



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

**The evaluation of gender income
inequality by means of the Gini
index decomposition**

Michele Costa

Quaderni - Working Paper DSE N°1130



The evaluation of gender income inequality by means of the Gini index decomposition

Michele Costa

Abstract This paper proposes to measure and to evaluate gender gaps and gender inequalities by means of the decomposition of an inequality measure. A three-terms decomposition of the Gini index is applied, thus allowing to take into account also the role of overlapping between female and male subpopulations. We develop an unified framework for the evaluation of gender gap, linking traditional measures, based on subgroups income means, to the approach related to inequality decomposition, and showing how overlapping component represents a key issue in gender gap analysis. An analysis of the income distribution of the Italian households shows how gender gaps represent a major source of inequality, without particular improvements during the last 20 years.

Key words: Gender gap, Gender income inequality, Inequality decomposition, Gini index

JEL classification: D63; J1; J7

Michele Costa
Department of Economics, University of Bologna, e-mail: michele.costa@unibo.it

Non-technical summary

Gender inequalities and gender gaps are a worldwide concern and represent the core of uncountable actions and policies developed by either governments and institutions. A growing literature aims at the evaluation of gender gap and of its effects, with an impressive escalation of contributions and proposals during the last years.

The existence and the relevance of gender gap originate from the income mean of the female subpopulation being strongly and systematically lower than the income mean of the male subpopulation. The immediate consequence was to measure the gender gap on the basis of the difference between the income means.

This an oversimplification, because, as well known, income distributions are characterized by a strong skewness and a right fat tail which prevent the use of the mean as a synthetic indicator. Consequently all mean-based methods are not suited for the analysis of income distributions. The same argument holds when some symmetric distribution, as the gaussian, is required.

A measurement system of income gender gap based on differences not only in average incomes, but in income female and male distributions can be developed from the decomposition of an inequality indicator. In particular, the Gini index, as a function of all pairwise differences between the units, allows to correctly evaluate all distributional characteristics, skewness in primis, and to take into account the presence of an heavy right tail.

The income distributions of female and male subpopulations differ in many aspects, but also share a strong overlap. Regarding this aspect, it is relevant to chose, among the many proposals, a Gini index decomposition which explicitly considers overlap. Our main contribution is to evaluate the overlap between male and female distributions and to include also this element in the gender gap analysis.

We show that, in absence of overlapping between the female and male income distributions, the subgroup means are fully informative and how their difference, appropriately weighted, leads to the same evaluation of gender gap as the inequality decomposition. However, when the subgroups overlap, income means underestimate the gender gap, with the underestimation being proportional to the degree of overlapping.

The analysis of the income distribution of the Italian households shows, on the basis of income means, and without taking into account the overlapping between female and male income distributions, a declining gender gap from 1993 to 2014. Within a framework based on the Gini index decomposition, which considers all distributional characteristics, and which also include overlapping information, we find that gender gap explained 31% of total inequality in 1993, a level which remains stable until 2014.

1 Introduction

Gender inequalities and gender gaps are a worldwide concern and represent the core of uncountable actions and policies developed by either governments and institutions. Gender inequalities are firstly a primary and fundamental issue of justice and represent the current expression of long standing questions about the sources of inequality advanced by philosophers, political scientists and economists during the centuries. Consequences of gender inequalities are frequently overlooked or underestimated, while it exists an interesting literature which analyzes the relation between gender inequality and welfare, pointing out gender gaps as a constraint for economic growth. Furthermore also overall inequality is positively related to gender gap.

A growing literature aims at the evaluation of gender gap and of its effects, with an impressive escalation of contributions and proposals during the last years (see e.g. Atkinson et al. [2], Bonnet et al. [4], Bucher et al. [5], Cupak et al. [11], Goraus et al. [23], Mussida and Picchio [25], Pittau et al. [28], Schneebaum et al. [29]).

We assess the role of gender in income inequality by decomposing the Gini inequality ratio following the approach introduced by Dagum in 1997 ([14]). The first advantage of our approach is to develop a measurement system of income gender gap based on differences not only in average incomes, but in income female and male distributions. As well known, income distributions are characterized by a right fat tail which prevents the use of the mean as a synthetic indicator. Consequently all mean-based methods are not suited for the analysis of income distributions. The same argument holds when some symmetric distribution, as the gaussian, is required. On the contrary, the Gini index, as a function of all pairwise differences between the observed units allows to correctly evaluate all distributional characteristics, skewness in primis, and to take into account the presence of an heavy right tail.

The necessity of evaluating the gender gap as a comparison between female and male income distributions and not only as a difference between average incomes is addressed, among the others, by Goraus et al. ([23]), who emphasize the problem of uninformative means, and by Selezneva and Van Kerm ([30]), who propose an innovative distribution-sensitive analysis of gender wage gap.

The evaluation of gender-related effects on income through inequality decomposition is still a novelty, even if some attempts have been made and the literature offers interesting contributions. Among the others, Checchi and Peragine ([7]) achieve a novel measure of opportunity inequality, Chantreuil and Lebon ([6]) develop an interesting proposal based on the Shapley approach, Larraz ([24]) introduces a refinement of the Dagum's decomposition for small populations and Bonnet et al. ([4]) follow the Gini index decomposition proposed by Lerman and Yitzhaki ([31]).

Besides being simple and intuitive, the Dagum's Gini index decomposition has a further feature which is extremely useful for the gender gap analysis. The income distributions of female and male subpopulations differ in many aspects, but also share a strong overlap. Regarding this aspect, the Dagum's decomposition is particularly suitable, as it is among the few decompositions that explicitly consider

overlap. Our main contribution is to evaluate the overlap between male and female distributions and to include also this element in the gender gap analysis.

We develop an unified framework for the evaluation of gender gap, linking traditional measures, based on subgroups income means, to the approach related to inequality decomposition, and showing how overlapping component represents a key issue in gender gap analysis.

We also contribute to literature on gender gap and on inequality decomposition by extending traditional approaches, thus achieving a more complete and effective assessment of the gender gap. We propose three generalizations of the usual decomposition framework.

First, we focus on the differences at the lower and at the upper end of the male and female distributions, by comparing the related decompositions and by analysing how gender gap affects different parts of the distributions.

Second, by means of inequality decomposition we are able to analyse the inequality structure and take into account other variables (such as area of residence or educational level), able to influence the gender gap. In this way it is possible to develop a broader approach to the analysis of gender gap, where other factors can be added to the income dimension. Since gender gap is a multidimensional issue (see e.g. [18], [1]), which cannot be fully explained by only the income dimension, the inclusion of further inequality factors is of great importance.

Third, we evaluate inequality between female and male distributions within a more realistic framework, which allows to fully understand the relevance of gender gap.

The next Section briefly outlines the main aspects of the Dagum's Gini index decomposition, while Section 3 illustrates the methods proposed for gender gap evaluation. Section 4 presents a case study for Italian households and Section 5 concludes.

