N} \alpha_i \overline{x}_i].$$

Finally the pay scheme that is symmetric and efficient and that satisfies the marginality axiom is evidently the corresponding Shapley value of the firm given by $\boldsymbol{\varphi}_i^F(\overline{x})$.

The residuals are computed as follows:

$$e_i^{sym} = \alpha_i \overline{x}_i - \frac{1}{|[i]|} \sum_{j \in [i]} \alpha_j \overline{x}_j;$$

$$e_i^{eff} = -\frac{1}{n}[F(\overline{x}) - \sum_{i \in N} \alpha_i \overline{x}_i].$$

Finally, the marginality residual is $e^{mrg} = v^{sym,eff} - \varphi^F$.

Intuitively, observe that the marginality residual is the weighted average of the difference between the corrected linear pay scheme and the marginal contribution under the firm's different configurations. The total residual is a weighted average of the difference between the linear pay scheme and the marginal contribution:

$$e_i^{sh} = \sum_{x \lhd \overline{x},\ x_i=0} \frac{(|x|)!(|\overline{x}|-|x|-1)!}{(|\overline{x}|)!}[\alpha_i \overline{x}_i - (F(x+\overline{x}_i e_i) - F(x))].$$

The distance of the linear pay scheme to the Shapley payoff at $\overline{x}$ is therefore:

$$||e^{sh}(\alpha)||^2 = \sum_{i\in N}\{ \sum_{x \lhd \overline{x},\ x_i=0} \frac{(|x|)!(|\overline{x}|-|x|-1)!}{(|\overline{x}|)!}[\alpha_i \overline{x}_i - (F(x+\overline{x}_i e_i) - F(x))]\}^2,$$

which is a function of the vector $(\alpha_i)_i$ of pay rates.

We now analyze the effect of increasing a worker i's pay rate α_i on the Shapley distance. We have:

$$\frac{\partial}{\partial \alpha_i}||e^{sh}||^2(\alpha) = 2e_i^{lc}\overline{x}_i.$$

This shows that the sign of the effect of a change in α_i entirely depends on the sign of e_i^{sh}. Furthermore, the magnitude of this effect depends on the effort level $\overline{x}_i$ and the residual e_i^{sh}. A necessary and sufficient condition for the residual e_i^{sh} to be positive is when the linear payoff that worker i is receiving is greater than what the worker would have received under the Shapley payoff: $\alpha_i\overline{x}_i > \varphi_i^F(\overline{x})$. Therefore, increasing the effort unit rate α_i increases the level of unfairness only if worker i is getting more than his fair pay.

We now determine how each component of the distance between the linear pay scheme and the Shapley payoff at $\overline{x}$ is affected by a change in α_i.

First of all, notice that the violation of the equal-treatment axiom is the variance of the average pay of symmetric workers:

$$||e^{sym}(\alpha)||^2 = \sum_{i\in N}[\alpha_i\overline{x}_i - \frac{1}{|[i]^f|}\sum_{j\in[i]^f}\alpha_j\overline{x}_j]^2.$$

The derivative of this measure with respect to α_i is:

$$\frac{\partial}{\partial \alpha_i}||e^{sym}(\alpha)||^2 = 2e_i^{sym}\frac{|[i]^f|-1}{|[i]^f|}\overline{x}_i - \frac{1}{|[i]^f|}\sum_{j\in[i],j\neq i} 2e_j^{sym}\overline{x}_i.$$

We note that the latter derivative depends on two components. One component is the additional lack of the equal-treatment property of worker i which is positive when $v_i^{lc} > v_i^{sym}$ (that is, when worker i receives under the linear pay scheme a payoff greater than the average payoff of the group of symmetric workers to which i belongs). The second component measures discrimination due to the payoff of the other workers symmetric to i which is smaller than the average: $v_j^{lc} < v_j^{sym}$ for $j \neq i$. It is clear, that an increase in α_i has a direct effect and an externality effect that depend on the relative position of the people within the group of workers who are symmetric to i.

