Economics Bulletin

Volume 30, Issue 1

Measuring equality of opportunity by Shapley value

Giuseppe Pignataro
Department of Economics, University of Bologna

Abstract

Equality of opportunity is a political ideal which requires that ex-ante inequalities, and only those inequalities, should be eliminated. Justice requires leveling the playing field by rendering everyone's opportunities equal in an appropriate sense, and then letting individual choices and their effects dictate further outcomes. In this paper we propose a methodology to decompose the path-independent Atkinson index of equality through Shapley value seeking for a measure of overall inequality produced by the marginal contributions of the opportunity and the responsibility component.

I would like to express my deep gratitude to Vitorocco Peragine for his encouragement and many invaluable suggestions and advice. I also wish to thank an anonymous referee, Roberto Cellini, Giuseppe Coco, Paolo Li Donni and Gilberto Turati for their insightful comments. Financial support for the Italian Ministry of Education and Science is gratefully acknowledged. Any errors and opinions expressed are the sole responsibility of the author

Citation: Giuseppe Pignataro, (2010) ''Measuring equality of opportunity by Shapley value'', *Economics Bulletin*, Vol. 30 no.1 pp. 786-798.
Submitted: Nov 03 2009. **Published:** March 25, 2010.

1 Introduction

According to the principle of equality of opportunity (EOp, hereafter), a distribution among individuals is fair if only if it requires that unchosen inequalities should be eliminated and that inequalities that arise from the choices of individuals, given equal initial conditions, should not be eliminated or reduced. So far as the Western European tradition is concerned, a society should split equally the means to reach a valuable outcome among its members. Once the set of opportunities have been equalized, which particular opportunity, the individual chooses from those open to her, is outside the scope of justice. Ex ante inequalities, and only those inequalities, should be eliminated or compensated for by public intervention.

This note represents an attempt to decompose the well-defined Shapley decomposition in accordance with the equality of opportunity principle. The procedure is developed by taking into account the "path independent property" of the multiplicative measures expressed in equality terms. Particularly, we characterize a *shapley-value* decomposition of the Atkinson index of equality with the aversion parameter $\epsilon = 1$ into an "opportunity" and a "responsibility" part. With the help of an illustration, we show that the marginal contributions produced by these two components perfectly correspond to the measurement of overall inequality in the society.

Section 2 defines the equality of opportunity model taken out from Peragine (2004) while section 3 reorganizes the modified shapley procedure in equality terms. Finally, the application of this methodology measuring equality of opportunity is proposed in section 4. Concluding remarks follow in section 5.

2 The model[1]

Given a society composed by N individuals, each individual income is a joint result of circumstances and effort. We define the circumstances as all factors for which individuals do not have control, belonging to a finite set $\Omega = \{T_1, ..., T_j, ..., T_m\}$, with $|\Omega| = m$. Instead, effort variable summarizes all factors for which individuals have some responsibility, denoted by a variable $E \in \Theta$. The value of the effort level E is not observable. Formally, individual

[1] This model is taken out from Peragine (2004).

income is generated by a function $g : \Omega \times \Theta \to \Re_+$ that assigns each individual income to combinations of effort E and circumstance T:

$$y = g(T, E) \tag{1}$$

Given the ordering defined over Ω, we can partition any population into m subgroups, each one identified by a variable $T_j \in \Omega = \{T_1, ..., T_j, ..., T_m\}$. For each $T_j \in \Omega$, we call "type-j" the set of individuals whose set of circumstances is T_j. Let N_j^Y be the number of people in type j of distribution Y, such that $\sum_{j=1}^{m} N_j^Y = N^Y$. We stand for $\mathbf{y}_j = \left\{ y_{j,1}, ..., y_{j,N_j^Y} \right\} \in \Re_+^{N_j^Y}$, the type-$j$ income distribution. Therefore, for the overall income profile $Y \in \Re_+^{N^Y}$, we consider the three following reference profiles[2]:

