Working Paper Series

Fairness and Gini decomposition

Domenico Moramarco

ECINEQ 2023 650

ECINEQ 2023 650
2023 June
www.ecineq.org

Fairness and Gini decomposition

Domenico Moramarco
University of Bari - DiEF

Abstract

We propose a new decomposition of the Gini coefficient inspired by the opportunity egalitarian paradigm. Under the assumption that income is function of circumstances out of individual control, effort and unobservable factors, we identify fair and unfair inequalities as components of the total inequality. Our nine-term decomposition of the Gini coefficient nests a structuralist and an individualistic definition of unfair inequality, as well as ex ante and ex post measures of inequality of opportunity, reflecting different weights assigned to circumstances and effort. We illustrate the new decomposition on Belgian data, highlighting the source of a consistent difference between the structuralist and individualistic views about unfair inequality. This result shows the existence of room for new normative principles to close the wedge between these two views.

Keyword: equality of opportunity, Gini decomposition, individualism, structuralism

JEL Cassification: D63, D31

Fairness and Gini decomposition *

Domenico Moramarco †

June 29, 2023

Abstract

We propose a new decomposition of the Gini coefficient inspired by the opportunity egalitarian paradigm. Under the assumption that income is function of circumstances out of individual control, effort and unobservable factors, we identify fair and unfair inequalities as components of the total inequality. Our nine-term decomposition of the Gini coefficient nests a structuralist and an individualistic definition of unfair inequality, as well as ex ante and ex post measures of inequality of opportunity, reflecting different weights assigned to circumstances and effort. We illustrate the new decomposition on Belgian data, highlighting the source of a consistent difference between the structuralist and individualistic views about unfair inequality. This result shows the existence of room for new normative principles to close the wedge between these two views.

Keywords: equality of opportunity, Gini decomposition, individualism, structuralism

1 Introduction

Since the seminal contributions of Rawls (1971), Sen (1980) and Dworkin (1981a,b), we have witnessed the development of a view of social justice that distinguishes inequalities due to individual responsibility from those which are not. This paradigm, known as Equality of Opportunity (EOp),

*This paper makes use of the MEqIn dataset, collected by a team of researchers from Université catholique de Louvain, KU Leuven, Université libre de Bruxelles, and University of Antwerp. The collection of the MEqIn data was enabled by the financial support of the Belgian Science Policy Office (BELSPO) through grant BR/121/A5/MEQIN (BRAIN MEqIn).

†DiEF, University of Bari "Aldo Moro" and ECARES, Univeristé Libre de Bruxelles. E-mail: domenico.moramarco@uniba.it. I gratefully aknowledge the GRINS foundation for the financial support. I thank Pedro Salas-Rojo, Flaviana Palmisano and Annaelena Valentini for their comments.

claims that a fair society should remove inequalities due to factors out of individual control (circumstances) while preserving those due to individual responsibility or conscious actions (effort) (see Roemer & Trannoy, 2015; Ramos & Van de gaer, 2016; Ferreira & Peragine, 2015, for recent surveys) The exact definition of what is circumstances and what is effort remains an open question in the literature: one which we don't address in this paper.

In a deterministic model where income is assumed to be increasing function of circumstances and effort alone, Sarkar (2023) compares individuals in terms of relative deprivation (RD). In a nutshell, suppose that an individual a has lower income than b, then a is relatively deprived with respect to b. The RD of a with respect to b is their income difference (in absolute value), while the RD of b with respect to a is normalized to zero. The RD of a with respect to b is considered fair if a has lower effort and better circumstances than b. Vice versa, the same RD is unfair if a has worse circumstances and higher effort than b. Since the Gini coefficient can be expressed as the normalized sum of all pairwise RDs, Sarkar (2023) proposes a three terms decomposition of the Gini into fair, unfair and residual inequality.

In this paper we refine this decomposition, and apply it to a model which is closer to the reality because it allows for individuals with the same circumstance and effort to have different income. Lefranc *et al.* (2009) attribute this event to luck but, more in general, this can be due to errors in the assumption about the income generating process or the existence of unobserved factors (effort or circumstances). Our refinement is motivated by the interest for isolating inequality among individuals with same circumstances (types) and inequality among those with same effort (tranches). In the opportunity egalitarian paradigm, the former (resp. latter) can be fair (resp. unfair) depending on the approach followed: ex ante or ex post. We define two strong measures of ex ante and ex post inequality of opportunity (IOp), which we can interpret, respectively, as upper and lower bounds for unfair inequality. We also distinguish between structuralist and individualistic unfair inequality (see Mayhew, 1980a,b, for a discussion), where the former emphasizes the role of circumstances in the income generating process, while the latter focuses on individual effort as justification for different income.

