

# A Decomposition of the Atkinson Index through the Shapley Value

Joss Sánchez-Pérez, Víctor Rosas-Martínez

Faculty of Economics, Autonomous University of San Luis Potosí, San Luis Potosí, México

Email: [joss.sanchez@uaslp.mx](mailto:joss.sanchez@uaslp.mx), [vick\\_tor\\_2000@hotmail.com](mailto:vick_tor_2000@hotmail.com)

Received September 17, 2012; revised October 19, 2012; accepted November 20, 2012

## ABSTRACT

This paper studies the economic inequality as a phenomenon resulting from two variables: The variable of responsibility which depends on the agents' performance in the economy, and the variable of opportunity which is independent of the individuals' performance in the economy. Identifying the contribution of the inequality of opportunity in the total inequality as the problem to be measured, we propose a decomposition of the Atkinson inequality index, by using one of the central solution concepts of cooperative game theory, called the Shapley value.

**Keywords:** Atkinson Index; Shapley Value; Opportunity Inequality

## 1. Introduction

“Social equality is an extremely complex notion” (Sen [1]). Walking on the streets of any city with a free market economic system, observing the behaviors and characteristics of the people, we find that economic inequality is a totally natural and usual aspect. The existence of such inequality is totally intuitive, because the citizens play economic activities with different characteristics of human capital and therefore they get different remunerations. If the problem of economic inequality were thought from this point of view, then there would be no need for questions or concerns at all, since it would only be a consequence of the individual decisions made by citizens, given their preferences. But if the economic inequality is very large, would be hard to believe that there are agents that would take decisions which would lead them to take certain spots of the economic structure, or worse, in an extreme case, to death for the lack of resources to survive; which tells us that the inequality is not necessarily a consequence of the different rational agents decisions. It is also important to mention that there is a close relationship between the rebellions propensity or fluctuations of insecurity in a population, and the economic inequality that this presents. Inequality must be therefore measured, because it is a strong indicator of the economic situation and could also describe more aspects of a defined territory or population. Following the current social welfare functions theory in Bergson-Samuelson [2], which defines that welfare must be based on the utility of the agents, breaking with the concept of welfare of Pareto which has no power of

discrimination and therefore no useful judgment to the measure of this problem (see for instance Sen [1]). In order to rank welfare functions, Atkinson [3] assumes increasing, concave and equal functions of utility among the individuals and establishes some desirable characteristics<sup>1</sup> that this ranking should have. He also looks for a measure of the total inequality which respects such ranking and based in this he defines the equally distributed equivalent level of income, or the level of income per head which if equally distributed would give the same level of social welfare as the present distribution, and using this concept he defines an inequality index. Levy *et al.* [4] found this index as the only one that satisfies some desire properties in order to give a health benefits analysis. So in this work we will use this measurement to get the inequality of a given income distribution.

Inequality is an aspect that has been treated and observed throughout human history. Awakening the interest of the society, and humanitarian feelings, this topic has originated some spirit of solidarity and total apathy. Trying to focus on the income inequality problem, and explaining it as a result of two variables (responsibility and opportunity), the present work arises.

Following the welfare economics theory, that looks to provide a judgment criteria for a given income distribution, and using one of the main concepts of the cooperative games theory, which has been used to decompose inequality indices; and considering that economic inequality is not necessarily a result of the agents' rational

<sup>1</sup>Such as the Daltons' principle of transfers (see Atkinson [3]).

actions, we address the following questions: Given a level of economic inequality in a population,

- 1) How much inequality is due to the difference in opportunity for the population? and
- 2) How much inequality is due to their decisions?

This paper provides an identification of the contribution of the inequality of opportunity in the total inequality as the problem to be measured, we propose a decomposition of the Atkinson inequality index, by using one of the central solution concepts of cooperative game theory, called the Shapley value. With this we explain the participation of each of these variables in the total inequality for a given income distribution.

The application of the Shapley value algorithm to decompose inequality and poverty measures has been proposed by Shorrocks [5] and by Sastre and Trannoy [6] for inequality measures. We propose a modified application of the Shapley value which takes into account a multiplicative class of indices expressed in equality terms.