2 The Dagum's Gini index decomposition

The Gini index [19] is one of the most important measure of inequality and, during its over 100 years of life, has experienced many different interpretations, expressions and formulas, starting from

$$G = \frac{1}{2n^2\bar{y}} \sum_{i=1}^n \sum_{r=1}^n |y_i - y_r| = \frac{\Delta}{2\bar{y}}$$

where \bar{y} is the arithmetic mean of Y in the overall population, y_i is the value of Y in the i -th unit and Δ is the Gini mean difference. For the case of a population disaggregated into k subgroups of size n_j , with $\sum_{j=1}^k n_j = n$, the Gini index can be expressed as follows

$$G = \frac{1}{2n^2\bar{y}} \sum_{j=1}^k \sum_{h=1}^k \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} |y_{ji} - y_{hr}| \quad (1)$$

where y_{ji} is the value of Y in the i -th unit of the j -th subgroup and, accordingly, y_{hr} is the value of Y in the r -th unit of the h -th subgroup.

A further expression of the Gini index which is extremely useful for the index decomposition is

$$G = \sum_{j=1}^k \sum_{h=1}^k G_{jh} p_j s_h \quad (2)$$

where G_{jj} is the Gini index of the j -th subgroup, G_{jh} is the Gini index between subgroup j and subgroup h , with

$$G_{jh} = \frac{1}{n_j n_h (\bar{y}_j + \bar{y}_h)} \sum_{i=1}^{n_j} \sum_{r=1}^{n_h} |y_{ji} - y_{hr}| = \frac{\Delta_{jh}}{\bar{y}_j + \bar{y}_h} \quad (3)$$

while $p_j = n_j/n$ and $s_j = (n_j \bar{y}_j)/(n\bar{y})$ are the population share and the income share of the j -th subgroup, respectively.

The literature on the Gini index is quite extensive and cannot be easily summarized, however for a detailed discussion of the Gini index see, e.g., [12],[21].

The information provided by the Gini index, related to the overall inequality level, can be successfully exploited by means of the Gini index decomposition. Our interest is specifically dedicated to this point, since it represents the key for our proposal to address gender gap analysis.

Among the many methods which allow to decompose the Gini index (see, e.g., [13],[22],[31]), we use the decomposition proposed by Dagum [14], which is really simple and intuitive.

Besides the two traditional components that characterize all the index decompositions, i.e. the inequality within subgroups G_w and the inequality between subgroups G_b , Dagum explicitly considers a third component, G_t , related to the overlapping, or transvariation (using the Gini's terminology), between subgroups. Since overlapping represents a major feature of many real situations, such as gender gap analysis, we think that it is necessary to include also this element into the decomposition, thus giving to overlapping and its interpretation the same status as inequality within and inequality between, without considering this third component as an unfortunate and undesirable residual term.

The appeal of the Dagum's proposal is its great simplicity: the n^2 differences $|y_{ji} - y_{hr}|$ in (1) are directly assigned to the three components of the decomposition.

First, inequality within G_w is derived from the differences $|y_{ji} - y_{hr}|$ belonging to the same subgroup, that is for $j = h$.

Second, when $j \neq h$, that is when the quantities y_{ji} and y_{hr} involve two different subgroups, the difference $|y_{ji} - y_{hr}|$ is assigned to the inequality between G_b if $\bar{y}_j \geq \bar{y}_h$ and $y_{ji} \geq y_{hr}$, that is when the sign of the difference is the same as the sign of the difference between the subgroups means.

Third, if $j \neq h$, $\bar{y}_j \geq \bar{y}_h$ and $y_{ji} < y_{hr}$ that is in case of an overlapping unit, the difference $|y_{ji} - y_{hr}|$ belongs to the transvariation component G_t .

Even if the focus of the paper is not on the Gini index decomposition, but on the advantages of its use on the gender gap measurement, it is however relevant to note how Dagum's contribution starts an impressive line of researches. Mussard [26], individually or together with co-authors, Ebert [15], Ogwang [27], just to mention some Authors, extensively studied the Dagum's decomposition, proposing generalizations under many aspects and advancing considerably the analysis of inequality.

The analysis of the gender gap implies the presence of $k = 2$ subgroups, as the total population is disaggregated into the female (f) and male (m) subgroups. The existence and the relevance of gender gap originate from \bar{y}_f , the income mean of the female subpopulation, being strongly and systematically lower than \bar{y}_m , the income mean of the male subpopulation. In the following, assuming $\bar{y}_f < \bar{y}_m$ as reference, it is illustrated the Dagum's decomposition for the analysis of gender gap, that is for $k = 2$ subgroups.

3 Gender gap analysis by means of inequality decomposition

When a population is divided into only 2 subgroups, the female (f) and male (m) subpopulations, the Gini index, defined as in (2), can be expressed as

$$G = G_f p_f s_f + G_m p_m s_m + G_{fm} p_f s_m + G_{mf} p_m s_f$$

The case of two subgroups also allows some important simplifications since

$$p_m = 1 - p_f, \quad s_m = 1 - s_f, \quad G_{fm} = G_{mf}.$$

By referring to the female and male subpopulations, we can write the Gini index between f and m as

$$G_{mf} = \frac{\Delta_{mf}}{\bar{y}_m + \bar{y}_f}$$

where

$$\Delta_{mf} = \frac{1}{n_m n_f} \sum_{i=1}^{n_m} \sum_{r=1}^{n_f} |y_{mi} - y_{fr}| = \frac{1}{n_m n_f} \sum_{i=1}^{n_m} \sum_{r=1}^{n_f} (y_{mi} - y_{fr}) + \frac{1}{n_m n_f} \sum_{i=1}^{n_m} \sum_{r=1}^{n_f} (y_{fr} - y_{mi})$$

$y_m > y_f$ $y_f > y_m$

Let be $F(y) = P(Y \leq y)$ the cumulative distribution function, $f(y)$ the probability distribution function and $E(Y)$ the mathematical expectation, then

$$G_{mf} = \frac{1}{E(Y_m) + E(Y_f)} \int_0^\infty \int_0^\infty |Y_m - Y_f| dF(Y_m) dF(Y_f)$$

where

$$\begin{aligned} & \int_0^\infty \int_0^\infty |Y_m - Y_f| dF(Y_m) dF(Y_f) = \\ & = \int_0^\infty dF(Y_m) \int_0^{y_m} (Y_m - Y_f) dF(Y_f) + \int_0^\infty dF(Y_m) \int_{y_m}^\infty (Y_f - Y_m) dF(Y_f) \end{aligned}$$

Given $E(Y_m) > E(Y_f)$, we have that

$$\int_0^\infty dF(Y_m) \int_0^{y_m} (Y_m - Y_f) dF(Y_f) > \int_0^\infty dF(Y_m) \int_{y_m}^\infty (Y_f - Y_m) dF(Y_f),$$

up to the case $\int_0^\infty dF(Y_m) \int_{y_m}^\infty (Y_f - Y_m) dF(Y_f) = 0$ when the two subgroups are not overlapping.

The measurement of gender gap is based on G_{fm} and requires to assess the relevance of the quantity $\int_0^\infty dF(Y_m) \int_0^{y_m} (Y_m - Y_f) dF(Y_f)$ related to the differences between the male and the female subgroups and also the amount of the quantity $\int_0^\infty dF(Y_m) \int_{y_m}^\infty (Y_f - Y_m) dF(Y_f)$ related to the overlapping.

