An Index of Unfairness*

Victor H. Aguiar†, Roland Pongou‡, Roberto Serrano§ and Jean-Baptiste Tondji¶

July 9, 2018

Abstract

Aguiar et al. (2018) propose the Shapley distance as a measure of the extent to which output sharing among the stakeholders of an organization can be considered unfair. It measures the distance between an arbitrary pay profile and the Shapley pay profile under a given technology, the latter profile defining the fair distribution. We provide an axiomatic characterization of the Shapley distance, and show that it can be used to determine the outcome of an underlying bargaining process. We also present applications highlighting how favoritism in income distribution, egalitarianism, and taxation violate the different ideals of justice that define the Shapley value. The analysis has implications that can be tested using real-world data sets.

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

Keywords: Shapley value, Shapley distance, (un)fairness, horizontal equity, efficiency, marginality.

---

*We are grateful to the editors of this volume for inviting us to contribute to it, and to an anonymous referee for helpful comments.

†Dept. Economics, University of Western Ontario, vaguiar@uwo.ca

‡Dept. Economics, University of Ottawa, roland.pongou@uottawa.ca

§Dept. Economics, Brown University, roberto_serrano@brown.edu

¶Dept. Economics and Finance, University of Texas Rio Grande Valley, tondji.jean@gmail.com

1 Introduction

Assume that an organization[1] compensates its agents using a pay scheme that possibly violates one or more of the following ideals of justice:

1- Symmetry: equally productive agents receive the same pay.

2- Efficiency: the entire output of the organization is shared among the agents.

3- Marginality: if the adoption of a new technology increases the marginal productivity of an agent, that agent's pay should not decrease relative to the old technology.

How can we measure violations of these ideals of justice for the compensation rule utilized by the organization? As an answer to this question, Aguiar et al. (2018) propose the *Shapley distance*, which, for a given production technology f, measures the distance between an arbitrary pay profile and the Shapley pay profile at f given by the Shapley value (Shapley (1953)). The Shapley value is the only pay scheme that satisfies all of the three aforementioned ideals (Young (1985)). In fact, 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), Serrano (2013)). Moreover, Aguiar et al. (2018) provide an orthogonal decomposition of the Shapley distance into terms that indicate violations of each of the Shapley axioms. This chapter continues this line of research by analyzing the properties characterizing the Shapley distance.

Our main contribution is to axiomatize the *Shapley distance* as a measure of injustice. We also show that the Shapley distance can be used to determine the outcome of a bargaining procedure. We imagine a situation in which agents have to implement a *fairness prescription* F, defined as the set of payoffs induced, under a fixed technology f, by a set of compensation rules $\mathcal{F}$ satisfying certain ideals of justice. There is an initial pay profile ϕ that works as a reference point. Agents may want to depart from ϕ, but they should implement an outcome that belongs to the fairness prescription F. This defines a bargaining function that maps any pair (F, ϕ) to an element of F. We show that the Shapley distance is the unique (up to monotone transformations) index defining a bargaining function that satisfies *Anonymity* and *Independence of Irrelevant Alternatives* (IIA), for the set of compensation rules that obey symmetry, efficiency, and marginality.[2]

Using several illustrations that include favoritism, egalitarianism, and tax distortions, we

[1] We consider an organization as a group of agents including the owner or the principal if any, with each member possibly performing an activity and endowed with a technology.

[2] Some of our ideas are reminiscent of Nash's (1950) pioneering axiomatic characterization of a bargaining solution; see Binmore et al. (1992), Thomson (1994), or Serrano (2008) for surveys of the bargaining literature.

show how the Shapley distance can be applied to determine the extent to which a given income distribution departs from the fair ideal, and how unfairness can be further unbundled to determine its origins.

Together with Aguiar et al. (2018), we contribute to the literature that studies economic inequality using game theory (e.g., Einy & Peleg (1991) and Nembua & Wendji (2016)). In particular, we provide an axiomatic foundation to a notion of unfairness, namely the Shapley distance. A similar axiomatic approach can be used to characterize the decomposition of this distance as provided in Aguiar et al. (2018)

The rest of this chapter is organized as follows. After dealing with preliminaries in section 2, section 3 introduces the Shapley distance and our notion of unfairness and contains our main results. Section 4 presents several applications showing the different ways in which favoritism, egalitarianism, and taxation distort fairness in revenue sharing. Section 5 concludes.

2 Preliminaries

2.1 Organization and Data Set

In this section, we introduce preliminary definitions. We follow Aguiar et al. (2018). Let N be a nonempty and finite set of agents, with $|N| = n$. A coalition is a nonempty subset C of agents: $C \subseteq N$, $C \neq \emptyset$.