(a) $Y = \{\mathbf{y}_1, ..., \mathbf{y}_j, ..., \mathbf{y}_m\}$
(b) $Y_B = \left\{ \mu_{\mathbf{y}_1} 1_{N_1}, ..., \mu_{\mathbf{y}_j} 1_{N_j}, ..., \mu_{\mathbf{y}_m} 1_{N_m} \right\}$
(c) $Y_W = \{\tilde{\mathbf{y}}_1, ..., \tilde{\mathbf{y}}_j, ..., \tilde{\mathbf{y}}_m\}$

where $\mu_{\mathbf{y}_j}$ is the mean income of the type-j vector, 1_{N_j} is the unit vector of size N_j and the individual income $\tilde{y}_j$ is obtained by rescaling each type-j income vector as

$$\forall j \in \{1, ..., m\}, \forall i \in \{1, ..., N_j\},$$

$$y_j^i \to \frac{\mu}{\mu_{\mathbf{y}_j}} y_j^i$$

where μ is the mean income of the overall income profile $Y \in \Re_+^{N^Y}$. In this case, **(a)** is the overall vector of income profile, **(b)** eliminates the equality within-types and **(c)** eliminates the between-types equality. On one side, this means that by measuring the equality in the *smoothed* distribution (Y_B), obtained by replacing each income i with its type mean income $\mu_{\mathbf{y}_j}$, we capture only and fully the between-types equality, which, in turn, in the type approach reflects the equality of opportunity. On the other side by rescaling all type distributions until all types have the same mean income, we are left

[2]This procedure identifies a methodology called *the type approach* proposed by Peragine (2004). It focus on *ex-ante* inequalities between individuals with the same circumstances.

with an income profile (Y_W) which express the equality within-types. This can be interpreted as equality due to individual responsibility.

Measuring equality of opportunity through these two distributions **(b)** and **(c)** leads to different evaluations for almost all indices used in the income inequality literature. Instead, Lasso de La Vega and Urrutia (2005) introduce the "path-independent decomposable" class of multiplicative indices for which the two profiles (directly or indirectly computed) yield the same results. This class reduces to a single measure of equality when the chosen reference income is the arithmetic mean, i.e., the Atkinson index E_A with the aversion parameter $\epsilon = 1$[3], such that:

$$E_A(Y_B) = \frac{E_A(Y)}{E_A(Y_W)} \tag{2}$$

3 Shapley decomposition in equality terms

The application of the Shapley value algorithm to decompose inequality and poverty measures into between- and within- parts has been proposed by Shorrocks (1999) and by Chantreuil and Trannoy (1999) and Sastre and Trannoy (2002) for inequality measures, while a Shapley decomposition by income sources was provided by Fournier (1999). The Shapley value is a solution concept first employed in game theory with the aim to divide a given surplus among members of a coalition.

We propose a modified application of the Shapley value which takes into account a multiplicative class of indices expressed in equality terms. Particularly, due to the equality of opportunity model suggested in the previous section, we focus on the path independent Atkinson index of equality E_A taking into account v_j income-types, $\forall j \in \Omega = \{1, ..., m\}$ which can be expressed as:

$$E_A = E_A \ [v_1, ..., v_j, ..., v_m] \tag{3}$$

The decomposition principle consists in assigning contribution C_j to each one of the variable v_j, $\forall j \in \Omega = \{1, ..., m\}$, allowing the E_A variable to be

[3] A proof of the path-independent decomposition of the Atkinson index with the aversion parameter $\epsilon = 1$ is proposed in the appendix.

expressed as the sum of the factor contributions. Each contribution refers to the marginal impact when all possible elimination tracks are taken into account. We define $E_A(S)$ where $S = \Omega \backslash \{v_j\}$ as the value of the index of equality when the factor v_j has been dropped. The Shapley procedure considers the elimination of all possible variables. Therefore, in a recursively way, the set S is the domain of variables remaining after the Shapley process of elimination has been developed and s is its cardinality, i.e., the number of variables remaining after the successive eliminations. The contribution of the j-variable to the inequality index is given by:

$$C_j(\Omega, I) = \left| \sum_{s=0}^{m-1} \sum_{S \subseteq J \backslash \{v_j\}} \frac{s!(m-s-1)!}{m!} \left[E_A\left(S \cup \{v_j\}\right) - E_A(S)\right] \right| \tag{4}$$

such that

$$\sum_{j=1}^{m} C_j(\Omega, I) = I_A = (1 - E_A) \tag{5}$$

where I_A is the Atkinson index of inequality. As proposed above in the computation of the Atkinson index of equality through the Shapley procedure, we assume that all mean incomes in the *smoothed* distribution (Y_B) are different between each other. To express the equality between types, we liken the income distribution where the mean incomes are different $(\mu_j \neq \mu)$ to the distribution when they are equal to the overall mean income $(\mu_j = \mu)$. As regards instead the equality within types, a comparison is developed between the distribution in which the individual incomes within a specific type differ from the mean income $(y_{ij} \neq \mu_j)$ and the distribution where all incomes of the type are equal $(y_{ij} = \mu_j)$[4].