We illustrate our decomposition on data from Belgian in 2016, to study the weight each of the nine components has on total inequality. Since our decomposition nests the one in Sarkar (2023), this paper is also the first application of Sarkar's proposal. We show a consistent wedge between structuralist and individualistic inequality, due to what we call ambiguous inequality. This result underlines the potential need and room for stronger normative principles to characterize an idea of social justice that balances these two approaches.

The rest of the paper is organized as follows. Section 2 delineates the model, introduces necessary notation and formalizes our Gini decomposition. Section 3 describes the data and performs the empirical exercise. Section 4 concludes.

2 The model

Let the function $f : \mathbb{N} \times \mathbb{N} \times \mathbb{R} \to \mathbb{R}_{++}$, increasing in all its arguments, represent the assumed[1] income generating process, such that $f(c, e, u)$ is the income of an individual with observed circumstances $c \in \{1, ..., \bar{c}\}$, observed effort $e \in \{1, ..., \bar{e}\}$, and random component u capturing unobserved factors. Individuals are assumed to be responsible for their effort but not for their circumstances.

For each pair $(j, k) \in \{1, ..., \bar{c}\} \times \{1, ..., \bar{e}\}$, let $\mathbf{x}_{jk} = (f(j, k, u_p))_{p=0}^{n_{jk}}$ be the income realizations of individuals with characteristics (j, k) in the population, with $\sum_{j=1}^{\bar{c}} \sum_{k=1}^{\bar{e}} n_{jk} = n$. For simplicity, let $f(j, k, u_p) = x_{jk}^p$ and denote with $X = (\mathbf{x}_{jk})_{j,k=1}^{\bar{c},\bar{e}}$ the income distribution.

Define the relative deprivation (RD) of an individual p with characteristics (j, k) against individual i with characteristics (c, e) as $RD_{jk}^p(c, e, i) = \max\left\{x_{ce}^i - x_{jk}^p, 0\right\}$. Then, the Gini index of the income distribution X can be defined as the average of all RDs, normalized by the total income. Formally,

$$G = \frac{\sum_{j=1}^{\bar{c}} \sum_{k=1}^{\bar{e}} \sum_{c=1}^{\bar{c}} \sum_{e=1}^{\bar{e}} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD_{jk}^p(c, e, i)}{n^2 \mu}$$

where μ is the average income in X and $RD_{jk}^p(c, e, i)$ is defined as above.

In this model we can formulate two intuitive fairness principles based, respectively, on equity and meritocracy. RDs related to worse circumstances are unfair, while RDs related to lower effort are fair. Clearly, partial observability of effort and circumstances does not allow us to establish causal links with income. Nevertheless, suppose an individual has bad circumstances and is poor despite exerting high effort then, conditional on the observables, we cannot justify his low income on the ground of low effort. Of course, one may emphasize the role of unobservable factors, among which we should also include empirical misspecification of the income generating process, to claim that we should remain agnostic about any RD. We find this view unappealing, as it would prevent us from conducting distributive analysis.

[1] f is not necessarily the *true* income generating process, rather it is the one assumed/inferred by the evaluator given the available information.

2.1 Classifying relative deprivation

Suppose that $x^p_{jk} < x^i_{ce}$, so that (j,k,p) is relatively deprived with respect to (c,e,i), that is $RD^p_{jk}(c,e,i) > 0$. This RD may be caused by observed circumstances, observed effort or unobservable factors. Depending on this, we classify (j,k,p)'s relative deprivation in nine possible ways.

We say that $RD^p_{jk}(c,e,i)$ is *unfair* (U) if (j,k,p) has worse circumstances and higher effort (i.e. $j < c$ and $k > e$). In line with our equity principle, this RD is induced by less favourable circumstances because the higher effort does not compensate for them. The same RD is *fair* (F) if (j,k,p) has better circumstances and lower effort (i.e. $j > c$ and $k < e$) because, in line with the meritocratic view, lower effort justifies lower income, particularly in presence of better circumstances.