An organization is a pair (N, f) where $f : 2^N \mapsto \mathbb{R}$ is a technology such that $f(\emptyset) = 0$. In what remains, we fix N, so that an organization is completely defined by a technology f. We denote by Γ the set of all organizations.

A pay scheme is a way to share the output produced by the grand coalition N of agents.[3]

Definition 1. (**Pay scheme**) A pay scheme is a function $\Phi : \Gamma \mapsto \mathbb{R}^n$ that maps any technology f to a vector $\Phi(f) = (\Phi_1(f), \Phi_2(f), \ldots, \Phi_n(f)) = \phi \in \mathbb{R}^n$ such that $\sum_{i \in N} \Phi_i(f) \leq f(N)$. ϕ is called a pay profile, and for each agent $i \in N$, $\phi_i \in \mathbb{R}$ is interpreted as the payoff of i out of the output $f(N)$. The set of all pay schemes is denoted Θ.

Notice that we allow for negative payoffs, interpreted as taxation. We also recall the notions of observation and data generating pay scheme introduced by Aguiar et al. (2018).

An **observation** is a pair (f, ϕ) where f is a technology and $\phi \in \mathbb{R}^n$ is a pay profile, defined as a distribution of the output generated by the grand coalition, formally $\phi_i \in \mathbb{R}$, for each

[3]Our framework also works if the organization is sharing total cost or total profit. The interpretation of the axioms will have to be done in terms of the context in those cases.

$i \in N$, and $\sum_{i \in N} \phi_i \leq f(N)$. In fact, any vector $\phi \in \mathbb{R}^n$ such that $\sum_{i \in N} \phi_i \leq f(N)$ is called a pay profile in the sequel, even if it is not the result of applying a pay scheme.

Definition 2. (**Data generating pay scheme**) We say that $\Phi : \Gamma \to \mathbb{R}^n$ is a data generating pay scheme if it is the unique pay scheme such that $\Phi(f) = \phi$ for any observation (f, ϕ).

In the context of a limited data set, given by a single observation, we do not have the details about how the data generating pay scheme Φ distributes the total output for a technology that is not the observed technology f. We only know the realized pay profile ϕ for f. However, we have full information on f, (i.e., we know the exact magnitudes of $f(C)$ for all $C \subseteq N$).

2.2 The Shapley Value as an Ideal for Fairness

In this subsection, we recall the definition of the Shapley value as well as its fundamental characterization as a fair pay scheme. This characterization provides an axiomatic basis for analyzing the different ways in which an arbitrary pay scheme might violate basic principles of fairness, as departures from the Shapley value prescription. The following definition will be needed for the statement of these characterizations.

Definition 3. Let $i, j \in N$ be two agents, and f be a technology.

1. The marginal contribution at f of agent $i \in N$ to a set $C \subseteq N$ such that $i \notin C$ is $f(C \cup \{i\}) - f(C)$, and it is denoted by $mc(i, f, C)$.
2. Agent i is a *null-agent at* f if for any set $C \subseteq N$ such that $i \notin C$, we have $mc(i, f, C) = 0$.
3. Agents i and j are said to be substitutes *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)***
A pay scheme Φ satisfies symmetry if for any technology f, and any agents i and j that are substitutes at f, $\Phi_i(f) = \Phi_j(f)$.

Axiom 2. ***(Efficiency)***
A pay scheme Φ is efficient if for any technology f, $\sum_{i \in N} \Phi_i(f) = f(N)$.

Axiom 3. ***(Marginality)***
A pay scheme Φ satisfies marginality if for any technologies f and g, any agent $i \in N$, $[mc(i, f, C) \geq mc(i, g, C); \forall C \subseteq N \setminus \{i\}] \Rightarrow [\Phi_i(f) \geq \Phi_i(g)]$.

The symmetry axiom is a no-discrimination condition (horizontal equity), requiring that agents who have identical marginal contributions under a technology f receive the same pay. Efficiency requires that the output of the grand coalition be fully shared among the various contributors, and it can also be justified in terms of Pareto optimality. Marginality means that, if a new technology increases the marginal productivity (or the vector of marginal contributions) of an agent, that agent's pay should not decrease relative to the old technology. This is an old property in neoclassical economic theory, requiring that the payoff of an agent depend only on his marginal productivity given other agents' inputs.

The result set out below establishes necessity and sufficiency to characterize the Shapley payoff function (defined by equation (1) below). The axioms just presented also establish the Shapley value a fairness ideal.

