

Invited paper to the 56th SIEDS Scientific Meeting – Ascoli Piceno 2019

THE GINI CONCENTRATION RATIO: BACK TO THE FUTURE

Giovanni Maria Giorgi

1. Introduction

On 29th May 1914, in the context of the Scientific Meeting of the Royal Venetian Institute of Sciences, Letters and Arts, Corrado Gini proposed a new synthetic measure of inequality which he called concentration ratio (G). Perhaps it will seem a bit strange that, more than a century later, scholars continue to discuss about G . Despite the weight of the years this index is still widely used not only in economics but also in fields that were unthinkable when it was originally proposed. In fact, new ways of using G actualize the past and form the basis for further future developments.

Inequality and its measurement are relevant because the perception of inequality itself influences both political and citizen choices.

One of the aspects of economic inequality that has negative repercussions on social status is, for example, the diversity of cognitive abilities of older people. That is, in other words, elderly people - in a weak socio-economic state and in poor health - show reduced intellectual capacity (Dal Bianco et al., 2013).

Aging is one of the main socio-economic problems of the 21st century. In this context the decline of cognitive abilities is relevant. Several factors have an impact on this decline, such as for instance retirement. Education also has a significant influence on the level of cognitive skills (Mazzonna-Peracchi, 2017).

The careful identification of these problems can be of great help for the health of the inhabitants of a country. In fact, if policies are promptly started to reduce the gender gap in education and the distribution of education among young people is improved, it is possible to reduce inequality of cognitive functions in the future (Oliveira et al., 2018). The information must be available quickly and at a cost that is not excessive and this leads to an increasing use of sample surveys. When the inequality is calculated on sample data, before generalizing the results, it is necessary to know and take into consideration the sample characteristics of the inequality index used, so that the information obtained constitutes a valid and reliable support for the

decisions of economic policy for reducing inequality (see, e.g., Giorgi, 1999; Giorgi and Gigliarano, 2017)

To preserve confidentiality, official statistical agencies publish the data in a grouped form. This aspect has also been taken into consideration by Gini (1914) who proposes various formulas for the calculation of his index and suggests different solutions to the problems that arise when the lower limit of the first class and / or the upper limit of the last class are not known. On the basis of some empirical tests, Gini also shows how the distortion of the estimate of his index increases with the increase in concentration and class size.

2. Genesis of the Gini concentration ratio

Now let us recall how Gini index was originally proposed. Between the end of the 19th century and the beginning of the 20th century, studies aimed at achieving a more equitable distribution of wealth characterize the political and economic debate. Vilfredo Pareto (1895, 1897) and Corrado Gini (1909) contribute to this type of study. The indices proposed by the two scholars give rise to a heated debate on their advantages and disadvantages, but Gini - despite considering his index better than that of Pareto - continues the studies on the subject, arriving in 1914 to the definition of the so-called concentration ratio also known in the literature as Gini coefficient, Gini ratio or simply as Gini index. He also highlights the connection with the Lorenz diagram (1905) and the mean difference (Gini, 1912), providing adequate empirical applications. In fact, he avoids excessive formalization and tries, first of all, to discover the underlying logic of events; only afterwards and for what is strictly necessary he uses mathematical tools. Gini's scientific production is not the result of a theoretical-formal elaboration as an end in itself, but comes from the need to solve concrete problems that arise in empirical applications ranging from economics to sociology, to demography and biology.

Despite the efforts of Gini and his collaborators to support the use of the concentration ratio, the index arouses the interest of non-Italian scholars only after some time its appearance in the literature. This can be deduced from an article by Gini (1921) published in the *Economic Journal* in which he points out that Dalton's (1920) proposal to measure inequality of economic welfare cannot be compared to the methods introduced by Italian scholars, including the concentration ratio.