The paper is organized as follows. We first recall the main basic features of games in characteristic function form and Atkinson index. The equality of opportunity model is introduced in Section 3, taken out from Peragine [7]. In Section 4 we propose a decomposition of the Atkinson index, using the Shapley value to find the participation of the total inequality explanatory factors. Finally, Section 6 concludes.

## 2. Framework

In this section we provide some concepts and notation related to  $n$ -person-person games in characteristic function form, as well as a brief subsection of preliminaries related to the Atkinson index.

### 2.1. Games in Characteristic Function Form

By an  $n$ -person game<sup>2</sup> in characteristic function form, in what follows just a game, we mean a pair  $(N, v)$ , where  $N = \{1, \dots, n\}$  is a finite set of players and  $v$  is a function  $v: 2^N \rightarrow R$  with the property that  $v(\emptyset) = 0$  ( $2^N$  denotes the set of subsets of  $N$ ). We usually refer to subsets  $S$  of  $N$  as coalitions and to the number  $v(S)$  as the worth of  $S$ . There are several interpretations for  $(N, v)$ , it depends on what people want to model.

We denote by  $G^N$  the set of all games with a fixed set of players  $N$ , *i.e.*,

$$G^N = \{v: 2^N \rightarrow R | v(\emptyset) = 0\}$$

Given  $v, w \in G^N$  and  $\lambda \in R$ , we define the sum and

the product,  $v + w$  and  $\lambda v$ , in  $G^N$  in the usual form, *i.e.*,

$$(v + w)(S) = v(S) + w(S) \text{ and } (\lambda v)(S) = \lambda v(S)$$

It is easy to verify that  $G^N$  is a vector space with these operations.

Next, we define some axioms which are common in the cooperative game theory frame-work.

**Axiom 1 (Linearity)** *The solution  $\phi$  is linear if  $\phi(v + w) = \phi(v) + \phi(w)$  and  $\phi(cv) = c\phi(v)$ , for all  $v, w \in G^N$  and  $c \in R$ .*

Let us consider the group  $S_n = \{\theta: N \rightarrow N | \theta \text{ is bijective}\}$ , the group of permutations of the set of players. For every  $\theta \in S_n$  and  $v \in G^N$  we define another game  $(N, \theta \cdot v)$  as follows

$$\theta \cdot v(S) = v(\theta^{-1}S)$$

**Axiom 2 (Symmetry)** *The solution  $\phi$  is said to be symmetric if and only if  $\phi(\theta \cdot v) = \theta \cdot \phi(v)$  for every  $\theta \in S_n$  and  $v \in G^N$ .*

The player  $i$  is said to be a null player in the game  $(N, v)$  if  $v(S \cup \{i\}) = v(S)$  for every  $S \subseteq N$ . The nullity axiom requests that every dummy player in  $(N, v)$  gets a zero payoff.

**Axiom 3 (Nullity)** *If the player  $i$  is a null player in  $(N, v)$ , then  $\phi_i(v) = 0$ .*

**Axiom 4 (Efficiency)** *The solution  $\phi$  satisfies the efficiency axiom if and only if  $\sum_{i \in N} \phi_i(v) = v(N)$  for every  $v \in G^N$ .*

**Theorem 5 (Shapley [9])** *There exists a unique solution  $Sh: G^N \rightarrow R^n$  that satisfies linearity, symmetry, nullity and efficiency axioms. Furthermore it is given by*

$$Sh_i(v) = \sum_{S \subseteq N/\{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} [v(S \cup \{i\}) - v(S)]$$

### 2.2. The Atkinson Index

Atkinson [3] defines the *equally distributed equivalent* level of income,  $y_{ede}$ , which is the per capita income level that distributed among the population in an egalitarian way, presents the welfare level of the given income distribution. That is, the level of income per head which if equally distributed would give the same level of social welfare as the present distribution, so  $y_{ede} = y$  if  $y$  satisfies

$$mU(y) = \sum_{i=1}^m U(y_i)$$

where  $y_i$  is the income level of the person  $i$ ,  $U$  is a linear or concave utility function, and  $m$  is the number of agents in the economy, and from it Atkinson defines the measure of inequality