An $RD^p_{jk}(c,e,i) > 0$ is *ambiguous* (A) if (j,k,p) has worse circumstances and lower effort (i.e $j < c$ and $k < e$). In this case one cannot establish whether deprivation is caused by bad circumstances or low effort. In a model like ours, where income increasing in both effort and circumstances, this is likely to be a relevant class of RDs. A meritocratic or individualistic view is likely to emphasize the normative relevance of effort and consider it fair. On the other side, a structuralist view may condemn the relative deprivation of individuals that already suffer from bad circumstances, which may also impact other life aspects, and include this RD among the unfair ones.

Moreover, $RD^p_{jk}(c,e,i) > 0$ is *erratic* (E) if (j,k,p) has better circumstances and higher effort (i.e. $j > c$ and $k > e$). Indeed, net of the possible influence of unobservable factors like "bad luck", this RD is not explainable by the assumed income generating process.

We say that $RD^p_{jk}(c,e,i)$ is *residual* (R) if (j,k,p) has same circumstances and same effort (i.e $j = c$ and $k = e$). This is the part of inequality that is clearly due to unobservable factors.

The next four classifications are of interest in light of the distinction, in the EOp literature, between ex ante and ex post approaches. The ex ante approach tends to express neutrality with respect to inequality within types. The ex post approach, instead, focuses on individuals with the same effort to claim that they should obtain equal incom (see Fleurbaey & Peragine, 2013, for a deep discussion). As we will see, these two interpretations of EOp are not always in line with the equity and meritocracy principles.

Let RD be *anti-meritocracy* (AM) if both have the same circumstances but (j,k,p) has higher effort (i.e. $j = c$ and $k > e$). This RD contradicts the meritocratic idea that higher effort justifies better outcomes. Conversely, RD is *pro-meritocracy* (PM) if both have the same circumstances but (j,k,p) has lower effort (i.e. $j = c$ and $k < e$). In the ex-ante approach, those two category define the inequality which has no influence on EOp.

We say that $RD^p_{jk}(c,e,i)$ is *pro-equity* (PE) if they both have the same effort but (j,k,p) has better circumstances (i.e. $j > c$ and $k = e$). In this case, the unobserved factors are compensating the effect of better circumstances. Conversely, we will have an *anti-equity* (AE) RD if they both have the same effort but (j,k,p) has worse circumstances (i.e. $j < c$ and $k = e$). In the ex post approach, those two category define the inequality which reduces EOp.

2.2 A decomposition of the Gini coefficient

Summing up the nine types of RD from the previous section we obtain a perfect decomposition of the Gini coefficient. Formally,

$$G = G_U + G_F + G_A + G_R + G_{AM} + G_{PM} + G_{PE} + G_{AE} + G_E \quad (1)$$

where

$$\begin{aligned}
\left(n^2\mu\right) G_U &= \textstyle\sum_{j=1}^{\bar{c}-1} \sum_{k=2}^{\bar{e}} \sum_{c=j+1}^{\bar{c}} \sum_{e=1}^{k-1} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(c,e,i) \\
\left(n^2\mu\right) G_F &= \textstyle\sum_{j=2}^{\bar{c}} \sum_{k=1}^{\bar{e}-1} \sum_{c=1}^{j-1} \sum_{e=k+1}^{\bar{e}} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(c,e,i) \\
\left(n^2\mu\right) G_A &= \textstyle\sum_{j=1}^{\bar{c}-1} \sum_{k=1}^{\bar{e}-1} \sum_{c=j+1}^{\bar{c}} \sum_{e=k+1}^{\bar{e}} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(c,e,i) \\
\left(n^2\mu\right) G_R &= \textstyle\sum_{j=1}^{\bar{c}} \sum_{k=1}^{\bar{e}} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{jk}} RD^p_{jk}(j,k,i) \\
\left(n^2\mu\right) G_{AM} &= \textstyle\sum_{j=1}^{\bar{c}} \sum_{k=2}^{\bar{e}} \sum_{e=1}^{k-1} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(j,e,i) \\
\left(n^2\mu\right) G_{PM} &= \textstyle\sum_{j=1}^{\bar{c}} \sum_{k=1}^{\bar{e}-1} \sum_{e=k+1}^{\bar{e}} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(j,e,i) \\
\left(n^2\mu\right) G_{PE} &= \textstyle\sum_{j=2}^{\bar{c}} \sum_{k=1}^{\bar{e}} \sum_{c=1}^{j-1} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(c,k,i) \\
\left(n^2\mu\right) G_{AE} &= \textstyle\sum_{j=1}^{\bar{c}-1} \sum_{k=1}^{\bar{e}} \sum_{c=j+1}^{\bar{c}} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(c,k,i) \\
\left(n^2\mu\right) G_E &= \textstyle\sum_{j=2}^{\bar{c}} \sum_{k=2}^{\bar{e}} \sum_{c=1}^{j-1} \sum_{e=1}^{k-1} \sum_{p=1}^{n_{jk}} \sum_{i=1}^{n_{ce}} RD^p_{jk}(c,e,i)
\end{aligned}$$