In the following the three components of the Dagum's Gini index decomposition, G_w , G_b and G_t , are derived and interpreted within the framework of gender gap evaluation.

3.1 The inequality within subgroups

The component of inequality within G_w can be obtained quite easily as a weighted sum of the Gini indexes of each subgroup, where the weights are given by the population share and the income share of the two subgroups:

$$G_w = G_f p_f s_f + G_m p_m s_m \quad (4)$$

Even if the measurement of G_w is the subject of some criticism (see, e.g. Frosini [17]), the Dagum's proposal is in line with the existing literature, where the majority of the decompositions obtains G_w as a function of G_j .

G_w allows to evaluate the contribution to total inequality related to the variability within the subgroups. Low values of G_w indicate homogenous subgroups and a reduced impact of variability within female and male subgroups on total inequality, while an high G_w provides the opposite indications.

3.2 The inequality between subgroups and the overlapping component

Given G and G_w , we can obtain the inequality between subgroups G_b and the overlapping component G_t as

$$G - G_w = G_b + G_t = G_{fm}p_f s_m + G_{mf}p_m s_f$$

For G_b and G_t , which in the original version of the Dagum's decomposition require some substantial computational effort, are also available [9] simplified expressions.

In order to derive the inequality between subgroups G_b and the overlapping component G_t , it is useful to start from the simplest case, that is only two non overlapping subgroups, and then to introduce the presence of overlapping.

3.2.1 The case of no overlapping

A population divided into two non overlapping subgroups, despite being a relevant departure from many real situations, offers a simple and straightforward solution for the measurement of the inequality between.

Since $G_t = 0$ we have

$$G_b = G_{fm}p_f s_m + G_{mf}p_m s_f = G_{fm}(p_f s_m + p_m s_f) = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} \sum_{i=1}^{n_f} \sum_{r=1}^{n_m} |y_{fi} - y_{mr}|$$

Given $\bar{y}_f < \bar{y}_m$, i.e. assuming the presence of gender gap, the absence of overlapping implies $|y_{fi} - y_{mr}| = (y_{mr} - y_{fi})$, that is

$$G_b = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} \sum_{i=1}^{n_f} \sum_{r=1}^{n_m} (y_{mr} - y_{fi}) = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} n_f n_m (\bar{y}_m - \bar{y}_f).$$

Analogously, since $G_t = 0$ implies $\int_0^\infty dF(Y_m) \int_{y_m}^\infty (Y_f - Y_m) dF(Y_f) = 0$, we have

$$\begin{aligned} G_b &= \frac{p_f s_m + p_m s_f}{E(Y_f) + E(Y_m)} \int_0^\infty \int_0^\infty (Y_m - Y_f) dF(Y_m) dF(Y_f) = \\ &= \frac{p_f s_m + p_m s_f}{E(Y_f) + E(Y_m)} \left(\int_0^\infty Y_m dF(Y_m) - \int_0^\infty Y_f dF(Y_f) \right) = \frac{p_f s_m + p_m s_f}{E(Y_f) + E(Y_m)} (E(Y_m) - E(Y_f)). \end{aligned}$$

Starting from this structure, based on the subgroups means, it is possible to further simplify the expression of G_b :

$$\begin{aligned} G_b &= \frac{p_f s_m + p_m s_f}{(\bar{y}_f + \bar{y}_m)} (\bar{y}_m - \bar{y}_f) = p_f (s_m + p_m \bar{y}_f / \bar{y}) \frac{(\bar{y}_m - \bar{y}_f)}{(\bar{y}_f + \bar{y}_m)} = p_f p_m \frac{(\bar{y}_m + \bar{y}_f)}{\bar{y}} \frac{(\bar{y}_m - \bar{y}_f)}{(\bar{y}_f + \bar{y}_m)} = \\ &= p_f p_m \bar{y}_m / \bar{y} - p_f p_m \bar{y}_f / \bar{y} = p_f (1 - s_f) - (1 - p_f) s_f = p_f - s_f. \end{aligned}$$

Therefore our final result for G_b allow to evaluate the inequality between as the difference between the population share and the income share of the female subgroup.

The Dagum's Gini index decomposition therefore results

$$G = G_w + G_b = (G_f p_f s_f + G_m p_m s_m) + (p_f - s_f)$$

where the gender gap can be evaluated by the difference $(p_f - s_f)$.

3.2.2 The case of overlapping

In the case of overlapping subgroups, some differences $(y_{mr} - y_{fi})$ result negative and therefore the quantity $(p_f - s_f)$, which is the sum of all differences $(y_{mr} - y_{fi})$, both positive and negative, will underestimate the inequality between G_b for an amount equal to the sum of the negative differences.

Since the negative differences correspond to the overlapping component, we obtain

$$G_b = \frac{p_f s_m + p_m s_f}{n_f n_m (\bar{y}_f + \bar{y}_m)} \left(\sum_{i=1}^{n_f} \sum_{r=1}^{n_m} (y_{mr} - y_{fi}) + \sum_{i=1}^{n_m} \sum_{r=1}^{n_f} (y_{fr} - y_{mi}) \right)_{y_f > y_m}$$

and

$$G_b = \frac{p_f s_m + p_m s_f}{E(Y_f) + E(Y_m)} \int_0^\infty \int_0^\infty (Y_m - Y_f) dF(Y_m) dF(Y_f) + \frac{p_f s_m + p_m s_f}{E(Y_f) + E(Y_m)} \int_0^\infty dF(Y_m) \int_{y_m}^\infty (Y_f - Y_m) dF(Y_f)$$

In a more compact form we can write

$$G_b = p_f - s_f + G_t$$

and the Dagum's Gini index decomposition results

$$G = G_w + G_b + G_t = (G_f p_f s_f + G_m p_m s_m) + (p_f - s_f + G_t) + G_t$$

from which it is possible to derive G_t as

$$G_t = (G - G_w - p_f - s_f) / 2$$

Inequality between G_b and overlapping component G_t allow to evaluate the contribution to total inequality attributable to the differences between the subgroups, that is, in our analysis, the gender gap.

The role of the two components is quite different. From one side, an high (low) G_b indicates a relevant (slight) gender gap, as total inequality is (is not) strongly influenced by inequality between. From the other side, an high (low) G_t points out to a slight (relevant) gender gap, since complete overlapping corresponds to the absence of gender gap, while $G_t = 0$ (female and male subgroups are perfectly separated) indicates a total stratification.

The gender gap can still be evaluated by means of G_b , but in the general case of possible overlapping, it is given by $(p_f - s_f + G_t)$.

It is worth to note how the use of G_b to evaluate the gender gap is direct and straightforward, and how it does not imply the inclusion of any additional assumption, or constraint, as frequently happens in the case of parametric methods, even simple ones, such as the linear ones.

3.2.3 The measurement of overlapping

Since overlapping represents a key issue in gender gap evaluation, its measurement requires a specific attention. At this regard we refer to two indicators, introduced by Gini (1959) [20], namely the probability and the intensity of transvariation.