<sup>2</sup>For a brief revision of the concepts of cooperative games that are mentioned here, see Driessen [8].

$$D = 1 - \frac{y_{ede}}{\mu}$$

or 1 minus the ratio of the equally distributed equivalent level of income to the mean  $\mu$  of the actual distribution. If  $D$  falls, then the distribution has become more equal—we would require a higher level of equally distributed income (relative to the mean) to achieve the same level of social welfare as the actual distribution. The measure  $D$  has, of course, the convenient property of lying between 0 (complete equality) and 1 (complete inequality).

Consider a society  $I = \{1, 2, \dots, m\}$ , where their incomes are given by the vector  $X = (x_1, x_2, \dots, x_m)$ . The Atkinson's equality index of the distribution  $X$  according to the value of parameter  $\varepsilon = 1$  is by Atkinson [3]<sup>3</sup>:

$$A(X) = \frac{\prod_{j=1}^m (x_j)^{\frac{1}{m}}}{\mu}$$

where  $\mu$  denotes the mean of the values in the income vector  $X$ .

### 3. The Model

The model will be based on the one proposed by Peragine [7], which focuses on the individuals economic opportunities differences, as a welfare ranking criteria for different income distributions.

We consider a society  $I$  with  $|I| = m$  individuals, each individual's income  $x$  is causally determined by two kinds of factors: Factors beyond the individual control, represented by a person's opportunity set  $O$  and factors for which the individual is fully responsible, represented by a scalar variable  $w$ : Income  $x$  is supposed to be continuously distributed, with cumulative distribution function  $F(x)$  and  $x \in [0, z]$ , where  $z$  is the maximum individual income. A person's opportunity set  $O$  is observable, and we denote by  $\Omega = \{O^1, \dots, O^l\}$  the set of all possible opportunity sets, considered to be finite. We assume that there is a general political agreement on the following ordering over all possible opportunity sets, assumed to be antisymmetric:

$$O^{i+1} \succ O^i \forall i \in \{1, \dots, l-1\}$$

The responsibility variable  $w$  is unobservable and individuals have the same degree of access to the set  $\Theta$ ,  $\Theta \subseteq R_+$ , of all possible values of the responsibility variable.

The individual income  $x$  is then given by

$$x = g(O, w)$$

where  $g : \Omega \times \Theta \rightarrow R_+$  is a function that assigns indi-

vidual incomes to combinations of responsibility and opportunities.

Now, there is a partition the population into types, where each type denotes the subset of the total population having the same opportunity set: Type  $i$  is defined as the set of individuals whose opportunity set is  $O^i$ , and is denoted by  $I^i$ . We denote by  $\Psi$  the set of income distributions. For any  $F \in \Psi$ , within each type  $i$  there will be a distribution of income  $F^i(x)$ , with density function  $f^i(x)$ .

Consider any individual  $k$  who is endowed with opportunity  $O^i$  and, after exercising responsibility  $w_k$ , ends up with ex-post income  $x_k^i = g(O^i, w_k)$ . Denoting by  $\mu_F^i$  the mean income of type  $i$ , defined as  $\mu_F^i = \int x f^i(x) dx$  and by  $r_k$  the residual terms, such that, within each type  $i$ ,  $\int r f(x) dx = 0$ , we can write the income level of individual  $k$  in type  $i$  as

$$x_k^i = \mu^i + r_k$$

Now, consider that income differences between the individuals within the same type will only be because of differences on their responsibility. On the other hand, it seems reasonable to interpret the average income reached within each type, as a characteristic of the type itself; that is to say that the average income is determined, within each type, by the common endowment of opportunity.

Formally:

#### Assumption 1

- 1)  $\forall i \in \{1, \dots, l\}$ ,  $\exists \mu : \Omega \rightarrow R$  such that  $\mu^i = \mu(O^i)$ ;
- 2)  $\forall i \in \{1, \dots, l\}$ ,  $\forall k \in I^i$ ,  $\exists r : \Theta \rightarrow R$  such that  $r_k = r(w_r)$ .