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

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

3 The Shapley Distance as a Measure of Unfairness

In this section, we provide an axiomatic characterization of the notion of the *Shapley distance* introduced in Aguiar et al. (2018). It measures the level of unfairness associated with any pay profile ϕ by the distance between that pay profile and the Shapley value. Aguiar et al. (2018) show that it can be decomposed into terms that indicate violations of the axioms that characterize the Shapley value. We recall this decomposition and illustrate it through several examples.

3.1 An Axiomatic Characterization of the Shapley Distance

In this section, we provide an axiomatic characterization of the *Shapley distance*. Let $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ be a distance in $\mathbb{R}^n$. 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 4. (**Shapley distance**) For any technology f, the Shapley distance of a pay profile $\phi \in \mathbb{R}^n$ for f, denoted $d(\phi, \boldsymbol{Sh}(f))$, is the distance between ϕ and the Shapley pay profile $\boldsymbol{Sh}(f) \in \mathbb{R}^n$ at f.

We axiomatize below the Shapley distance. First, we need some definitions.

We consider the set of fairness prescriptions of an arbitrary set of pay schemes.

Definition 5. (**Fairness prescription**) Given a technology f and a set of pay schemes $\mathcal{F} \subseteq \Theta$, a set of profiles $F \subseteq \mathbb{R}^n$ is a fairness prescription for f in $\mathcal{F}$ if, for each $\phi \in F$, there exists $\Phi \in \Theta$ such that $\phi = \Phi(f)$.

Our fairness index will be the result of a bargaining procedure, where an original pay profile ϕ works as a reference point. The intuition is that an arbitrator requires all agents to implement a fairness prescription, but the agents are free to choose a new pay profile. They may want to depart from the status-quo ϕ altogether. The result of this procedure is a fairness bargaining function.

Definition 6. (**Fairness bargaining function**) A fairness bargaining function is a mapping $C : \{F\} \times \{\phi\} \to F$ for any fairness prescription F and pay profile ϕ.

We propose an axiomatic approach to studying the properties that the fairness bargaining function ought to have.

Let $\sigma : N \to N$ be a permutation of agents. We define $\sigma(F)$ as the set of fairness prescriptions such that $\varphi \in \sigma(F)$ is a permutation of an element $\eta \in F$. The first axiom requires that the fairness bargaining function is invariant with respect to permutations of the prescriptions and the reference pay profile ϕ.

Axiom 4. ***(Anonymity)***
For all $F \subseteq \mathbb{R}^n$, all $\phi \in \mathbb{R}^n$, and any permutation σ on N, $(C_{\sigma(i)}(F,\phi))_{i\in N} = C(\sigma(F), (\phi_{\sigma(i)})_{i\in N})$.

The second condition requires that the solution to the fairness bargaining problem be optimal.

Axiom 5. ***(Independence of Irrelevant Alternatives (IIA))***
For any set $S \subseteq F \subseteq \mathbb{R}^n$ and any $\phi \in \mathbb{R}^n$, $C(F,\phi) \in S$ implies $C(F,\phi) = C(S,\phi)$.

Without loss of generality,[4] we also assume that any $F \subseteq \Theta$ is convex and closed.

Lemma 1. *The only fairness bargaining function that satisfies Anonymity and IIA is the minimal distance bargaining function*

$$C(F,\phi) = argmin_{v\in F} d(v,\phi).$$

[4] Our analysis also works for cases where this is not the case, but the particular case we study for a Euclidean Shapley Distance with the given fairness prescriptions is covered in this restriction.

Proof. To check that the minimal distance bargaining function satisfies Anonymity and IIA is trivial. To prove uniqueness, we observe that anonymity axiom implies the following two axioms: Invariance to permutations (IP) and Nash symmetry (NS).

(i) A fairness prescription is closed to permutations whenever: $\phi \in F$ implies $(\phi_{\sigma(i)})_{i \in N} \in F$ for any permutation of the set of agents $\sigma : N \to N$.

Invariance to permutations (IP): If F is closed to permutations then $C_i(F, \phi) = C_j(F, \phi)$ for all $i, j \in N$.

(ii) A fairness prescription F is called symmetric if the set F is symmetric around the 45 degree line.

Nash symmetry (NS): If F is symmetric and $\phi_i = \phi_j$ for all $i, j \in N$, then $C_i(F, \phi) = C_j(F, \phi)$ for all $i, j \in N$.

If axioms (IP) and (NS) hold then the symmetry axiom to a line property in Rubinstein & Zhou (1999) holds. A line is $< a, \alpha >$ where $a \in \mathbb{R}^n$ is a reference and $\alpha \in \mathbb{R}^n$ is a direction, such that the points of the line are $a + t\alpha$ where t is a real number.