Probability of transvariation refers to overlapping which occurs when $y_{fi} > y_{mr}$ while $\bar{y}_{mef} < \bar{y}_{mem}$, that is the gender gap is defined by the relation between the medians. Null differences ($y_{mr} - y_{fi}$) are equally divided between overlapping and non-overlapping differences. Probability of transvariation pt is then computed as the number of overlapping differences to its maximum, that is as the ratio

$$pt = (2 \cdot {}_1nt + {}_2nt) / (n_f n_m)$$

where ${}_1nt$ and ${}_2nt$ are, respectively, the number of negative and the number of null differences ($y_{mr} - y_{fi}$). The maximum of pt is reached when the two subgroups completely overlap and the median of the female subgroup equals the median of the male subgroup. Probability of transvariation ranges between 0, when there is no transvariation and no difference is overlapping, and 1, when overlapping component reaches its maximum.

The second indicator introduced by Gini, the intensity of transvariation, refers to overlapping which occurs when $y_{fi} > y_{mr}$ while $\bar{y}_f < \bar{y}_m$, that is the gender gap is defined by the relation between the means. With respect to the probability of transvariation, which takes into account the number of overlapping differences, the intensity of transvariation is based on their size, that is on the quantities $|y_{fi} - y_{mr}|$ for $y_{fi} < y_{mr}$, $i = 1, \dots, n_f$ and $r = 1, \dots, n_m$.

Let be T the sum of overlapping differences:

$$T = \sum_{i=1}^{n_f} \sum_{r=1}^{n_m} |y_{fi} - y_{mr}|$$

for $y_{fi} < y_{mr}$ and $\bar{y}_f < \bar{y}_m$.

It is straightforward to observe how T is increasing as the difference $(\bar{y}_m - \bar{y}_f)$ is decreasing; furthermore, for $\bar{y}_m = \bar{y}_f$, T reaches its maximum, that is

$$\max T = \frac{1}{2} \sum_{i=1}^{n_f} \sum_{r=1}^{n_m} |y_{fi} - y_{mr}|.$$

Intensity of transvariation is obtained as the ratio of the sum of overlapping differences $|y_{fi} - y_{mr}|$ to its maximum, that is as the ratio

$$it = \frac{2T}{\sum_{i=1}^{n_f} \sum_{r=1}^{n_m} |y_{fi} - y_{mr}|}$$

By jointly using intensity and probability of transvariation it is possible to obtain a wide information set about the relevance and the extent of the overlapping component.

3.3 The informative content of subgroups means

The use of the Dagum's Gini index decomposition for the evaluation of the gender gap also allows to shed some light on the debate about the informative content of the subgroups means \bar{y}_m and \bar{y}_f .

The first gross evaluation of the gender gap based on \bar{y}_m and \bar{y}_f is given by the absolute difference

$$gap_1 = (\bar{y}_m - \bar{y}_f)$$

which is usually complemented by some relative difference such as

$$gap_2 = (\bar{y}_m - \bar{y}_f) / \bar{y}_f$$

or

$$gap_3 = (\bar{y}_m - \bar{y}_f) / \bar{y}$$

The extreme case $\bar{y}_m = \bar{y}_f = \bar{y}$ of null gender gap implies $p_f = s_f$ and $p_m = s_m$, that is the equidistribution. It follows that, when $p_i = s_i$, the gender gap is equal to 0, while increasing differences $(p_i - s_i)$ report an increasing inequality. The case $p_f > s_f$, that is when the population share of the female subpopulation is greater than the income share of the female subpopulation, implies $\bar{y}_m > \bar{y}_f$ and indicates the existence of a gender gap. In the following we include also this indicator in our analysis:

$$gap_4 = (p_f - s_f).$$

It is crucial to assess to what extent the subgroups means \bar{y}_m and \bar{y}_f are informative and if they allow to correctly evaluate the gender gap.

We are able to propose a solution by resorting to the pioneering measure of the inequality between proposed by Bhattacharia and Mahalanobis [3]

$$G_{bBM} = \sum_{j=1}^k \sum_{h=1}^k p_j p_h |\bar{y}_j - \bar{y}_h| / 2\bar{y} \quad j \neq h$$

which evaluates the differences between the subgroups on the basis of the subgroups means only. For the case of two subgroups, m and f , we have

$$G_{bBM} = p_f p_m (\bar{y}_m - \bar{y}_f) / \bar{y}$$

and it is possible to show that

$$\begin{aligned} p_f p_m (\bar{y}_m - \bar{y}_f) / \bar{y} &= p_f p_m \bar{y}_m / \bar{y} - p_f p_m \bar{y}_f / \bar{y} = p_f s_m - p_m s_f = \\ &= p_f (1 - s_f) - (1 - p_f) s_f = p_f - p_f s_f - s_f + p_f s_f = p_f - s_f \end{aligned}$$

where the difference $(\bar{y}_m - \bar{y}_f)$ is linked to the difference $(p_f - s_f)$.

Since the quantity $(p_f - s_f)$ corresponds (see Section 3.2.1) to the inequality between in the Dagum's Gini index decomposition for the case of two non overlapping subgroups, we are able to affirm that the subgroups means are fully informative in absence of overlapping. However, their informative content declines for increasing levels of overlapping.

From Section 3.2.2 we also derive that, when the female and male income distributions overlap, the use of the means leads to underestimate the gender gap, with the underestimation being proportional to the degree of overlapping.

3.4 The inequality structure

Inequality decomposition is a powerful tool to get a deep understanding of the inequality structure. We exploit this property to provide insight on gender gap.

Usually the Gini index decomposition is applied to all n available observations on a variable y (such as income, expenditure, etc.) disaggregated on the basis of some inequality factor of interest, in our case the gender of the head of the household.

From the decomposed Gini index

$$G_y = G_{wy} + G_{by} + G_{Ty}$$

we analyze the effect of the inequality factor by means of the three ratios

$$\frac{G_{wy}}{G_y}, \quad \frac{G_{by}}{G_y}, \quad \frac{G_{Ty}}{G_y}.$$

In particular, by referring to the measurement of the gender gap, we add to our set of measures the second ratio,

$$gap_5 = G_b / G$$

which is based on the inequality between, that is the component of the inequality able to evaluate the relevance of the underlying inequality factor.

3.4.1 The effect of low and high income values

A first possible extension with respect to the traditional framework is to compare the decomposition obtained by using all n observations to the decompositions obtained by referring only to subsamples of observations.

In particular, it is useful to analyze the decompositions for the lower values of y ,

$$G_{y|ymin} = G_{wy|ymin} + G_{by|ymin} + G_{ty|ymin}$$

as well as for the higher values of y ,

$$G_{y|ymax} = G_{wy|ymax} + G_{by|ymax} + G_{ty|ymax}.$$

Then we evaluate the inequality structure by analysing the relations

$$\begin{aligned} \frac{G_{wy}}{G_y} &= \frac{G_{wy|ymin}}{G_{y|ymin}} = \frac{G_{wy|ymax}}{G_{y|ymax}}, \\ \frac{G_{by}}{G_y} &= \frac{G_{by|ymin}}{G_{y|ymin}} = \frac{G_{by|ymax}}{G_{y|ymax}}, \\ \frac{G_{ty}}{G_y} &= \frac{G_{ty|ymin}}{G_{y|ymin}} = \frac{G_{ty|ymax}}{G_{y|ymax}}. \end{aligned}$$

When the structure of the decomposed indices $G_{y|ymin}$ and $G_{y|ymax}$ is similar, that is, when the equivalences hold, we get that the underlying inequality factor operates uniformly on y . On the contrary, different structures indicate that particular regions of y are more affected by the inequality factor.