This is a granular decomposition of the Gini coefficient which nests the one of Sarkar (2023), where the fair inequality is $G_F + G_{PM}$, the unfair inequality is $G_U + G_{AE}$ and the residual is $G_A + G_{AM} + G_{PE}$.[2]

The ex ante unfair inequality

$$G_{ex-ante} = G - (G_{AM} + G_{PM}) \quad (2)$$

is obtained as the inequality remaining after removing the differences between individuals with the same circumstances: the only tolerable inequality in this strong version of ex ante EOp. Equation (2) provides an upper bound for the inequality between type's average, often used as

[2]In the theoretical model of Sarkar (2023), G_E and G_R are zero by construction.

measure of ex ante EOp (for example Checchi & Peragine, 2010). This is because, our measure accounts for the whole income distribution within a type, rather than its expectation alone (see also Dagum, 1998).

The ex post unfair inequality

$$G_{ex-post} = G_{AE} + G_{PE} \tag{3}$$

considers only the differences between individuals with the same effort, remaining silent about how to treat differences in incomes at different levels of effort. Instead, the individualistic unfair inequality

$$G_{ind} = G_U + G_{AM} + G_{PE} + G_{AE} \tag{4}$$

augments equation (3) with the differences across circumstance groups that violate meritocracy.

The structuralist unfair inequality

$$G_{str} = G_U + G_A + G_{AE} \tag{5}$$

considers all the relative deprivation suffered by individuals with worse circumstances. As we can see, the unfair RDs are condemned in both structuralist and individualistic approaches. This is because, contrary to $G_{ex-ante}$ and $G_{ex-post}$, they perform comparison of individuals with different circumstances or effort. G_{str} includes the ambiguous RDs because it emphasizes the negative impact of circumstances on income, while Ambiguous RDs are acceptable under the meritocratic approach that emphasizes the role of effort.

3 Application

In this section we illustrate the above decomposition using the Measuring Equivalent Income (MEqIn) data, collecting information on individual income, time-use, and other aspects of life for people above 18 years old in Belgium 2016.

We focus on individual income as outcome of interest, with a sample of 2,791 observations. We have information on gender, age, country of birth, parents education and chronic diseases, and we use a data driven technique, Conditional Inference Trees (see Brunori *et al.*, forthcoming 2023; Salas-Rojo & Rodríguez, 2022, for recent applications), to define a partition of the population in types. We obtain 9 groups, which are ranked according to expected income. Effort groups are defined in two ways. The first one uses Conditional Inference Trees with education and time dedicated to paid work as input; the algorithm identifies 6 realizations of effort which, again, are ranked according to expected income. The second methodology

Figure 1: Decomposition - share of total inequality

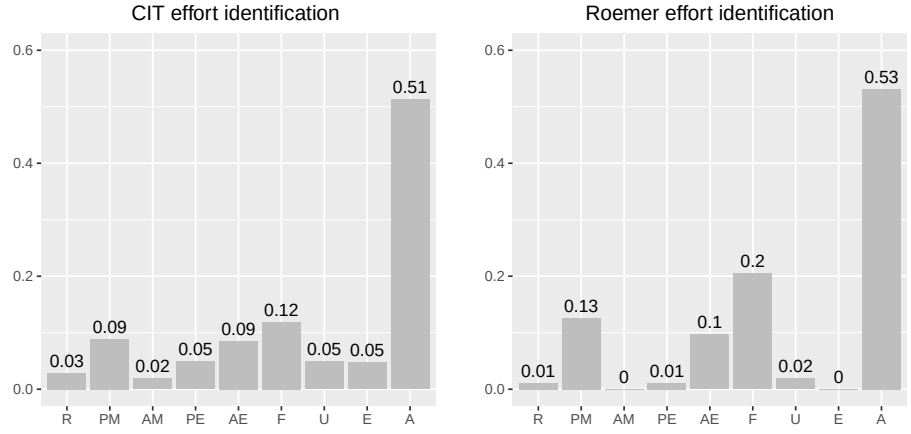

Table 1: Unfair inequalities

Gini Index	CIT effort	Roemerian effort
Structuralist	0.168	0.168
Individualistic	0.053	0.033
Ex ante	0.231	0.227
Ex post	0.035	0.028

identifies effort via the quantile (from 1 to 5) individuals occupy in they type's income distribution; this approach, proposed by Roemer (1998), is now standard in the literature and overcomes the implicit assumption of separability between effort and circumstances in the first approach.