We say that F is symmetric to a line $< a, \alpha >$ if for every orthogonal direction β ($\beta'\alpha = 0$), $a + t\alpha + \beta \in F$ implies that $a + t\alpha - \beta \in F$. The condition of Rubinstein & Zhou (1999) requires that if F is symmetric to a line $< \phi, \alpha >$ where ϕ is a pay profile, then $(C_i(F))_{i \in N} \in < \phi, \alpha >$. In fact, axiom (IP) means that if F is symmetric with respect to the line $(t, \cdots, t)'$ for any real number t, then $(C_i(F, \phi))_{i \in N} \in (t, \cdots, t)'$. Axiom (NS) requires that, if F is symmetric with respect to the 45 degree line and $\phi_i = \phi_j$, then $(C_i(F, \phi))_{i \in N} \in (t, \cdots, t)'$ (i.e., $C_i(F, \phi) = C_j(F, \phi)$ for all $i, j \in N$). Axioms (IP) and (NS) imply that if F is symmetric for any line going through ϕ, then the solution will be on that line. This is proved in Rubinstein & Zhou (1999). Then by Proposition 2.1 in Rubinstein & Zhou (1999), we establish that (IP), (NS) and (IIA) axioms imply that

$$C(F, \phi) = argmin_{v \in F} d(v, \phi).$$

. ■

Next we define our fairness index.

Definition 7. (**Fairness index**) A fairness index is a mapping ($\rho : \mathbb{R}^n \times \{\phi\} \mapsto \mathbb{R}_+$) such that there exists a fairness bargaining function C defined as follows:

$$C(F, \phi) = argmin_{v \in F} \rho(v, \phi).$$

We are ready to present our main result.

Theorem 1. ***(Shapley Distance)****. Let C be a bargaining function that satisfies Anonymity and IIA. Then the Shapley distance is the unique (up to monotone transformations) value of the fairness index defining C at any point (F, ϕ) where F is induced by the set $\mathcal{F}$ of pay schemes that satisfy symmetry, efficiency, and marginality.*

Proof. By Lemma 1, the bargaining function C is defined by the minimal distance function: $C(F, \phi) = argmin_{v \in F} d(v, \phi)$ for any convex and closed set F. By Claim 1 (see also Young (1985)), we know that, for any technology f, the fairness prescription F induced by the set of pay schemes that satisfy symmetry, efficiency, and marginality is the singleton $\{\boldsymbol{Sh}(f)\}$, which is a convex and closed set. It follows that $C(F, \phi) = argmin_{v \in \{\boldsymbol{Sh}(f)\}} d(v, \phi)$. But $min_{v \in \{\boldsymbol{Sh}(f)\}} d(v, \phi) = d(\boldsymbol{Sh}(f), \phi)$, which completes the proof. ■

Different choices of the distance function provide different fairness indices. We focus now on a particular choice, the Euclidean distance, which is shown by Aguiar et al. (2018) to have an additive (and orthogonal) comparability property in terms of the different axioms of fairness, hence justifying its use. As recalled below, the square of the Shapley distance has a unique **decomposition** into terms that measure violations of the classical axioms of the Shapley value. This approach is analogous to that of Aguiar & Serrano (2017) who study departures of a demand function from rationality. Despite the similarities in the two approaches, in this paper we address a different question in a different environment.

Moreover, in finite data sets, these terms can be used to make partial inferences about the violations of the axioms defined for complete data sets, and to make complete inference about the violations of the axioms defined for a fixed technology, for the subset of monotone technologies, (see also Aguiar & Serrano (2016)). This is of interest because the observer usually does not have information about a pay scheme under different technologies making it practically impossible to check the validity of the axioms that require comparisons between different technologies.

3.2 A Decomposition of the Shapley Distance with Limited Datasets

We now present a decomposition of the Euclidean Shapley distance, or Shapley Distance for short. In this section, we follow the set-up in **?**. Let f be a technology and $\phi \in \mathbb{R}^n$ 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 technology f and an error term $\phi = \boldsymbol{Sh}(f) + e^{sh}$, by defining $e^{sh} = \phi - \boldsymbol{Sh}(f) \in \mathbb{R}^n$. Moreover, we 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.

Aguiar et al. (2018) find this orthogonal decomposition to be the result of the following procedure. First they find the closest pay scheme to ϕ that satisfies sym; then they find the closest pay scheme to ϕ that satisfies eff in addition to sym; and finally they find the closest pay scheme to ϕ that satisfies mrg in addition to sym and eff, which is simply the Shapley value itself. The described order, in which these constraints are imposed, is the only one 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.