For the analysis of the gender gap we are specifically interested to the equivalences related to G_b , which allow to evaluate more thoroughly the relation between income and gender gap, understanding in particular if gender gap is more related to low or high incomes or if it maintains the same size for any value of y .

3.4.2 The effect of further inequality factors

The previous approach, where the decomposition obtained by using all n observations is compared to decompositions related to subsamples of observations, can be generalized to include additional inequality factors in the gender gap analysis.

In order to evaluate the influence of an inequality factor x , we can rank y on the values of x , select two subsamples of observations related to particular values of x (usually the lowest and highest) and analyze the decomposition for female and male subgroups.

In this case we analyse the inequality structure by means of the relations

$$\frac{G_{wy}}{G_y} = \frac{G_{wy|xmin}}{G_{y|xmin}} = \frac{G_{wy|xmax}}{G_{y|xmax}},$$

$$\frac{G_{by}}{G_y} = \frac{G_{by|xmin}}{G_{y|xmin}} = \frac{G_{by|xmax}}{G_{y|xmax}},$$

$$\frac{G_{ty}}{G_y} = \frac{G_{ty|xmin}}{G_{y|xmin}} = \frac{G_{ty|xmax}}{G_{y|xmax}}$$

Similar decompositions suggest that the inequality factor x does not provide any additional information on gender gap, while different decompositions indicate a relation between x and the gender gap. In the latter case we are able to exploit the additional information on gender gap provided by x , thus moving gender gap evaluation to a multidimensional framework.

3.4.3 The evaluation of the inequality between

A further extension with respect to the traditional inequality decomposition which is of interest for the gender gap analysis refers to the evaluation of the inequality between. Usually the importance of G_b is measured on the basis of the ratio G_w/G , where G acts as the maximum of G_b .

The scenario $G_b = G$ implies $G_w = G_t = 0$. The hypothesis $G_t = 0$, that is the absence of overlapping, even if representing a limit case, is not inconsistent with gender gap analysis and therefore doesn't cause particular difficulties. On the contrary, assuming $G_w = 0$, that is the equidistribution of y within each subgroup, implies a clearly unrealistic situation, where y presents only two values, one for the female units and the other for the male ones.

In order to achieve an evaluation of G_b more coherent with the observed data ([16], [10]), it is possible to keep $G_t = 0$ but to replace $G_w = 0$ with $G_w = G_{wmin}$, which is the minimum inequality within compatible with the observed data. In this case G_b is evaluated as

$$gap_6 = G_b / (G - G_{wmin})$$

and we interpret this ratio as a further measure for the evaluation of the gender gap.

We can obtain G_{wmin} in many ways. In order to preserve the structure of the original data we can keep the p_i or s_i initially observed. First, we sort y in ascending order, second we obtain all the possible subgroups combinations by permutating the initial p_i or s_i . Given $k = 2$, we have as possible cases $\{p_f, p_m\}$ and $\{p_m, p_f\}$ by using the population shares, and $\{s_f, s_m\}$ and $\{s_m, s_f\}$ by using the income shares.

For example, let be $\{p_f, p_m\} = \{0.3, 0.7\}$ and $\{s_f, s_m\} = \{0.2, 0.8\}$ the observed data: the female units represent the 30% of the total population and possess the 20% of the income. By using the observed p_i and s_i , we can achieve 4 new cases. In the first case we set, $\{p_1, p_2\} = \{0.3, 0.7\}$, where the first of the two new subgroups collects the poorest 30% of the observations. In the second case we have $\{p_1, p_2\} = \{0.7, 0.3\}$, where the first of the two new subgroups collects the poorest 70% of the observations. In the third case we refer to $\{s_1, s_2\} = \{0.2, 0.8\}$, where the first of the two new subgroups collects the poorest units possessing up to the 20% of the total

income. In the fourth case we refer to $\{s_1, s_2\} = \{0.8, 0.2\}$, with the units of the first subgroup being the poorest and possessing up to the 80% of the total income.

In the first 2 cases way we keep the original p_i structure, while in the latter 2 cases we keep the original s_i structure. In the last step to obtain G_w , we calculate for the 4 cases the decomposed Gini index and adopt as G_{wmin} the minimum of the 4 G_w calculated.

Since gender gap is evaluated by means of the importance of G_b on overall inequality, the extension illustrated on this Section is of particular interest in order to assess the true effect of gender on inequality.

4 Gender income inequality among Italian households

The Dagum's decomposition of the Gini index presented in Section 2 is extremely useful to analyze the relevance of gender in income inequality. The component G_w allows to evaluate how the income variability existing within the female and male subpopulations influence total inequality, while the contribution attributable to the differences between the female and male subpopulations is given by G_b and G_t . The meaning of G_b is straightforward, but as far as G_t it is useful to point out that high levels of overlapping indicate a small contribution of gender to income inequality, while low levels of overlapping suggest a stronger contribution.

4.1 The data

The data used in this study are from the Survey on Households Income and Wealth, a multidimensional survey on Italian households performed every two years by the Bank of Italy. The analyses and the results illustrated in this section refer to three waves of the survey: 1993, 2004 and 2014.

Table 1 reports the basic information set for the analysis of income gender gap: the population share, the income share, the income mean and the Gini index for the Italian households by gender of the head of the household. We observe how, during 20 years, a female head of the household, initially only slightly more than a quarter of the population, becomes a consolidated reality, actually representing more than 35% of the Italian households.

The household income distribution by gender is illustrated in Figures 1-3 for 1993, 2004 and 2014, respectively: we clearly observe the well known positive skewness which usually strongly characterizes income distributions and also the strong overlapping between the distribution of men and women.

Table 1 also includes further information about the income distribution by means of the median \bar{y}_{me} and the coefficient of variation cv , given by the ratio of the standard deviation by the mean. The distance between \bar{y} and \bar{y}_{me} confirms the relevance of the skewness; we can also note how the women income distribution is greater