This assumption implies that the individual income is additively separable in responsible and non-responsible factors. From this we can say that the income average of each type depends on the opportunity, so since all of the individuals inside of  $I^i$  had the same opportunity  $O^i$  then we could say that their average income will indicate the part of the income which all of them could get. And therefore the rest of the income will be explained by a residual term which depends on the performance of the individuals in the economy (responsibility variable).

#### Assumption 2

$$\forall i \in \{1, \dots, l\}, \forall k \in I^i, \mu(O^i) > \mu(O^{i-1}) \text{ and } r'(w_k) > 0$$

This assumption implies that the type means are ordered according to the opportunities order. Notice that the assumption also implies that the residual term increases with the responsibility variable.

### 4. The Decomposition

Decomposition analysis provides help in the difficult task

<sup>3</sup>Defined also in Lasso de la Vega and Urrutia [10].

of explaining how different factors affect an income distribution.

In the context of difference in opportunities and responsibility by the agents, where the opportunities are the set of circumstances in which agents are developed, and the responsibility refers to the efforts and actions of agents for income generation, we identify the inequality caused by the difference in opportunities as a problem, and not the one caused by the performance of the agents. Lets see it from the productivity point of view, if we face some income inequality that only depends on the performance of the agents, there will be no reason to complain at all, because we could say that every one gets what they deserve and if some individual wants to increase his income, then he would have to increase his performance, however if we start giving people a different payback depending in some other aspect such as an opportunity level, this could disincentive productivity because the agents could lose the opportunity of getting what they first wanted. Since we would like to preserve or increase productivity incentives in the economy, we try to highlight the difference in opportunities roll in the total income inequality, with this we could measure the size of such issue and therefore give a more complete analysis about an income distribution. So what we try to explain is: given a level of inequality is, how much inequality is caused due to the difference in opportunities? and, how much inequality is caused due to the variable of responsibility? In other words, this section seeks to find the level of participation of these variables in the total inequality and facilitate the analysis and measurement of the inequality in opportunities problem for a given population. To answer these questions we will use the approach of the Shapley value, looking for the advantages already set out in the coalitional games section.

To decompose the Atkinson index we will assume that the cardinality is equal between each type (in a case of homogenous type partitions), *i.e.*

$|I^1| = |I^2| = \dots = |I^{l-1}| = |I^l|$  and we define the following profiles:

1) The original income distribution:

$$X = (x_1, x_2, \dots, x_m)$$

2) Each agent gets the average of their type; this payoff vector shows the opportunity differences, because of the assumption that establishes the types averages dependency only on the opportunity level:

$$X_o = (\mu^1, \mu^2, \dots, \mu^m)$$

where  $\mu^j$  denotes the income mean of the opportunity set where  $j$  belongs.

3) A payoff vector which is a payment reescalation such that all types have the general mean (or same op-

portunity), showing the responsibility difference:

$$X_R = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m)$$

where  $\hat{x}_j = (\mu/\mu^j)x_j$  and  $\mu$  being the income mean of the original distribution  $X$ .

Now, we reach the decomposition by getting the marginal contribution of the income level determining variables (opportunity and responsibility) to the total inequality, to be able to assign the average of their marginal contributions. The last part of this process gets done by computing the Shapley value of a certain game, such that we can know the contribution of each variable.

Considering that the total inequality of the original income is given by  $1 - A(X)$ , we can take advantage of the efficiency property in the Shapley value, since we can establish the exact participation of the factors in the exact total inequality.

Formally, there are two players or factors  $N = \{O, R\}$  influencing the inequality (denoting the factors of opportunity and responsibility, respectively); so the worth of the grand coalition is  $1 - A(X)$  and it will be the amount to be distributed among the players  $O$  and  $R$ , which are the profiles that exhibit the difference in opportunities  $X_o$  and in responsibility  $X_R$ .

