GENDER INEQUALITIES FROM AN INCOME PERSPECTIVE

Marcella Mazzoleni[1], Angiola Pollastri[2] and Vanda Tulli[2]

[1] Center on Economic, Social and Cooperation dynamics (CESC), University of Bergamo, (e-mail: marcella.mazzoleni@unibg.it)

[2] Department of Statistics and Quantitative Methods, University of Milano-Bicocca (e-mail: angiola.pollastri@unimib.it, vanda.tulli@unimib.it)

ABSTRACT: The difference between females' and males' income is one of the main topics in the analysis of gender gap, as it is known that, even with a higher educational level, females earn less than males do. To inspect this, we analyse and estimate the distribution of the ratio of females' income over males' income using the methodology based on the distribution of the ratio of two Dagum with three parameters. We applied this method to the Bank of Italy Survey on Household Income and Wealth (SHIW) data to evaluate the deciles, the density functions, and the cumulative distribution functions of the ratio of the females' income over males' income in different age classes, Italian areas, and years.

KEYWORDS: Dagum distribution, distribution of the ratio of two Dagum random variables, distribution of the ratio of female and male income.

1 Introduction and method

It is well known that even with a higher educational level, women earn less than men do. The differences between men' and women' income on average are decreasing in the recent years but income parity has not yet been achieved.

The purpose of this paper is to estimate the distribution of the ratio of females' income over males' income. The methodology used to study the ratio is based on the distribution of the ratio of two Dagum with three parameters (Pollastri and Zambruno 2010). The distribution of this ratio studied in two different situations can reveal the gender inequality concerning income in different groups or times, accordingly the distribution of the ratio is analysed to reveal the gender inequality with applications to the income in different age classes, areas, and times.

In literature we have many examples which confirm that the model proposed in 1977 by Camilo Dagum fits very well to many distributions of economic variables. Supposing that X is a type I Dagum, then $X \sim D(a, b, p)$ with $a, b, p > 0$. The distribution function for $x > 0$ is defined as (Kleiber and Kotz 2003):

$$F_X(x) = \left[1 + \left(\frac{x}{b}\right)^{-a}\right]^{-p}$$

While the density function for $x > 0$ is:

$$f_X(x) = \frac{apx^{ap-1}}{b^{ap}\left[1 + \left(\frac{x}{b}\right)^a\right]^{p+1}}$$

The Dagum distribution parameters are estimated using the function *dagum* implemented in the VGAM package in the software R. This function estimates the parameters using the maximum likelihood estimation method proposed by Kleiber and Kotz (2003). Domanski and Jedrzejczak (1998) showed, through a simulation study, that estimation method performance is good for a and p when $n > 2000$ or 3000, while for the scale parameter b the bias tends to 0 when $n > 4000$.

The purpose of this paper is to analyse the ratio:

$$U = \frac{X}{Y}$$

where $X \sim D(a_1, b_1, p_1)$ and $Y \sim D(a_2, b_2, p_2)$ with X and Y independent.

Following the definition of the density function of the ratio of two random variables in Mood, Graybill and Boes (1974), applying the independence of X and Y and the density function of a type I Dagum, it is possible to obtain the density function for the ratio U:

$$f_U(u) = \int_0^{+\infty} y \left\{ \frac{a_1 p_1 (uy)^{a_1 p_1 - 1}}{b_1^{a_1 p_1}\left[1 + \left(\frac{uy}{b_1}\right)^{a_1}\right]^{p_1+1}} \right\} \times \left\{ \frac{a_2 p_2 y^{a_2\, p_2 - 1}}{b_2^{a_2 p_2}\left[1 + \left(\frac{y}{b_2}\right)^{a_2}\right]^{p_2+1}} \right\} dy$$

Using the definition of the cumulative distribution function, it is possible to obtain:

$$F_U(u) = \frac{a_1 p_1 a_2 p_2}{b_1^{a_1 p_1} b_2^{a_2 p_2}} \int_0^u t^{a_1 p_1 - 1} \int_0^{\infty} y^{a_1 p_1 + a_2\, p_2 - 1} \left[1 + \left(\frac{ty}{b_1}\right)^{a_1}\right]^{-p_1 - 1} \times$$

$$\left[1 + \left(\frac{y}{b_2}\right)^{a_2}\right]^{-p_2 - 1} dydt$$

In Pollastri and Zambruno (2010) a graphical analysis of this method performance is exposed comparing the empirical and the computed density function, where the empirical one is created with the ratios of all the possible couples.