Begin by fixing a pair consisting of an observed pay profile and a technology (f, ϕ) and consider the Shapley distance of ϕ at this point, which is:

$$||e^{sh}|| = ||\phi - \boldsymbol{Sh}(f)||.$$

Let $\boldsymbol{v}^{sym}$ be the closest pay scheme to ϕ that satisfies symmetry (pointwise under the chosen norm) (i.e., $\boldsymbol{v}^{sym} \in argmin_{\boldsymbol{v} \in \boldsymbol{\Theta}} ||\phi - \boldsymbol{v}(f)||$ s.t. $\boldsymbol{v}$ satisfies sym).[5] Aguiar et al. (2018) prove that each entry evaluated at f is given by v_i^{sym} that corresponds to the average pay according to ϕ among the agents who are substitutes of i under f. They then establish that ϕ 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:

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

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). **?** 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 ϕ divided by the number of agents in N. It follows that v^{sym} can be uniquely written as:

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

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

Finally, remark 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

[5] Existence is easy to verify noticing that the space of symmetric pay schemes is convex and closed.

be the Shapley value because of the uniqueness established in Claim 1. Thus $v^{sym,eff,mrg} = \boldsymbol{Sh}(f)$. Thus, we let $e^{mrg} = v^{sym,eff} - \boldsymbol{Sh}(f)$. Notice that we can always decompose ϕ (pointwise) as:

$$\phi = \boldsymbol{Sh}(f) + e^{sh},$$

because ϕ and $\boldsymbol{Sh}(f)$ belong to the same vector space. With all this in hand, Aguiar et al. (2018) establish the following main result.

Theorem 2. *(Aguiar et al. (2018)) For any given observation (f, ϕ), we have the unique pointwise decomposition:*

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

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

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

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

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

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

To the best of our knowledge, $||e^{sh}||^2$, introduced in (**?**), is the first measure of departures from the Shapley axioms. It has the advantage of providing a unified treatment of the three axioms in the form of a numerical and additive decomposition. Furthermore, in the decomposition analysis, 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.

4 Some Applications

In this section, we feature several applications of our analysis. They are attempts to enhance our understanding of inequality, and answer the question of when income inequality can be considered unfair. The different applications show how favoritism, egalitarianism, and taxation distort fairness in revenue distribution.

4.1 Favoritism

Consider the following simple example:

Example 1. The nephew's problem. Let an organization consist of a set of agents $N = \{1, 2, 3\}$ and a technology f defined as follows: $f(N) = 10$, $f(\{1, 2\}) = 4$, $f(\{1, 3\}) = f(\{2, 3\}) = 9$, $f(\{i\}) = 0$ for $i = 1, 2, 3$. The environment describes a firm owned by agent 3, who employs a nephew (agent 1). Agent 2 is also employed in the firm, with no family connections to the other two people. Although from the point of view of productivities, agents 1 and 2 are substitutes, agent 3, exhibiting favoritism toward agent 1, allows him to show up to work only half of the time, leading to output waste. In addition, the uncle has set the pay scheme $\Phi(f) = (2, 1, 4)$. Note that the Shapley value yields the pay profile $\boldsymbol{Sh}(f) = (2.5, 2.5, 5)$. Thus, the overall Shapley distance is 3.5, decomposed as 0.5 (attributed to the violation of symmetry) and 3 (attributed to the violation of efficiency). No violation of marginality is observed, after one corrects for the other two failures: the moves in $\mathbb{R}^3$ describe a first transition from $(2, 1, 4)$ to $(1.5, 1.5, 4)$–correcting for symmetry–, and then to $(2.5, 2.5, 5)$ -correcting for efficiency-, which is the Shapley value. In this example, favoritism causes an efficiency flaw that, according to our measure, is 6 times as important as the lack of symmetry.

4.2 Egalitarianism versus Fairness

Our second illustration relates to the egalitarian pay scheme. Before showing it, we need to present a generalization, due to Hsiao & T.E.S. (1993), Pongou et al. (2017), and Pongou & Tondji (2018), of the framework of an organization, to an environment where agents have more than two options (i.e, active or inactive). A **production environment** is modeled as a list $\mathcal{G} = (N, L, G)$ where $N = \{1, 2, \ldots, n\}$ is a nonempty finite set of agents of cardinality n; $L = \{0, 1, 2, \ldots, l\}$ is a nonempty finite set of hours of labor or effort levels that an agent can supply, with 0 denoting a situation of inaction; and G is a production function that maps each action profile $x = (x_1, \ldots, x_n) \in L^n$ to a real number –output– $G(x)$. The function G can also be interpreted as the aggregate profit or cost function. Interpreting it as the profit function might be useful in certain settings, in that it could be incorporating both production and cost functions. Regardless of the interpretation, we assume that $G(0,0,\ldots,0) = 0$, which means that no output is produced when all the agents are inactive.