The violation of efficiency is simply the square of the wasted output divided by the number of workers:

$$||e^{eff}(\alpha)||^2 = \sum_{i\in N}[\frac{1}{n}[\sum_{i\in N}\alpha_i\overline{x}_i - F(\overline{x})]]^2 = \frac{[\sum\limits_{i\in N}\alpha_i\overline{x}_i - F(\overline{x})]^2}{n}.$$

The effect of increasing the pay rate α_i of worker i on the efficiency violation is:

$$\frac{\partial}{\partial \alpha_i}||e^{eff}(\alpha)||^2 = \frac{2[\sum\limits_{i\in N}\alpha_i\overline{x}_i - F(\overline{x})]}{n}\overline{x}_i.$$

This effect is always nonpositive due to the fact that $\sum_{i \in N} \alpha_i \overline{x}_i \leq F(\overline{x})$. It follows that increasing a worker's pay rate always increases efficiency. Together with the findings on the effect of increasing a worker's pay rate on the symmetry violation, this finding suggests that the linear pay scheme trades off horizontal fairness and efficiency under certain configurations.

The marginality violation is equal to:

$$||e^{mrg}(\alpha)||^2 = \sum_{i \in N} \left(\sum_{x \triangleleft \overline{x},\ x_i = 0} \varphi(x, \overline{x}) \left(\frac{1}{|[i]|} \sum_{j \in [i]} \alpha_j \overline{x}_j + \frac{1}{n}[F(\overline{x}) - \sum_{i \in N} \alpha_i \overline{x}_i] - mc(i, F, x, \overline{x}) \right) \right)^2 .$$

Taking the derivative of $||e^{mrg}(\alpha)||^2$ with respect to α_i yields:

$$\frac{\partial}{\partial \alpha_i} ||e^{mrg}(\alpha)||^2 = 2e_i^{mrg}(\tfrac{1}{|[i]|}\overline{x}_i - \tfrac{1}{n}\overline{x}_i) + \sum_{j \in [i]^f, j \neq i} 2e_j^{mrg}(\tfrac{1}{|[i]^f|}\overline{x}_i - \tfrac{1}{n}\overline{x}_i) + \sum_{k \in N, k \notin [i]} e_k^{mrg}(-\tfrac{1}{n}\overline{x}_i).$$

A sufficient condition for this derivative to be positive in the first two components is that the symmetry and efficiency payoffs are greater for i and for $j \in [i]$ than their fair share: $v_i^{sym,eff} > \varphi_i^F(\overline{x})$ for $i \in [i]$; this means that increasing the effort rate of worker i increases unfairness. The final component is positive if the workers outside the equivalence class of worker i have symmetry and efficiency payoffs that are below their fair payoffs, that is, $v_k^{sym,eff} < \varphi_k^F(\overline{x})$.

In summary, increasing a worker i's payoff increases the violation of marginality when the worker himself or workers who are symmetric to him are receiving more than they should receive under the Shapley pay scheme and when other workers who are different from i receive less than their Shapley wage.

By a simple rule of derivation, we note that the total effect of a change in the effort rate α_i is also additively decomposable into the terms that we have presented:

$$\frac{\partial}{\partial \alpha_i} ||e^{lc}(\alpha)||^2 = \frac{\partial}{\partial \alpha_i} ||e^{sym}(\alpha)||^2 + \frac{\partial}{\partial \alpha_i} ||e^{eff}(\alpha)||^2 + \frac{\partial}{\partial \alpha_i} ||e^{mrg}(\alpha)||^2.$$

7.2.2 Ordinary Least Squares: Shapley Relative Importance of Explanatory Variables for the Goodness-of-Fit

In this application, we provide an axiomatic test of ordinary least squares, viewed as a method for estimating the relative importance of a finite set of input variables in the production of an output variable. Consider a scalar dependent random variable y, a set of independent random variables K, and an unobserved random scalar variable ϵ. The dependent variable is related to the other variables by the following linear equation:

$$y = \beta_0 + \sum_{x_j \in K} \beta_j x_j + \epsilon.$$

The statistician is usually interested in the total explanatory power of the variables in K, captured by the goodness-of-fit R^2.

The (population) $R^2(K) = \frac{SS_R}{SS_T} = \frac{Var(\hat{y})}{Var(y)}$, where $\hat{y} = \beta_0 + \sum_{x_j \in K} \beta_j x_j$. We are interested in finding an assignment of the goodness-of-fit index among the independent variables in K.