2 Applications to Survey on Household Income and Wealth

We apply this method to the individual net incomes in 2016 from the Bank of Italy Survey on Household Income and Wealth (SHIW). We compare the ratio of the females' and males' income in different groups:

- males and females divided in three age classes: *young* $(age < 40)$, *adult (*$40 \leq age < 70$) and *old (*$age \geq 70$)

- males and females divided in three areas: *North, Centre,* and *South and Islands*
- males and females in two different years: 2016 and 1998

The dataset is composed of 11,844 subjects. Of these 50.98% are males, and 49.02% are females. Concerning the division by ages 15.21% are aged less than 40 years, 54.41% are aged between 40 and 70 years, and 30.39% are aged equal or more than 70 years. For the division by area, 43.90% of the subjects come from the North, 22.40% from the Centre, and 33.70% from the South and the Islands. The 1998 dataset, it is composed of 12,616 subjects, of these 56.52% are males and 43.48% are females.

After estimating the Dagum parameters, we evaluate the cumulative distribution function and the deciles of the ratio of the females' income over the males' income, comparing the results of the ratio distributions for different ages, areas, and times.

Comparing the ratio of the females' income over males' income in different age classes, we observe higher value of deciles for younger subject, lower for adult group, and even lower for the older group. This confirms that the income at the beginning of the career is similar between the two genders but increasing the age and the position achieved, the gap rises.

For the deciles of ratio of the females' income over males' income comparing different areas, we observe close and higher value for the subjects that live in North and Centre of Italy, and lower value for the subjects that live in South and Islands. This can be related to the different economical and social situation in the Islands and in the South of Italy.

We observe that the deciles of the ratio of the females' income over males' income are higher in 2016 with respect to the ratio in 1998. This confirms that the differences between men' and women' income are decreasing in the recent years, but income parity has not yet been achieved.

3 Conclusions

In this paper we propose to use the ratio of two type I Dagum random variables for analysing the difference of the income of females and males. We observe that this method gives us interesting conclusions and can be applied to different dataset comparing also the ratio of females' over males' income in different countries, in order to highlight the differences concerning gender gap.

This method is used to analyse the Italian situation and to compare the ratio of females' over males' income in different ages, areas and times. As a matter of fact, in the applications we observe less diversity for females' and males' income in the younger group, but the diversity rises increasing the subjects' age, passing from young to adult, and from adult to old group. In the division by areas, the deciles of the ratio of females' income over males' income for the North and Centre are close, while for the South and Islands a wider difference between genders is observed. The difference of the income for males' and females' is decreasing over the years. In fact, the deciles

of the ratio of females' income over males' income are higher for 2016 dataset than the deciles of the 1998 dataset.

References

ALLEVA G. (2017). Indagine conoscitiva sulle politiche in materia di parità tra uomini e donne. Istat, Rome.

BANK OF ITALY (2016). I bilanci delle famiglie italiane nell'anno 2016. Supplementi al Bollettino Statistico, XVII, Centro Stampa Banca d'Italia, Roma.

BANK OF ITALY (1998). I bilanci delle famiglie italiane nell'anno 1998. Supplementi al Bollettino Statistico, XVII, Centro Stampa Banca d'Italia, Roma.

DAGUM C. (1977). A new model of personal income distribution: specification and estimation. *Economie Appliquée*, **30** (3), 413-437.

DAGUM C. (1990). Generation and properties of income distribution functions. [In:] *Studies in Contemporary economics. Income and wealth distribution, inequality and poverty*, C. Dagum, M. Zenga (Eds.), Springer, Berlin.

DOMANSKI, C., JEDRZEJCZAK A. (1998) Maximum likelihood estimation of the Dagum model parameters. *International Advances in Economic Research*, **4**, 243–252

KLEIBER, C., KOTZ S. (2003). Statistical Size Distributions in Economics and Actuarial Sciences. Hoboken, NJ, USA: Wiley-Interscience.

MOOD A., GRAYBILL F., BOES D. (1974). Introduction to the theory of statistics. Wiley, New York.

POLLASTRI A., ZAMBRUNO G. (2010) Distribution of the ratio of two independent Dagum random variables. *Operations Research and Decisions*, **3** (20), 95-102.

YEE T. (2019). VGAM: Vector Generalized Linear and Additive Models. R package.