So the game has the next characteristic function:

$$v(S) = \begin{cases} 0, & \text{if } S = \emptyset \\ 1 - A(X), & \text{if } S = N \\ 1 - A(X_o), & \text{if } S = \{O\} \\ 1 - A(X_R), & \text{if } S = \{R\} \end{cases}$$

The next table shows the marginal contributions for the previous game and the Shapley value for it:

$\theta \in S_2$	$i = O$	$i = R$
$OR$	$v(\{O\}) = 1 - A(X_o)$	$v(\{O, R\}) - v(\{O\}) = A(X) - A(X_o)$
$RO$	$v(\{O, R\}) - v(\{R\}) = A(X) - A(X_R)$	$v(\{R\}) = 1 - A(X_R)$
$Sh_i(v)$	$v(\{O\}) + \frac{1}{2}[v(\{O, R\}) - v(\{O\}) - v(\{R\})]$	$v(\{R\}) + \frac{1}{2}[v(\{O, R\}) - v(\{O\}) - v(\{R\})]$

**Example 6** Let  $X = (6, 10, 20, 40, 42, 80)$  be the income distribution of a given population that presents two types of opportunity  $I^1 = \{1, 4, 5\}$  and  $I^2 = \{2, 3, 6\}$ . Notice that the mean for types  $I^1$  and  $I^2$  are  $88/3$  and  $110/3$ , respectively. Now, as an illustration,

$$\hat{x}_1 = \frac{\mu}{\mu^1} x_1 = \frac{33}{88/3} (6) = 27/4. \text{ Then, in summary:}$$

Profile	$A(\cdot)$
$X = (6, 10, 20, 40, 42, 80)$	0.70699468
$X_o = (88/3, 110/3, 110/3, 88/3, 88/3, 110/3)$	0.99380799
$X_r = (27/4, 9, 18, 45, 189/4, 72)$	0.711399674

where,

$$A(X) = \frac{(6 \times 10 \times 20 \times 40 \times 42 \times 80)^{1/6}}{33} = 0.70699468$$

$$A(X_o) = \frac{\left(\frac{88}{3} \times \frac{110}{3} \times \frac{110}{3} \times \frac{88}{3} \times \frac{88}{3} \times \frac{110}{3}\right)^{1/6}}{33} = 0.99380799$$

$$A(X_r) = \frac{\left(\frac{27}{4} \times 9 \times 18 \times 45 \times \frac{189}{4} \times 72\right)^{1/6}}{33} = 0.711399674$$

And the game is given by

$S$	$v(S)$
$\emptyset$	0
$\{O\}$	0.00619201
$\{R\}$	0.288600326
$\{O, R\}$	0.293005319

where for example,

$$(100, 120, 140, 160, 230, 200, 300, 310, 340, 350, 380, \dots, 390, 390, 450, 410, 433, 455, 476, 500, 600, 800)$$

and there are five opportunity types for the individuals,

$$I^1 = \{1, 2, 3, 4, 5\}, I^2 = \{6, 7, 8, 9, 10, 11, 12, 13, 14\}, I^3 = \{15, 16, 17, 18\}, I^4 = \{19, 20\}, I^5 = \{21\}$$

this based on the services they were endowed. Given this data, now it is possible to compute the Atkinson index decomposition. The game is given by

$S$	$v(S)$
$\emptyset$	0
$\{O\}$	0.099163
$\{R\}$	0.019765
$\{O, R\}$	0.11697

Finally,

$$Sh_o(v) = 0.0981827 (83.94\%)$$

$$Sh_R(v) = 0.0187848 (16.06\%)$$

And we have a greater participation of the difference

$$v(\{R\}) = 1 - A(X_r) = 0.288600326$$

Finally,

$$Sh_o(v) = 0.005298501 (1.809\%)$$

$$Sh_R(v) = 0.287706817 (98.191\%)$$

Now we have answer our two questions by giving the participation level of each variable in the total income inequality.