The traditional approach to this problem is to use the standardized ordinary least squares (OLS) coefficients to measure the relative importance of each $x_j \in K$. Let $\sigma_{x_j} = \sqrt{Var(x_j)}$ be the squared root of the variance of x_j, $\sigma_y = \sqrt{Var(y)}$, and $\sigma_{x_j x_k} = Cov(x_j, x_k)$. We define the standardized β coefficient by:

$$\overline{\beta}_{x_j} = \beta_j \frac{\sigma_{x_j}}{\sigma_y}.$$

The quantity $\overline{\beta}^2_{x_j}$ is a measure of the relative importance of variable $x_j \in K$ in the goodness-of-fit $R^2(K)$. The $R^2(K)$ is given by:

$$R^2(K) = \sum_{x_j \in K} \overline{\beta}^2_{x_j} + \sum_{x_i, x_k \in K, x_i \neq x_k} 2\overline{\beta}_{x_i} \overline{\beta}_{x_k} \sigma_{x_i x_k}.$$

For simplicity, assume that the independent variables are pairwise independent (or non-redundant). Thus:

$$R^2(K) = \sum_{x_j \in K} \overline{\beta}^2_{x_j}.$$

We now present the Shapley assignment of the goodness-of-fit index to each of the independent variables in K. We want to quantify the departures from the axioms of symmetry and marginality. The axiom of marginality here is very natural as we want to assign a higher relative importance to variables that have a higher marginal contribution to the goodness-of-fit. The equal treatment requires that if two variables are perfect substitutes in terms of marginal explanatory power, they should receive the same weight. This is a natural requirement if the practitioner is trying to not introduce her subjective beliefs about the relative importance of variables in the measure. Under the independence assumption we have made, efficiency is always satisfied.

To define the Shapley assignment of $R^2(K)$, we define the constrained model:

$$y = \beta_0 + \sum_{x_j \in T} \beta_j x_j + \epsilon,$$

for $T \subseteq K$ (we abuse notation and keep the unobserved random variable ϵ the same across the different models).

Each of these models has an associated goodness-of-fit index $R^2(T)$. We identify the characteristic function $f : 2^K \mapsto \mathbb{R}$, $T \mapsto f(T)$ for all $T \subseteq K$, such that $f(T)$ is numerically equivalent to $R^2(T)$. (Note that the R^2 is monotone in the number of random explanatory variables. Hence our results apply.)

The Shapley assignment is given by:

$$\varphi_{x_j}(f) = \frac{1}{|K|!} \sum_{r \in R(K)} mc(x_j, f, r), \quad mc(x_j, f, r) = f(T(r^{x_j}) \cup \{x_j\}) - f(T(r^{x_j})),$$

where $T(r^{x_j}) = \{x_k \in K | x_k r x_j\}$ where $r \in R(K)$ is a linear order of introduction in the restricted model of the independent variables in K.

The Shapley distance is given by:

$$||e^{sh}||^2 = \sum_{x_j \in K} (\varphi_{x_j}(f) - \overline{\beta}^2_{x_j})^2.$$

The result of finding the symmetrized assignment is $v^{sym}_{x_j} = \frac{1}{|[x_k]|} \sum_{x_k \in [x_j]} \overline{\beta}^2_{x_j}$ for the equivalence class $[x_j]$ of variables that are symmetric in their marginal contributions. The first component of the orthogonal decomposition is:

$$||e^{sym}||^2 = \sum_{x_j \in K} (\frac{1}{|[x_k]|} \sum_{x_k \in [x_j]} \overline{\beta}^2_{x_j} - \overline{\beta}^2_{x_j})^2.$$

This quantity is the within-variance of the squared of the standardized β coefficients among those variables that are perfect substitutes according to their marginal contributions to the goodness-of-fit. This quantity tells us the extend to which the "label" of the variable matters. To fix ideas, we assume $|K| = 2$ and consider that both variables are symmetric (i.e., $x_1 \sim x_2$) and independent. Also let $\sigma_{x_j} = 1$, so that $\overline{\beta}_{x_j} = \frac{cov(y,x_j)}{\sigma_y}$ for $j = 1, 2$. Then:

$$||e^{sym}||^2 = \frac{1}{2}(\frac{cov(y,x_1)^2 - cov(y,x_2)^2}{\sigma^2_y})^2$$

is roughly the difference between the squared covariances of the independent variables and y. This means that in most cases, the equal treatment axiom is going to fail as two covariances are rarely the same.

The second component is zero because efficiency holds under the independence assumption. The third component is the difference between the symmetrized assignment and the Shapley assignment:

$$||e^{mrg}||^2 = \sum_{x_j \in K} (\frac{1}{|[x_k]|} \sum_{x_k \in [x_j]} \overline{\beta}^2_{x_j} - \varphi_{x_j})^2.$$

With $|K| = 2$ independent variables that are symmetric, we have:

$$||e^{mrg}||^2 = \frac{1}{4}(\frac{cov(y,x_1)^2 + cov(y,x_2)^2}{\sigma^2_y} - R^2(K))^2 = 0.$$

We cannot reject marginality in this case.