We denote by e_i the i^{th} unit vector $(0,0,\ldots,0,1,0,\ldots,0)$, 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, $(1,7,5,0,\ldots,0) \triangleleft (1,7,5,1,5,0,\ldots,0)$. We denote by $|x| = |\{i \in N : x_i > 0\}|$ the number of agents 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 $G(x) \leq G(y)$ whenever $x \trianglelefteq y$.

For any production environment $\mathcal{G} = (N, L, G)$, a pay scheme for the production maps any effort profile $\overline{x} \in L^n$ to a non null payoff profile $\Phi^G(\overline{x}) = (\Phi_1^G(\overline{x}), \Phi_2^G(\overline{x}), \ldots, \Phi_n^G(\overline{x}))$, where for all $i \in N$, $\Phi_i^G(\overline{x}) \in \mathbb{R}$ is interpreted as the payoff earned by i out of the output $G(\overline{x})$. In the production environment, an **observation** is a triple $(\overline{x}, G, \Phi^G(\overline{x}))$ where $\phi = \Phi^G(\overline{x})$ is an observed pay profile for any production function G and for any effort profile $\overline{x}$.

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

$$\boldsymbol{Sh}_i^G(\overline{x}) = \sum_{x \triangleleft\, \overline{x},\; x_i = 0} \frac{(|x|)!(|\overline{x}| - |x| - 1)!}{(|\overline{x}|)!}[G(x + \overline{x}_i e_i) - G(x)],\ \textit{for all}\ i \in N. \tag{2}$$

Aguiar et al. (2018) show that, for a fixed level of efforts $\overline{x}$, all the information given by the production environment can be equivalently expressed using a technology.

We now show how the egalitarian pay scheme distorts fairness in revenue distribution. This pay scheme is the benchmark that implements perfect equality. It divides the output in equal parts to each agent. So, this pay scheme is clearly efficient. Evidently, given different levels of efforts and productivities, the egalitarian pay scheme may not be fair, failing marginality. Our aim is to measure the divergence of the egalitarian pay scheme from the Shapley value and to identify the sources of this divergence. We do this through the following example in which, for simplicity, we assume two agents, with each choosing his effort level from a set that contains two levels.

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

$$G(x) = \begin{cases} 1 & if \quad x \neq (0,0) \\ 0 & if \quad x = (0,0) \end{cases} \tag{3}$$

Consider the egalitarian pay scheme $\boldsymbol{Eq}$ defined as follows:

$$\boldsymbol{Eq}_1(x) = \frac{1}{2}G(x) \text{ and } \boldsymbol{Eq}_2(x) = \frac{1}{2}G(x), \text{ for each } x \in L^2.$$

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

In order to quantify the violations of the properties that characterize the Shapley value, let us first derive the Shapley payoff of each agent at each vector $\overline{x}$. The Shapley payoff profile at each $\overline{x}$ is given by the following matrices: $\boldsymbol{Sh}^G(\overline{X}) = \begin{pmatrix} (0,0) & (0,1) \\ (1,0) & (\frac{1}{2},\frac{1}{2}) \end{pmatrix}$, 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 agent 1, and the second component denoting the effort level of agent 2.

The egalitarian payoff profile is given by: $\boldsymbol{Eq}(\overline{X}) = \begin{pmatrix} (0,0) & (\frac{1}{2},\frac{1}{2}) \\ (\frac{1}{2},\frac{1}{2}) & (\frac{1}{2},\frac{1}{2}) \end{pmatrix}$.

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

Note that Theorem 2 applies for each fixed effort level, equivalently for each entry of the matrix $\overline{X}$.

We now determine how the amount by which the violation of each property characterizing the Shapley value contributes to the total violation of fairness by an egalitarian payoff for any production function and any number of agents. We know that:

$$\boldsymbol{Eq}(\overline{x}) = \boldsymbol{Sh}^G(\overline{x}) + e^{sym} + e^{eff} + e^{mrg}.$$

1. Let $e^{sym} = Eq - v^{sym} = 0$. For all effort levels x, because $\boldsymbol{Eq}$ satisfies symmetry trivially.

2. Let $e^{eff} = v^{sym} - v^{sym,eff} = 0$. For all effort levels x, because $\boldsymbol{Eq}$ satisfies efficiency trivially.

3. Let $e^{mrg} = v^{sym,eff} - \boldsymbol{Sh}^G = \boldsymbol{Eq} - \boldsymbol{Sh}^G$. This means that the Shapley distance in general for this case is equal to $||\boldsymbol{Sh}^G - \boldsymbol{Eq}||^2 = ||e^{mrg}||^2$. This implies that a perfectly egalitarian pay profile may still be unfair given certain productivity and effort levels.