[4]Such decomposition refers to the case of homogenous type partitions. The heterogeneous case can be developed adding the marginal contributions suggested by different sizes between types. If instead, we consider the possibility of heterogeneous partition, then the impact of differences in size must be analyzed. In the latter case the analysis further requires the comparison among the situation in which the types have different sizes $(s_j \neq 1/m)$, $\forall j \in \{1, ..., m\}$, to the case where all types have the same dimension $(s_j = 1/m)$.

4 Measuring equality of opportunity

Let us take as a simple illustration the case where there are six individual incomes $(2, 8, 16, 30, 40, 60)$, where type A includes the incomes $(8, 16, 60)$ and type B $(2, 30, 40)$. The mean income of type A is 28 and that of type B is 24. The average income μ in this population is 26. Such analysis is characterized by 4 steps.

1) Applying the Shapley value on the income profile representation in the type strategy, we look at the reference income vectors showed in the previous section. Starting by the overall income vector $Y = \{y_1, ..., y_j, .., y_m\}$, we can express $E_A(\mu_j \neq \mu;\ y_{ij} \neq \mu_j)$ as the overall Atkinson index of equality. It refers to the case where the mean incomes of the types are different $\left(\mu_j \neq \mu\right)$ and the individual incomes within a given partition differ from the mean income $\left(y_{ij} \neq \mu_j\right)$. It follows that:

$$E_A(\mu_j \neq \mu; y_{ij} \neq \mu_j) = E_A\left(2, 8, 16, 30, 40, 60\right) = \frac{\left(\prod_{i=1}^{N} y_i\right)^{\frac{1}{N}}}{\mu} =$$

$$= \frac{(2 \times 8 \times 16 \times 30 \times 40 \times 60)^{\frac{1}{6}}}{26} = \frac{(18432000)^{\frac{1}{6}}}{26} = 0,6251 \qquad (6)$$

2) As a second step, we can describe $Y_B = \left\{\mu_{\mathbf{y}_1} 1_{N_1}, ..., \mu_{\mathbf{y}_j} 1_{N_j}, ..., \mu_{\mathbf{y}_m} 1_{N_m}\right\}$ where $\mu_{\mathbf{y}_j} 1_{N_j}$ is the mean income of the type-j income vector. $E_A(\mu_j \neq \mu;\ y_{ij} = \mu_j)$ reflects the equality in the distribution where the within-types equality has been eliminated and the mean incomes of the types are different[5].

$$E_A(\mu_j \neq \mu; y_{ij} = \mu_j) = E_A\left(24, 28, 28, 24, 24, 28\right) = \frac{\left(\prod_{i=1}^{N} y_i\right)^{\frac{1}{N}}}{\mu} =$$

$$= \frac{(24 \times 28 \times 28 \times 24 \times 24 \times 28)^{\frac{1}{6}}}{26} = \frac{(303464448)^{\frac{1}{6}}}{26} = 0,997 \qquad (7)$$

3) In this step, we compute the Atkinson index of equality for the income profile $Y_W = \left\{\tilde{\mathbf{y}}_1, ..., \tilde{\mathbf{y}}_j, ..., \tilde{\mathbf{y}}_m\right\}$. $E_A(\mu_j = \mu; y_{ij} \neq \mu_j)$ refers to the case where there is within-types equality. In this case, the mean type incomes are equal. This means that each original income is multiplied by the ratio $\frac{\mu}{\mu_{\mathbf{y}_j}} y_j^i$, $\forall j \; \epsilon \left\{1, ..., m\right\}, \forall h \; \epsilon \left\{1, ..., N_j\right\}$[6]:

$$E_A(\mu_j = \mu; y_{ij} \neq \mu_j) = E_A\left((2 \times 26/24)\,; (8 \times 26/28)\,; ...\right) =$$

[5] It is interesting to point out that we obtain the same result if we directly apply to the income profile Y_B, the between-group component E_A^B derived from the path independent multiplicative decomposition of the Atkinson index of equality E_A:

$$E_A^B\left(24, 28, 28, 24, 24, 28\right) = \frac{\prod_{j=1}^{m} \left(\mu_j\right)^{\frac{N_j}{N}}}{\mu} =$$

$$= \frac{(24)^{\frac{1}{2}} \times (28)^{\frac{1}{2}}}{26} = \frac{[4,8989 \times 5,2915]}{26} = 0,997$$