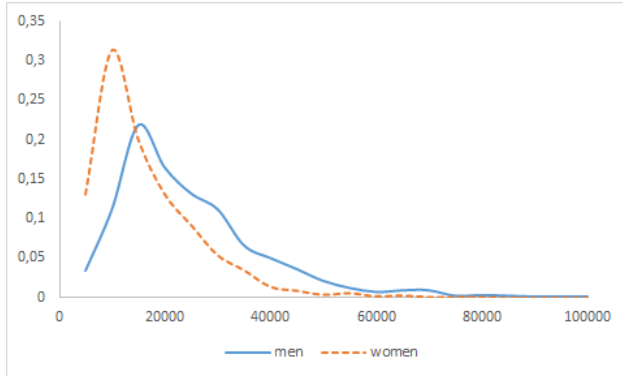


Fig. 1 Income distribution, men and women, Italy 1993, euros

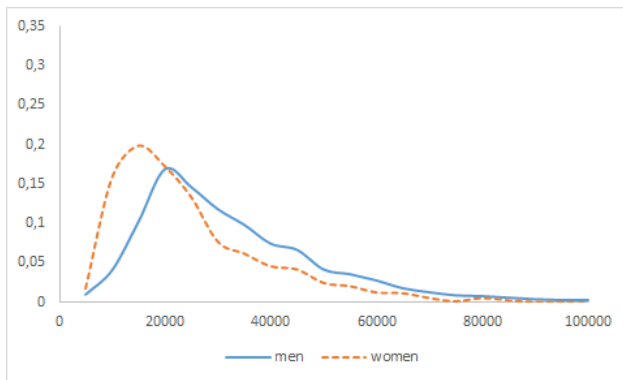


Fig. 2 Income distribution, men and women, Italy 2004, euros

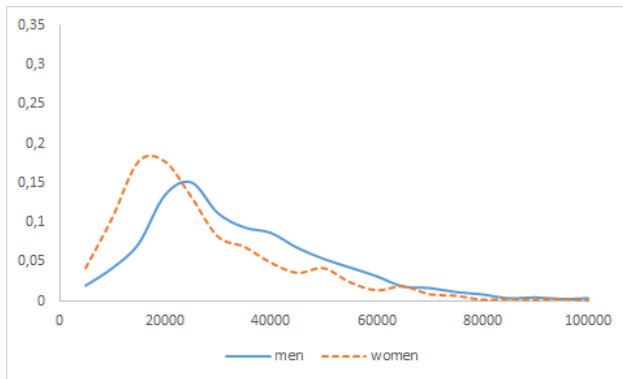


Fig. 3 Income distribution, men and women, Italy 2014, euros

affected by this aspect. Overall the difference $\bar{y} - \bar{y}_{me}$ and the cv point out to relevant differences between men and women distributions, thus suggesting that the use of an indicator based on pairwise differences between of the observed units, such as the Gini index, should be a preferred choice for income distribution analysis.

Table 1 Population share, income share, mean, median, coefficient of variation and Gini index for the Italian households by gender of the head of the household

	1993			2004			2014		
	female	male	tot	female	male	tot	female	male	tot
p	0.275	0.725	1.000	0.302	0.698	1.000	0.352	0.648	1.00
s	0.193	0.807	1.000	0.238	0.762	1.000	0.293	0.707	1.00
\bar{y}	14372	22730	20432	23204	32200	29483	25393	33319	30525
\bar{y}_{me}	11155	18901	16578	18450	26200	23833	19958	28251	25104
cv	0.794	0.700	0.748	0.734	0.873	0.869	0.871	0.661	0.733
G	0.382	0.342	0.366	0.362	0.339	0.353	0.375	0.328	0.350

Given the importance of the overlapping between female and male income distributions, in Table 2 are reported the probability and the intensity of transvariation for the Italian female and male income distributions. From 1993 to 2014 it is possible to observe a steadily increase of both indicators, thus indicating a stronger role of the overlapping between the two income distributions.

Table 2 Probability and intensity of transvariation for the Italian households by gender of the head of the household

	1993	2004	2014
pt	0.631	0.692	0.695
it	0.538	0.596	0.628

4.2 The gender gap by subpopulations means

The first, basic, traditional evaluation of gender gap is based on the subgroups means \bar{y}_m and \bar{y}_f . In Table 3 and Figure 4 it is reported the absolute difference ($\bar{y}_m - \bar{y}_f$), as well as some relative differences between \bar{y}_m and \bar{y}_f , showing a declining income gender gap for the Italian households from 1993 to 2014. We can observe a decrease both in absolute and in relative values, with gap_2 and gap_3 which fall by 46% and 36% respectively.

The same indication is provided by $gap_4 = (p_f - s_f)$, which is null when the gender gap is equal to 0, while it increases for increasing level of gender gap: from 1993 to 2014 this indicator goes from 0.082 to 0.059, with a decrease of 28%.

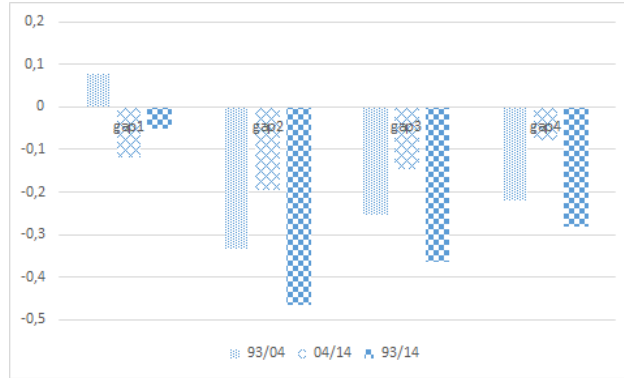


Fig. 4 Gender gap by subpopulations means: variations 93/04, 04/14, 93/14

Overall the aggregate data of Table 3 suggest the presence of a gender gap, but also its reduction over time.

Table 3 Gender income gap for the Italian households by income means

	gap_1 $(\bar{y}_m - \bar{y}_f)$	gap_2 $(\bar{y}_m - \bar{y}_f)/\bar{y}_f$	gap_3 $(\bar{y}_m - \bar{y}_f)/\bar{y}$	gap_4 $(p_f - s_f)$
1993	8358	0.582	0.409	0.082
2004	8996	0.388	0.305	0.064
2014	7927	0.312	0.260	0.059

4.3 The gender gap by inequality decomposition

Moving from the aggregate and means-based evaluation provided in the previous Section to the more detailed and accurate information contained on the decomposed Gini index (Table 4 and Figures 5-6), we obtain a different picture on gender income inequality.

First, the importance of inequality within on total inequality strongly decreases (from 60% in 1993, to 58% in 2004 and to 54% in 2014), thus indicating a weaker variability within the female and male subpopulations and a stronger effect of the underlying inequality factor, in our case the gender.

Second, the overlapping between the female and male subpopulations increases: the importance of the G_t component rises from 8.7% in 1993, to 11.7% in 2004 and to 14.6% in 2014. A greater overlapping represents a positive signal for the reduction of the gender gap, since it suggests that the distributions of the subpopulations share a larger area.

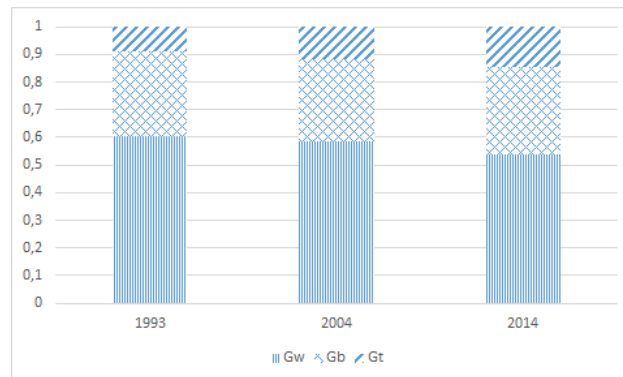


Fig. 5 Decomposed Gini index (male and female subgroups)

Third, the inequality between is stable, as the importance of the G_b component goes from 31% in 1993, to 29.9% in 2004 and to 31.5% in 2014. The relevance of the inequality between is fully understandable by relaxing the hypothesis $G_w = 0$ and by comparing G_b to its maximum compatible with the observed data (last column of Table 4): in this case G_b represents 48.2% of total inequality in 1993, moving to 42.6% in 2004 and to 44.7% in 2014.