**Remark 7** We previously assumed that there was the same number of people in each opportunity type. To get a more realistic decomposition (considering the possibility of heterogeneous partitions), we must change the reescalation made in the profile that shows the difference in responsibility (3):

3'. Once again, this vector represents the differences in the variable of responsibility, because all of the types will have the general mean:

$$X_2 = (\bar{x}_1, \dots, \bar{x}_m)$$

where  $\bar{x}_j = \frac{\mu |I_j|}{\sum_{k \in I_j} x_k} x_j$ , denoting by  $I_j$  the

opportunity set where  $j$  belongs.

**Example 8** Somewhere in the world there exists a society of 21 people that grew up under different conditions, since not all of them had the money to afford different services or to do human capital investment. The income vector for this people is given by

in opportunity level occupying more than 80% of the total income inequality.

### 5. Concluding Remarks

Despite the fact that economic inequality has always been treated with seriousness, not all inequality is necessarily a problem. The context of equal opportunities, allows us to understand how far should we get in the compensation of economic differences. And thus takes us to a less drastic, but more strategic solutions design.

The use of the cooperative game theory, allows us to get an idea about allocations or assignments, that use the axiomatic way, with strongly natural and intuitive criteria. So the use of the Shapley value that complies with its already seen four axioms, is convincing for the realization of the inequality index decomposition.

The measurement of inequality is intended to search for the improvement of social conditions, because it tries to rank the distribution functions in terms of social welfare, so in the decomposition, can be used as a criteria because it can be used to rank income distribution functions with the same inequality level but different contribution of the opportunities variable.

The decomposition of the inequality index is very useful to rule out the views that consider the treatment of economic inequality as a policy that disincentives productivity; *i.e.*, an incentive to decrease the responsibility variable, because it defines the problem to be treated, and seeks to set aside the inequality caused by economic productivity. Therefore, operating from the framework of the difference in opportunities, we can forget about so many debates and controversy caused by the inequality treatment, highlighting the ground for the development of more effective policies that seek to solve the problem in concrete and encourage a population's productivity.

## 6. Acknowledgements

J. Sánchez-Pérez acknowledges support from CONACYT research grant 130515. V. Rosas-Martínez acknowledges support from CONACYT grant 17665. We also wish to thank the editor and an anonymous referee. All possible remaining errors and short-comings are still ours.

## REFERENCES

- [1] A. K. Sen, "The Economic Inequality," Fondo de Cultura Económica, Mexico City, 2001.
- [2] A. Bergson-Samuelson, "A Reformulation of Certain Aspects of Welfare Economics," *Journal of Economics*, Vol. 52, No. 2, 1938, pp. 310-334. [doi:10.2307/1881737](https://doi.org/10.2307/1881737)
- [3] A. B. Atkinson, "On the Measurement of Inequality," *Journal of Economic Theory*, Vol. 2, No. 3, 1970, pp. 244-263.
- [4] J. Levy, S. Chemerynski and J. Tuchmann, "Incorporating Concepts of Inequality and Inequity into Health Benefits Analysis," *International Journal for Equity in Health*, Vol. 5, 2006, p. 2.
- [5] A. F. Shorrocks, "Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value," University of Essex, Colchester, UK.
- [6] M. Sastre and A. Trannoy, "Shapley Inequality Decomposition by Factor Components: Some Methodological Issues," *Journal of Economics*, Vol. 9, No. 1, 2002, pp. 51-90. [doi:10.1007/BF03052500](https://doi.org/10.1007/BF03052500)
- [7] V. Peragine, "Measuring and Implementing Equality of Opportunity for Income," *Social Choice and Welfare*, Vol. 22, No. 1, 2004, pp. 187-210. [doi:10.1007/s00355-003-0282-2](https://doi.org/10.1007/s00355-003-0282-2)
- [8] T. Driessen, "Cooperative Games, Solutions and Applications," Springer-Verlag, Dordrecht, 1988.
- [9] L. S. Shapley, "A Value for n-person Games," In: H. W. Kuhn and A. W. Tucker, Eds., *Annals of Mathematic Studies*, Princeton University Press, Princeton, 1953, pp. 307-317.
- [10] M. C. Lasso de la Vega and A. M. Urrutia, "A New Factorial Decomposition for the Atkinson Measure," *Economics Bulletin*, Vol. 4, No. 29, 2003, pp. 1-12.