[6] The same result is obtained if we directly apply to the income profile Y_W, the within-group component E_A^W derived from the path independent multiplicative decomposition of the Atkinson index of equality E_A :

$$E_A^W\left[2, 8, 16, 30, 40, 60\right] = \prod_{j=1}^{m} E_j^{\frac{N_j}{N}} = \prod_{j=1}^{m} \left[\left(\prod_{i=1}^{N_j} y_{ji}\right)^{\frac{1}{N_j}} / \mu_j\right]^{\frac{N_j}{N}} =$$

$$= \left\{\left[\frac{(8 \times 16 \times 60)^{\frac{1}{3}}}{28}\right]^{\frac{1}{2}} \left[\frac{(2 \times 30 \times 40)^{\frac{1}{3}}}{24}\right]^{\frac{1}{2}}\right\} = 0,626$$

$$= \frac{\left(\prod_{i=1}^{N} y_i\right)^{\frac{1}{N}}}{\mu} = \frac{[2,166 \times 7,428 \times 14,857 \times 32,5 \times 43,33 \times 55,71]^{\frac{1}{6}}}{26} = 0,626 \quad (8)$$

4) As a last step, we compute $E_A(\mu_j = \mu; y_{ij} = \mu_j)$ represents the profile where the individual incomes within a given type are equal. In this case it is also assumed that the mean incomes of the different types are equal. It should then be clear that in this distribution all the incomes are equal to the overall mean income such that:

$$E_A(\mu_j = \mu; y_{ij} = \mu) = E_A\,(26, 26, 26, 26, 26, 26) =$$

$$= \frac{\left(\prod_{i=1}^{N} y_i\right)^{\frac{1}{N}}}{\mu} = \frac{(26 \times 26 \times 26 \times 26 \times 26 \times 26)^{\frac{1}{6}}}{26} = \frac{(308915776)^{\frac{1}{6}}}{26} = 1 \quad (9)$$

According to the Shapley value, let us now compute the contributions to the overall equality of the within groups equality C_{Wtype} such that:

$$C_{Wtype} = \left| \begin{array}{c} \frac{1}{2}\left[E_A(\mu_j \neq \mu; y_{ij} \neq \mu_j) - E_A(\mu_j \neq \mu; y_{ij} = \mu_j)\right] \\ +\frac{1}{2}\left[E_A(\mu_j = \mu; y_{ij} \neq \mu_j) - E_A(\mu_j = \mu; y_{ij} = \mu)\right] \end{array} \right|$$

$$C_{Wtype} = \left| \frac{1}{2}\left\{[0,6251 - 0,997] + [0,626 - 1]\right\} \right| = 0,37295 \quad (10)$$

This represents the impact of equality in the within component of the Atkinson index of equality. On the other side, the between component C_{Btype} is given by:

$$C_{Btype} = \left| \begin{array}{c} \frac{1}{2}\left[E_A(\mu_j \neq \mu; y_{ij} \neq \mu_j) - E_A(\mu_j = \mu; y_{ij} \neq \mu_j)\right] \\ +\frac{1}{2}\left[E_A(\mu_j \neq \mu; y_{ij} = \mu_j) - E_A(\mu_j = \mu; y_{ij} = \mu)\right] \end{array} \right|$$

$$C_{Btype} = \left| \frac{1}{2}\left\{[0,6251 - 0,626] + [0,997 - 1]\right\} \right| = 0,00195 \quad (11)$$

Summing eq. (10) and eq. (11), the overall impacts of equality will be equal to:

$$C_{Wtype} + C_{Btype} = 0,37295 + 0,00195 = 0,3749 \quad (12)$$

which perfectly corresponds to the Atkinson index of inequality $I_A = (1 - E_A)$ such that:

$$I_A = (1 - E_A) = (1 - 0,6251) = 0,3749 \tag{13}$$

5 Concluding remarks

In this paper we provide an application of the *Shapley-value* procedure in accordance with the equality of opportunity principle. In the case of decomposition by population subgroups, we focus on the between- and within-components which refer to the opportunity and the responsibility factors directly computed. Rather different than the traditional decomposition, the opportunity egalitarian decomposition using the Shapley value produces different marginal contributions between- and within- types. Summing up these two contributions perfectly reflect the overall inequality in the society. This clearly represents a simplification of the model in the homogenous type case. Further development in this direction can be addressed taking into account differences in size among types. Some extensions in these directions can be left for future research.