Table 4 Income inequality decomposition by gender of the head of the household

	Gw	Gb	Gt	Gw/G	$\frac{gap_5}{Gb/G}$	Gt/G	$\frac{gap_6}{Gb/(G-Gwmin)}$
1993	0.221	0.113	0.032	0.603	0.310	0.087	0.482
2004	0.206	0.105	0.041	0.585	0.299	0.117	0.426
2014	0.189	0.110	0.051	0.539	0.315	0.146	0.447

Overall, the decrease of inequality within is balanced by the increase of the overlapping component. While a greater G_t alleviates the role of gender as inequality factor, this increase is not sufficient to reduce G_b , which shows a stable gender income inequality during the period 1993-2014. Our empirical findings are in line with existing literature on gender gap: also Larraz (2015) ([24]), working on Spanish data, points out that inequality within accounts for 50% of total inequality and how the overlapping component plays a relevant role. Figure 6 illustrates the dynamic of gender gap by taking into account also the overlapping between female and male distributions: its reduction is much lower than indicated by using the subgroup means.

In order to better understand the results of Table 4, we focus on the left and the right tail of the income distribution, taking into account the bottom and the top 20% of the income. Table 5 reports the population share, the income share and the Gini

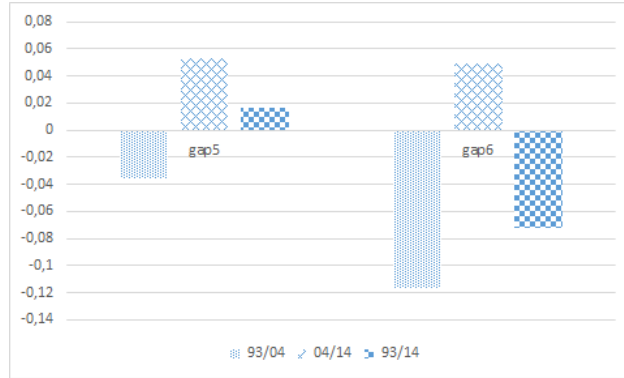


Fig. 6 Gender gap by inequality decomposition: variations 93/04, 04/14, 93/14

index for the female and male subpopulations for the two cases and it is possible to observe some relevant differences.

Overall, the female subpopulation share p_f increases steadily from 1993 to 2014, but the changes are not homogenous in the two subgroups: p_f decreases in the bottom 20%, while it assumes a far greater relevance for the top 20% incomes. In 1993 the two subgroups are quite different, with the bottom 20% characterized by an unicum $p_f > p_m$ and the top 20% where the female subpopulation represents only a small fraction of the total. 20 years later, in 2014, the two subgroups are more similar, for the bottom 20% is as usually $p_f < p_m$, while for the top 20% the female subpopulation share is doubled since 1993.

Table 5 Population share, income share and Gini index for the Italian households by gender of the head of the household

	1993			2004			2014		
	female	male	tot	female	male	tot	female	male	tot
20% bottom income									
p	0.558	0.442	1.000	0.530	0.470	1.000	0.431	0.569	1.000
s	0.548	0.452	1.000	0.510	0.490	1.000	0.426	0.574	1.000
G	0.186	0.211	0.199	0.171	0.157	0.167	0.244	0.214	0.227
20% top income									
p	0.118	0.882	1.000	0.195	0.805	1.000	0.246	0.754	1.000
s	0.112	0.888	1.000	0.179	0.821	1.000	0.244	0.756	1.000
G	0.171	0.180	0.179	0.155	0.218	0.207	0.177	0.168	0.170

We analyze the bottom and the top of the income distribution, with the aim to compare the structure of the decomposed Gini indices.

By comparing the decomposed Gini indexes for the bottom and the top incomes (Table 6), we note that the two decompositions, initially quite different, are more or less similar in 2014. We can attribute the reduction of the importance of G_w , already

observed on Table 2, to the 20% top income. The overlapping component, initially stronger for the bottom incomes, represents around the 20% of total inequality for both subgroups, and the increase of its weight is quite evenly distributed between the bottom and top incomes. On the contrary, the stability of the importance of the inequality between observed on Table 4 is the results of an increase related to the top incomes and a slight decrease observed for the 20% bottom income.

Table 6 Income inequality decomposition by gender of the head of the household

	Gw/G	Gb/G	Gt/G	Gw/G	Gb/G	Gt/G
	20% bottom income			20% top income		
1993	0.497	0.275	0.227	0.790	0.118	0.083
2004	0.494	0.313	0.193	0.722	0.177	0.101
2014	0.505	0.258	0.237	0.623	0.195	0.182

A further analysis of the gender income inequality refers to the study of particular population characteristics, such as educational level and geographical area of residence, chosen among the main inequality factors acknowledged by the literature. The Gini index decomposition is applied not to all n observations of the overall population, but only to the subsample of households with the particular characteristic which we are analyzing. More specifically, we compare the female/male decompositions obtained on two subgroups related to two different values of the character under examination. When the two decompositions are substantially similar, the underlying factor is not relevant for the interpretation of the gender inequality, while, on the contrary, different decompositions indicate an influence on gender inequality.

Table 7 illustrates the population share, the income share and the Gini index for two subgroups: for the educational level we compare the up-to-elementary school group to the group with a university degree, for the geographical area the group living in the north to the group living in the south or islands.

The related decompositions of the Gini index for the analysis of the gender gap are shown in Table 8. The comparison between the decompositions suggests that the educational level influences the gender income inequality more than the geographical area. We also confirm the decrease of the importance of the inequality within, together with an increase of the relevance of the overlapping component and of the inequality between, especially for the more affluent subgroups.

5 Conclusions

The decomposition of an inequality index allows powerful insights on the inequality structure and can be extremely useful into the study of the gender income inequality, where the decomposition refers to the female and male subpopulations.

Table 7 Population share, income share and Gini index for the Italian households by gender of the head of the household

	1993			2004			2014		
	female	male	tot	female	male	tot	female	male	tot
up to elementary school									
p	0.388	0.612	1.000	0.435	0.565	1.000	0.488	0.512	1.000
s	0.289	0.711	1.000	0.348	0.652	1.000	0.402	0.598	1.000
G	0.355	0.323	0.353	0.296	0.279	0.303	0.279	0.276	0.294
with university degree									
p	0.172	0.828	1.000	0.304	0.696	1.000	0.400	0.600	1.000
s	0.143	0.857	1.000	0.236	0.764	1.000	0.331	0.669	1.000
G	0.281	0.301	0.302	0.306	0.338	0.339	0.332	0.322	0.335
south islands									
p	0.247	0.753	1.000	0.302	0.698	1.000	0.357	0.643	1.000
s	0.172	0.828	1.000	0.242	0.758	1.000	0.308	0.692	1.000
G	0.352	0.360	0.371	0.385	0.321	0.348	0.411	0.326	0.361
north									
p	0.298	0.702	1.000	0.294	0.706	1.000	0.343	0.657	1.000
s	0.205	0.795	1.000	0.233	0.767	1.000	0.277	0.723	1.000
G	0.383	0.318	0.353	0.346	0.324	0.337	0.340	0.309	0.328

Table 8 Income inequality decomposition by gender of the head of the household, Italy 1993-2014

	Gw/G	Gb/G	Gt/G	Gw/G	Gb/G	Gt/G
up to elementary school						
1993	0.512	0.384	0.104	0.730	0.182	0.088
2004	0.488	0.400	0.113	0.594	0.305	0.101
2014	0.473	0.410	0.117	0.517	0.345	0.138
south islands						
north						
1993	0.645	0.278	0.077	0.569	0.349	0.083
2004	0.569	0.301	0.130	0.589	0.296	0.115
2014	0.526	0.305	0.169	0.546	0.327	0.127

We develop an unified framework for the gender gap analysis, where the approach based on the subgroups means is linked to the methods related to the inequality decomposition. We show how the two approaches are equivalent when female and male income distributions are not overlapping, while in the case of overlapping income means lead to underestimate the gender gap.