4.3 Taxes

In our third example, we illustrate how a tax levied over a fair wage can alter the fairness in an economy.

Example 3. Consider a small economy of two agents 1 and 2 who have to work to produce goods and services. Each agent has two options, either go to work (option W), or stay at home (option H). The production function is given by: $f(H,H) = 0$, $f(H,W) = 2$, $f(W,H) = 1$, and $f(W,W) = 5$. We observe that both agents work (i.e., we observe the effort profile (W,W)). This implies that the Shapley wage function allocates a payoff of 2 dollars to agent 1, and a payoff of 3 dollars to agent 2.

We assume that both agents have to contribute for a public good. For simplicity, we assume that the benefits from the public good are not received immediately and we can ignore them in the payoff profile. The vector $\Phi = (2(1-\alpha), 3(1-\alpha))$ represents the revenues of agents net of contributions, given that each agent contributes a positive proportion α of his/her revenue. How far is $\Phi(f)$ from the Shapley allocation $\boldsymbol{Sh}^f = (2,3)$?

The Shapley distance is given by $||e^{sh}||^2 = ||\boldsymbol{Sh}^f - \Phi||^2 = 13\alpha^2$. We now determine how the amount by which the violation of each fairness property characterizing the Shapley value contributes to the total violation of $13\alpha^2$.

1. $e^{sym} = \Phi - v^{sym}$. Since agents are not identical, it follows that $\Phi_i = v_i^{sym}$ and $e^{sym} = (0,0)$.

2. $e^{eff} = \Phi - v^{sym,eff}$. For each $i \in \{1,2\}$, $v_i^{sym,eff} = \Phi_i + \frac{5-\sum \Phi_i}{2}$. After calculations, $v^{sym,eff} = (\frac{4+\alpha}{2}, \frac{6-\alpha}{2})$, and $||e^{eff}||^2 = \frac{25\alpha^2}{2}$.

3. $e^{mrg} = \boldsymbol{Sh}^F - v^{sym,eff} = (-\frac{\alpha}{2}, \frac{\alpha}{2})$. Then, $||e^{mrg}||^2 = \frac{\alpha^2}{2}$. A quick verification confirms that $||e^{mrg}||^2 + ||e^{eff}||^2 = 13\alpha^2$. In general, we observe that the tax has an increasing and nonlinear distortion of fairness. When $\alpha \to 0$ there is no unfairness in the economy, and when $\alpha \to 1$ the unfairness level reaches its maximum.

Assuming that each agent contributes half of his/her revenue (i.e., $\alpha = \frac{1}{2}$), the departure from the Shapley allocation is $||e^{sh}|| = 1.8$ dollars. In addition, 96.15 percent of this value is explained by the violation of efficiency, and 3.85 percent by lack of marginality.

The previous example provides an upper bound to the cost of fairness. However, we made the strong assumption that there is no enjoyment of the public good by the agents. Here we relax that assumption and provide a lower bound of the cost of fairness.

Example 4. We consider the same economy defined in Example 3, but we assume that there is monetary (equivalent) benefit of the public good that can be enjoyed by both agents immediately. The total tax revenue is given by 5α dollars. We assume that each agent enjoyment of the public good is $\frac{5}{2}\alpha$ dollars. This implies that the adjusted payoff after taxes and considering the public good utility is $\Phi = (1-\alpha)(2,3) + \alpha(\frac{5}{2}, \frac{5}{2})$. In other words, the government is able to implement a convex combination of the Shapley wage and the egalitarian wage using a fully efficient tax to provide a public good that produces the same enjoyment to both agents.

The Shapley distance is given by $||e^{sh}||^2 = ||\boldsymbol{Sh}^f - \Phi||^2 = \frac{\alpha^2}{2}$. We notice that the new pay scheme is both efficient and symmetric, hence $||e^{sh}||^2 = ||e^{mrg}||^2$, which coincides with the marginality error in the previous example. In this example, the government is able to eliminate the efficiency loss and only the marginality loss remains. Note that when $\alpha \to 1$ there is a loss of $||e^{sh}|| = \frac{1}{\sqrt{2}} \approx 0.707$ dollars in terms of unfairness to produce a fully egalitarian income. This is 14.14% of the total output. This is of course a lower bound to the cost of fairness (while the previous example represented an upper bound).

For our final example, we consider a different tax scheme and explore its implications for fairness.