If we relax the symmetry assumption such that x_1 and x_2 are not symmetric, then $||e^{sym}||^2 = 0$, and we find a marginality violation:

$$||e^{mrg}||^2 = \frac{1}{2}(\frac{cov(y,x_1)^2 - cov(y,x_2)^2}{\sigma^2_y})^2 > 0,$$

where $R^2(K)$ is the goodness-of-fit of the model. The Shapley assignment, under the symmetry assumption, is just the equal split of this index. We can observe now that marginality is violated. In fact, the random assignments weighted by $\gamma \in \Delta(R)$, for $\theta_{x_1} = \gamma(r_{12})(R^2(\{x_1\})) + \gamma(r_{2,1})(R^2(K) - R^2(\{x_1\}))$ and $\theta_{x_2} = \gamma(r_{1,2})(R^2(K) - R^2(\{x_2\})) + \gamma(r_{2,1})(R^2(\{x_2\}))$. The Shapley value is: $\varphi(\{x_i\}) = \frac{1}{2}R^2(K)$.

This means that the assignment $\theta_{x_i} = \frac{1}{2}\frac{cov(y,x_i)^2}{\sigma_y^2}$ cannot be extended to an assignment that satisfies marginality (under the asymmetry assumption between x_1 and x_2).

We conclude that, in general, the standardized ordinary least squares violates symmetry and marginality, and the magnitude of these violations for a fixed model are determined by the level of substitutability among the different independent variables.

7.2.3 Intra-firm Bargaining and Firm Unfairness

We consider an at-will firm with two identical players. The firm and the players bargain over payoffs using a Rodolex procedure in the spirit of Stole and Zwiebel (1996). The bargaining is done bilaterally and $p \in [0,1]$ is the probability that if the firm rejects an offer from a player, the negotiation breaks down and $1-p$ is the probability that the negotiation proceeds to the next stage. Thus, p is a proxy for the firm's negotiating power. When $p \to 0$, the firm has no negotiating power at all. We let the firm be indexed as 0 and the players be indexed as $1, 2$. We spare the details of the bargaining protocol and redirect the reader to the work of Brugemann et al. (2015).

In this example, the negotiation finishes with the following pay vector:

$$w_1(p) = b + \frac{1}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b],$$

$$w_2(p) = b + \frac{1-p}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b],$$

and with the firm receiving:

$$w_0(p) = \pi(p) = \pi_0(p) + \frac{(1-p)^2}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b].$$

With $\pi_0(p) = y_0 + \frac{1-p}{2-p}[y_1 - y_0 - b]$, b is the outside option of both players and, y_i the production of the firm with $i \in \{0, 1, 2\}$ players being active.

The result of letting $p \to 0$ is that the pay scheme converges to the Shapley value. In particular:

$$\varphi_1 = w_1(0) = b + \frac{1}{3}[y_2 - \pi_0(0) - 2b],$$

$$\varphi_2 = w_2(0) = b + \frac{1}{3}[y_2 - \pi_0(0) - 2b],$$

and

$$\varphi_0 = w_0(0) = \pi(0) = \pi_0(0) + \frac{1}{3}[y_2 - \pi_0(0) - 2b].$$

We observe that players are symmetric and different from the firm in general. We compute:

$$v_0^{sym}(p) = w_0(p) = \pi_0 + \frac{(1-p)^2}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b],$$

and:

$$v_1^{sym}(p) = v_2^{sym}(p) = \frac{1}{2}[w_1(p) + w_2(p)] = b + \frac{(1 - \frac{1}{2}p)}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b].$$

The corresponding error terms are:

$$e_1^{sym}(p) = \frac{\frac{1}{2}p}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b],$$

and

$$e_2^{sym}(p) = \frac{-\frac{1}{2}p}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b],$$

(with $e_0^{sym}(p) = 0$).