We also extend inequality decompositions methods, thus allowing a more complete and effective assessment of gender gap. First we evaluate the size of the gender gap for different values of the income, second we assess the relevance on gender gap of further inequality factors, third we measure the inequality between by weakening the unrealistic hypothesis of null inequality within.

The analysis of the income distribution of the Italian households shows how gender gap explained 31% of total inequality in 1993, a level which remains stable until 2014. The scenario is even worse when evaluating inequality attributable to

the differences between female and male subpopulations without the traditional assumption of null inequality within: in this case gender gaps are accountable for up 40% of total inequality from 1993 to 2014. Inequality decomposition also allows to evaluate the relation between gender and other inequality factors: educational level of the head of the household and geographical area of residence are taken into account, with the former showing a greater influence on gender income inequality.

On the basis of the information provided by the average incomes, and without taking into account the overlapping between female and male income distributions, we obtain for the Italian case a declining gender gap from 1993 to 2014. Within a framework based on the Gini index decomposition, which considers all distributional characteristics, we reach a different conclusion, which can be extremely helpful for gender gap evaluation.

References

1. Amici, M., Stefani, M.L.: A gender equality index for the Italian regions. *Questioni di Economia e Finanza*, **190**, Banca d'Italia (2013)
2. Atkinson, A.B., Casarico, A., Voitchovsky, S.: Top incomes and the gender divide. *Journal of Economic Inequality*, **16**, 225–256 (2018)
3. Bhattacharya, B., Mahalanobis, B.: Regional disparities in household consumption in India. *Journal of the American Statistical Association* **62**, 143–161 (1967)
4. Bonnet, C., Meurs, D., Rapport, B.: Gender inequalities in pensions: different components similar levels of dispersion. *Journal of Economic Inequality*, **16**, (2018)
5. Bucher Koenen, T., Lusardi, A., Alessie, R., Van Rooij, M.: How financially literate are women? An overview and new insights. *Journal of Consumer Affairs*, **51**, 255–283 (2017)
6. Chantreuil, F., Lebon, I.: Gender contribution to income inequality. *Economic Letters*, **133**, 27–30 (2015)
7. Checchi, D., Peragine, V.: Inequality of opportunity in Italy. *Journal of Economic Inequality*, **8**, 429–450 (2010)
8. Costa, M.: Gini index decomposition for the case of two subgroups. *Communications in Statistics*, **37**, 631–644 (2008)
9. Costa, M.: Transvariation and inequality between subpopulations in the Dagum's Gini index decomposition. *Metron*, **67**, 134–120 (2009)
10. Costa, M.: The evaluation of the inequality between population subgroups. In Verde, R., Racioppi, F., Petrucci, A.: *Statistics and Data Science: new challenges, new generations*. Firenze University Press, Florence, 313–318 (2017)
11. Cupak, A., Fessler, P., Schneebaum, A., Silgoner, M.: Decomposing gender gaps in financial literacy: new international evidence. *Economics Letters*, **168** (2018)
12. Dagum, C.: Gini ratio. In: *The New Palgrave Dictionary of Economics*. Mac Millian Press, London (1987)
13. Dagum, C., Zenga M.: *Income and Wealth Distribution, Inequality and Poverty*. Springer, Berlin (1990)
14. Dagum, C.: A new decomposition of the Gini income inequality ratio. *Empirical Economics*, **22**, 515–531 (1997)
15. Ebert, U.: The decomposition of inequality reconsidered: Weakly decomposable measures. *Mathematical Social Sciences*, **60**, 94–103 (2010)
16. Elbers, C., Lanjouw, P., Mistiaen J.A., Ozler, B.: Reinterpreting Between-Group Inequality. *Journal of Economic Inequality*, **6**, 231–245 (2008)
17. Frosini, B.V.: Approximation and decomposition of Gini, Pietra-Ricci and Theil inequality measures. *Empirical Economics*, **43**, 175–197 (2012)

18. Gender Equality Index Report, European Institute for Gender Equality (2013)
19. Gini, C.: Sulla misura della concentrazione e della variabilit dei caratteri. *Atti R Ist Veneto Sci Lett Arti* **73**, 1203–1248 (1914); English translation in *Metron* **63**, 3–38 (2005)
20. Gini, C.: *Transvariazione*. Libreria goliardica, Roma (1959)
21. Giorgi, G.M.: Gini’s scientific work: an evergreen. *Metron*, **63**, 299–315 (2005)
22. Giorgi, G.M.: The Gini inequality index decomposition. An evolutionary study. In Deutsch, J., Silber, J.: *The measurement of individual well-being and group inequalities*. Routledge, London (2011)
23. Goraus, K., Tyrowicz, J., van der Velde, L.: Which gender gap estimate to trust? A comparative analysis. *The Review of Income and Wealth*, **63**, 118–146 (2017)
24. Larraz, S.: Decomposing the Gini inequality index: an expanded solution with survey data applied to analyze income inequality. *Sociological Methods and Research*, **44**, 508–531 (2015)
25. Mussida, C., Picchio, M.: The gender wage gap by education in Italy. *Journal of Economic Inequality*, **12**, 117–147 (2014)
26. Mornet, P., Zoli, C., Mussard, S., Sadefo-Kamdem, J., Seyte, F., Terraza, M.: The α, β - multi-level α -Gini decomposition with an illustration to income inequality in France 2005. *Economic Modelling*, **35**, 944–963 (2013)
27. Ogowang, T.: A convenient method of decomposing the Gini index by population subgroups. *Journal of Official Statistics*, **30**, 91–105 (2014)
28. Pittau, M.G., Yitzhaki, S., Zelli, R.: The ”make-up” of a regression coefficient: gender gaps in the European labor market. *Review of Income and Wealth*, **61**, 401–421 (2015)
29. Schneebaum, A., Rehm, M., Mader, K., Hollan, K.: The gender wealth gap across European countries. *Review of Income and Wealth*, **64**, 295–331 (2018)
30. Selezneva, E.; Van Kerm, P.: A distribution-sensitive examination of the gender wage gap in Germany. *Journal of Economic Inequality*, **14**, 21–40 (2016)
31. Yitzhaki, S., Lerman, R.: Income stratification and income inequality. *Review of Income and Wealth*, **37**, 313–329 (1991) Mussard Erbert Ogowang



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
<http://www.dse.unibo.it>