Example 5. We consider the same economy defined in Example 3, but we assume that the investment in the public good is done by using a lump-sum tax scheme, as opposed to the proportional tax scheme. Specifically, each agent contributes the amount t_i, $i \in \{1,2\}$, such that $t_1+t_2 = X$, where X represents the worth of the public good. The vector $\Phi = (2-t_1, 3-t_2)$ represents the revenues net of taxes. What could be the values of t_i, such that the vector Φ is close to the Shapley payoff vector $\boldsymbol{Sh}^f = (2,3)$? The distance between the two vectors Φ and $\boldsymbol{Sh}^f$ is given by the numerical expression $d(t_1,t_2) = t_1^2 + t_2^2$. To answer the question posed, we should solve the following minimization problem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & t_1^2 + t_2^2 \\ \text{subject to} \quad & 0 \le t_1 \le 2 \; ; 0 \le t_2 \le 3 \; ; t_1 + t_2 = X \; ; 0 < X \le 5. \end{aligned} \tag{4}$$

Solving problem 4 yields $t_1^* = \min(2, \frac{X}{2})$ and $t_2^* = \min(3, X - t_1^*)$. Assume that the amount of the public good X equals 4.5 dollars, then agent 1 contributes $t_1^* = 2$ dollars, agent 2 contributes $t_2^* = 2.5$ dollars. The payoff vector is $\Phi = (0, 0.5)$ net of taxes. The distance between both allocations Φ and $\boldsymbol{Sh}^f$ is $||e^{sh}|| = 3.20$ dollars. The vector Φ does not violate the symmetry property, since agents are not identical. The violation of efficiency, measured by $||e^{sh}|| = 3.18$ dollars, represents 98.78 percent of the total measure of unfairness ($\frac{||e^{eff}||^2}{||e^{sh}||^2} = 98.78$), whereas only ($\frac{||e^{mrg}||^2}{||e^{sh}||^2} = 1.22$) of unfairness is explained by the lack of marginality. Again, this is an upper bound of the cost of fairness. Due to the decomposition, it is easy to see that the way to reduce the important cost of fairness is to reduce the efficiency error. This can be done by taking into account the benefits of the public good. If the benefits of the public good are fully internalized, only the marginality error will matter, and that is smaller than in the tax schemes of previous examples.

5 Conclusions

We have provided an axiomatic characterization of the *Shapley distance*, which is a measure of unfairness in revenue distribution introduced by Aguiar et al. (2018). It is defined as the distance between an arbitrary pay profile and the Shapley pay profile under a given technology. Aguiar et al. (2018) provide a decomposition of this distance into terms that measure violations of each of the Shapley axioms. In this chapter, we have shown that the Shapley distance is the unique (up to monotone transformations) index defining a bargaining function that satisfies *Anonymity* and *IIA* for the set of pay schemes that obey symmetry, efficiency, and marginality. The analyses are illustrated through examples showing the different ways in which favoritism, egalitarianism, and taxation distort fairness in revenue sharing. We have also identified a tax scheme that minimizes this distortion.

References

Aguiar, V., Pongou, R., & Tondji, J.-B. (2018). A non-parametric approach to testing the axioms of the Shapley value with limited data. *Games and Economic Behavior.*

Aguiar, V. H. & Serrano, R. (2016). Classifying bounded rationality in limited data sets: A slutsky matrix approach. *Forthcoming in SERIEs, Journal of the Spanish Economic Association.*

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

Binmore, K., Osborne, M. J., & Rubinstein, A. (1992). Noncooperative models of bargaining. *Handbook of Game Theory with Economic Applications*, 1, 179–225.

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

Hsiao, C. & T.E.S., R. (1993). Shapley value for multichoice cooperative games, I. *Games and Economic Behavior*, (5), 240–256.

Nembua, C. C. & Wendji, C. M. (2016). Ordinal equivalence of values, Pigou–Dalton transfers and inequality in TU-games. *Games and Economic Behavior*, 99, 117–133.

Pongou, R., Tchantcho, B., & Tedjeugang, N. (2017). Properties of ladder tournaments. *European Journal of Operational Research*, 263(1), 203–213.

Pongou, R. & Tondji, J.-B. (2018). Valuing inputs under supply uncertainty: The Bayesian Shapley value. *Games and Economic Behavior*, 108, 206–224.

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

Rubinstein, A. & Zhou, L. (1999). Choice problems with areference'point. *Mathematical Social Sciences*, 37(3), 205–209.

Serrano, R. (2008). Bargaining. *The New Palgrave Dictionary of Economics (2nd edition), s. Durlauf and L. Blume (eds.) McMillian, London.*

Serrano, R. (2013). Lloyd Shapley's matching and game theory. *The Scandinavian Journal of Economics*, 115(3), 599–618.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28), 307–317.

Thomson, W. (1994). Cooperative models of bargaining. *Handbook of Game Theory with Economic Applications*, 2, 1237–1284.

Yaari, M. E. (1981). Rawls, Edgeworth, Shapley, Nash: Theories of distributive justice reexamined. *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.