The error of marginality is :

$$e_1^{mrg}(p) = v_1^{sym}(p) - \varphi = [\frac{p(\frac{1}{2} - \frac{1}{3}p)}{(1 + (1-p) + (1-p)^2)}][y_2 - 2b] + \frac{1}{3}\pi_0(0) - \frac{(1 - \frac{1}{2}p)}{1 + (1-p) + (1-p)^2}\pi_0(p),$$

and

$$e_2^{mrg}(p) = e_1^{mrg}(p).$$

Note that:

$$e_0^{mrg}(p) = \pi_0(p) - \pi_0(0) + \frac{(1-p)^2}{1 + (1-p) + (1-p)^2}[y_2 - \pi_0(p) - 2b] - \frac{1}{3}[y_2 - \pi_0(0) - 2b].$$

Without loss of generality, we fix $y_1 - y_0 - b = 0$ such that $\pi_0(p) = \pi_0(0) = y_0$. Then:

$$e_0^{mrg}(p) = e_1^{mrg}(p) = e_2^{mrg}(p) = [\frac{p(\frac{1}{2} - \frac{1}{3}p)}{(1 + (1-p) + (1-p)^2)}][y_2 - y_0 - 2b].$$

We also fix $b = 0$ and $y_0 = 0$ (with no consequence for the insights that we derive but with gains in tractability) where the outside options of the firm and the players are zero and we compute the goodness of fit index.

The distances to each property are:

$$||e^{sym}||^2 = \frac{\frac{1}{2}p^2}{(1 + (1-p) + (1-p)^2)^2}y_2^2,$$

and

$$||e^{mrg}||^2 = \frac{3p^2(\frac{1}{2} - \frac{1}{3}p)^2}{(1 + (1-p) + (1-p)^2)^2}.$$

We compute a relative goodness-of-fit measure by dividing this distances by the norm of the given pay profile:

$$||\theta||^2 = \frac{1 + (1-p)^2 + (1-p)^4}{(1 + (1-p) + (1-p)^2)^2}y_2^2.$$

We then have:

$$\begin{aligned}\frac{||e^{sh}(p)||^2}{||\theta(p)||^2} &= \frac{||e^{sym}(p)||^2}{||\theta(p)||^2} + \frac{||e^{mrg}(p)||^2}{||\theta(p)||^2} \\ &= \frac{\frac{1}{2}p^2}{1 + (1-p)^2 + (1-p)^4} + \frac{3p^2(\frac{1}{2} - \frac{1}{3}p)^2}{1 + (1-p)^2 + (1-p)^4}.\end{aligned}$$

Observe that:

$$\frac{||e^{sh}(p)||^2}{||\theta(p)||^2} \to 0 \ as \ p \to 0.$$

We notice that the errors of symmetry are more important than the errors of marginality depending on the value of p (the bargaining power of the firm). For lower values of p, a violation of symmetry is worse than a violation of marginality, but, for high enough values of p, the inverse is true. In fact:

$$\frac{||e^{sym}(p)||^2}{||\theta(p)||^2} > \frac{||e^{mrg}(p)||^2}{||\theta(p)||^2},$$

for $p \in [0, \frac{1}{2}(3 - \sqrt{6}))$.

Also

$$\frac{||e^{sym}(p)||^2}{||\theta(p)||^2} = \frac{||e^{mrg}(p)||^2}{||\theta(p)||^2},$$

at $p = \frac{1}{2}(3 - \sqrt{6})$,

and

$$\frac{||e^{sym}(p)||^2}{||\theta(p)||^2} < \frac{||e^{mrg}(p)||^2}{||\theta(p)||^2},$$

for $p \in (\frac{1}{2}(3 - \sqrt{6}), 1]$.

More importantly, the derivative of the distance to the Shapley value with respect to p is always positive. This means that the higher the bargaining power of the firm, the more the pay scheme differs from the Shapley payoffs. For $p \in (0, 1]$, we have:

$$\begin{aligned} \frac{\partial}{\partial p}\frac{||e^{sh}(p)||^2}{||\theta(p)||^2} &= \frac{\partial}{\partial p}\frac{||e^{sym}(p)||^2}{||\theta(p)||^2} + \frac{\partial}{\partial p}\frac{||e^{mrg}(p)||^2}{||\theta(p)||^2} \\ &= p + \frac{1}{6}p(9 + 2p(-9 + 4p)) > 0. \end{aligned}$$

Note that each term is positive in its specified domain.

These findings show that a sufficiently powerful firm can induce its players to increase their contributions to profits but only at the cost of creating inequality among identical players, and, even more importantly, at the cost of marginality. It remains to be seen if these insights into the effects of the firm's bargaining power on Shapley unfairness can be extended to the case of n players. Nonetheless, the analysis of a two-player firm provided above is very suggestive.