Figure 1 reports the incidence of each component on the Gini coefficient (0.26). As expected, changing the effort identification strategy changes most of the components of our decomposition. The reader may notice that AM in the second panel is zero by construction, while the lower PE is in line with the lower residual RDs induced by Roemer's identification strategy. Nevertheless, in both cases, ambiguous RDs account for more than half of the total inequality. This is at the basis of the disparity between G_{str} and G_{ind}, and between $G_{ex-ante}$ and $G_{ex-post}$ that are more extreme views. Table 1 reports the four Gini coefficients introduced in the previous section, as expected we have $G_{ex-ante} \geq G_{str} \geq G_{ind} \geq G_{ex-post}$, with results that are fairly stable across the two effort identification strategies. Interestingly, the inequality between type averages, which is the most common measure of ex ante IOp, is 0.112 in our sample.

4 Conclusion

Using the concept of relative deprivation, we proposed a new decomposition of the Gini coefficient into nine sources of inequality, and combined them to identify four measures on unfair inequality, each reflecting different weights assigned to circumstances and effort. Our illustration on Belgian data showed that the wedge between the four measures is mainly driven by what we called ambiguous inequality, that are relative deprivations in correspondence of both lower effort and worse circumstances. While this result is qualitatively not surprising in a model where income is increasing function of both effort and circumstances, the fact that ambiguous inequality account for more than half of the total income inequality in Belgium underlines the potential need and room for stronger normative principles to define an idea of social justice that balances structuralist and individualistic views.

References

Brunori, Paolo, Hufe, Paul, & Mahler, Daniel. forthcoming 2023. The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests. *Scandinavian Journal of Economics*.

Checchi, Daniele, & Peragine, Vito. 2010. Inequality of opportunity in Italy. *The Journal of Economic Inequality*, **8**(4), 429–450.

Dagum, Camilo. 1998. A new approach to the decomposition of the Gini income inequality ratio. *Pages 47–63 of: Income inequality, poverty, and economic welfare*. Springer.

Dworkin, Ronald. 1981a. What is equality? Part 1: Equality of welfare. *Philosophy & public affairs*, 185–246.

Dworkin, Ronald. 1981b. What is equality? Part 2: Equality of resources. *Philosophy & public affairs*, 283–345.

Ferreira, Francisco HG, & Peragine, Vito. 2015. *Equality of opportunity: Theory and evidence*. The World Bank.

Fleurbaey, Marc, & Peragine, Vito. 2013. Ex ante versus ex post equality of opportunity. *Economica*, **80**(317), 118–130.

Lefranc, Arnaud, Pistolesi, Nicolas, & Trannoy, Alain. 2009. Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France. *Journal of Public Economics*, **93**(11-12), 1189–1207.

Mayhew, Bruce H. 1980a. Structuralism versus individualism: Part 1, shadowboxing in the dark. *Social Forces*, **59**(2), 335–375.

Mayhew, Bruce H. 1980b. Structuralism versus individualism: Part II, ideological and other obfuscations. *Soc. F.*, **59**, 627.

Ramos, Xavier, & Van de gaer, Dirk. 2016. Approaches to inequality of opportunity: Principles, measures and evidence. *Journal of Economic Surveys*, **30**(5), 855–883.

Rawls, John. 1971. *A theory of justice*. Harvard Press, Cambridge.

Roemer, John E. 1998. *Theories of distributive justice*. Harvard University Press.

Roemer, John E, & Trannoy, Alain. 2015. Equality of opportunity. *Pages 217–300 of: Handbook of income distribution*, vol. 2. Elsevier.

Salas-Rojo, Pedro, & Rodríguez, Juan Gabriel. 2022. Inheritances and wealth inequality: a machine learning approach. *The Journal of Economic Inequality*, **20**(1), 27–51.

Sarkar, Sandip. 2023. Gini decomposition: An inequality of opportunity perspective. *Economics Letters*, 110975.

Sen, Amartya. 1980. Equality of What? *The Tanner Lecture on Human Values*, **1**, 197–220.