A The Appendix

On the basis of the "path-independent decomposable" class of multiplicative equality measures suggested by Lasso de La Vega and Urrutia (2005), it follows that the Atkinson index of equality E_A is given by:

$$E_A = \frac{\left[\prod_{i=1}^{N} y_i\right]^{\frac{1}{N}}}{\mu} \tag{14}$$

Let E_A^B be the between-group component which can be interpreted as the equality associated with a population of m types:

$$E_A^B = \frac{\prod_{j=1}^{m} (\mu_j)^{p_j}}{\mu} \tag{15}$$

where $p_j = N_j/N$ is the population share. Hence, we define the Atkinson's equality index within type-j E_j as follows:

$$E_j = \frac{\left[\prod_{i=1}^{N_j} y_{ji}\right]^{\frac{1}{N_j}}}{\mu_j} \tag{16}$$

From eq. (14) and eq. (16), the Atkinson index of equality E_A can be further portrayed into:

$$E_A = \frac{\left[\prod_{i=1}^{N} y_i\right]^{\frac{1}{N}}}{\mu} = \frac{\prod_{j=1}^{m}\left[\frac{\left(\prod_{i=1}^{N_j} y_{ji}\right)^{\frac{1}{N_j}}}{\mu_j}\mu_j\right]^{p_j}}{\mu} = \frac{\prod_{j=1}^{m}(E_j\mu_j)^{p_j}}{\mu} \tag{17}$$

By construction, the within part of the Atkinson index of equality is defined as the product of the equality index for each type-j:

$$E_A^W = \frac{E_A}{E_A^B} = \frac{\prod_{j=1}^{m} \left(E_j \mu_j\right)^{p_j}}{\mu} \frac{\mu}{\prod_{j=1}^{m} \left(\mu_j\right)^{p_j}} = \prod_{j=1}^{m} E_j^{p_j} \tag{18}$$

Finally, it follows that:

$$E_A = E_A^W E_A^B \tag{19}$$

References

[1] Atkinson, A.B. (1970), "On the measurement of inequality", *Journal of Economic Theory*, **2**, 244–263

[2] Blackorby, C., Donaldson D., Auersperg M. (1981), "A new procedure for the measurement of inequality within and among population subgroups", *Canadian Journal of Economics*, **14**, 665–685

[3] Chantreuil, F., and Trannoy, A. (1999), "Inequality Decomposition Values: The Trade-off Between Marginality and Consistency", Working Paper 9924, THEMA.

[4] Checchi, D. and Peragine, V. (2009), "Inequality of opportunity in Italy", *Journal of Economic Inequality*, forthcoming

[5] Cowell, F.A. (2000), "Measurement of inequality". In: Atkinson, A.B., Bourguignon, F. (eds.) Handbook of Income Distribution, pp. 87–166. North-Holland, Amsterdam, The Netherlands

[6] Deutsch, J. and J. Silber (1999), "Inequality Decomposition by Population Subgroups and the Analysis of Interdistributional Inequality", in J. Silber, editor, Handbook on Income Inequality Measurement, Kluwer Academic Press, Dordrecht and Boston, 363-397.

[7] Foster, J., Shneyerov, A. (1999), "A general class of additively decomposable inequality measures", *Economic Theory,* **14**, 89–111

[8] Foster, J., Shneyerov, A. (2000), "Path independent inequality measures". *Journal of Economic Theory*, **91**, 199–222

[9] Lasso de la Vega, C., Urrutia, A. (2003), "A new factorial decomposition for the Atkinson measure", *Economics Bulletin*, **4** (29), 1–12

[10] Lasso de la Vega, C., Urrutia, A. (2005), "Path independent multiplicatively decomposable inequality measures", *Investigaciones. Economicas,* XXIX(2), 379–387

[11] Peragine V. (2004), "Measuring and implementing equality of opportunity for income", *Social Choice and Welfare*, **22**, 1-24

[12] Roemer, J.E. (1998), *Equality of opportunity*, Cambridge, MA: Harvard University Press.

[13] Sastre M. and A. Trannoy (2002), "Shapley inequality decomposition by factor components: Some methodological issues", in P. Moyes, C. Seidl and A.F. Shorrocks (eds.), Inequalities: Theory, Experiments and Applications, *Journal of Economics,* **9**, 51-90

[14] Shorrocks, A.F. (1984), "Inequality decomposition by factor components", *Econometrica,* **50**, 193-211

[15] Shorrocks, A.F. (1984), "Inequality decomposition by population subgroups", *Econometrica,* **52**, 1369–1385

[16] Silber, J., (1989), "Factors Components, Population Subgroups and the Computation of the Gini Index of Inequality", *The Review of Economics and Statistics*, LXXI:107-115.