In order to spread his index at an international audience, Gini (1926) publishes in the *Journal of the Royal Statistical Society* a paper in which he highlights the contributions of Italian scholars to modern statistical methods. Furthermore, his conferences abroad always underline the importance of G and the other results of the so-called Italian statistical school. One of the difficulties in spreading these results

is due to the fact that Gini and his collaborators wrote almost exclusively in Italian and, therefore, quite often published in journals with a limited audience. The synthesis of these results provided by Gini, also through prestigious journals, was not sufficient for a widespread diffusion and a thorough knowledge of these results. These syntheses were essentially—on Gini's work, since he believed his results coincided with Italian Statistics. At that time, almost all the main scientific results in Statistics by Italian scholars were achieved by him and his collaborators; and the latter limited themselves to developing and deepening Gini's suggestions (see, e.g., Giorgi and Gubbiotti, 2017).

3. New applications and extensions

The computational simplicity of Gini coefficient (G) have stimulated new applications and extensions (see, e.g., Giorgi, 2020) In fact, for example, G has been used in criminology for measuring and reporting crime concentration. Bernasco and Steenbeek (2017) suggested a generalized Lorenz curve and Gini coefficient in order to correct the bias when crimes are sparse, that is when there are fewer crimes than places. For further developments of the topic see also Bernasco et al. (2017). Very recently, an interesting application of this methodology has been made by Hasisi et al. (2019) for investigating the spatial characteristic of vehicular terror attacks in Israel. The authors analyze the 71 vehicular attacks carried out in Israel between 2000 and 2017. They identify the “hot routes” at which attacks occurred as well as the estimated journey to attack routes.

Another application of G is to study the distribution of citations in scientific articles and therefore to evaluate the scientific productivity of departments, universities and countries from a qualitative and quantitative point of view (Rousseau, 1998). In particular, Zheng et al. (2008) use Gini coefficient to reflect the inequality degree of publications in China and other countries. Gini index is also used to measure the inequality of political representation among voters (Pretolani, 2014), as well as a tool to assess some aspects of a survival table (Shkolnikov et al., 2003).

Other interesting applications of the concentration ratio are in geology, where it is used to detect the damage caused by earthquakes to buildings thus providing an important contribution to the evaluation of disasters (Tu et al., 2017).

In robotics, the use of multi-robot systems in very dangerous and polluted environment allows the use of Gini coefficient for evaluating the available energy for robots (Wu et al., 2018).

Gini index is also applied to evaluate the impact of human activities on the hydrological cycle (Zhang et al., 2015). In other cases, G is used to investigate

insurance issues, for example to assess whether an alternative insurance score is useful for detecting differences between loss distributions and premiums (Frees et al., 2011; 2014).

The concentration ratio, according to some very recent studies (e.g., O'Hagan, 2018), may contribute to the development of biology and medicine, for example to study the distribution of proteins between different cells. Furthermore, Gini index also offers the possibility to analyze the collection of pharmacological data and the risk management strategy for new drugs (Torres-Garcia et al., 2017).

The rapid development of Internet has given a great contribution to the so-called big data era. In this case, it is relevant to understand which part of this huge information has to be considered to provide a correct interpretation of the data. To overcome this problem, it is possible to use systems that filter information and provide customized results to different users. These techniques are used by economic operators and social networks to have the user's contact information and help him/her to find new friends by connecting to various sites (such as Facebook). In this context, the crowding that occurs when some objects are recommended to many users can play a substantial role. A significant contribution to the solution of this problem is given by the Gini index that allows to evaluate the ability of an algorithm to avoid crowding (Ren et al., 2014).

4. Some further considerations on G

Gini index has aroused the interest of scholars of different disciplines such as, for example, economists, statisticians, biologists, sociologists and even mathematicians. This attention by scholars with different cultural backgrounds has inevitably given rise to a heated debate that starts from the appearance in the literature of G .

This debate was animated by scholars of the caliber of Antony Atkinson and Amartya Sen. The former (Atkinson, 1970) criticizes G and other indices of inequality because they do not rank income distributions according to strictly concave social utility functions. Sen (1972, 1973) points out that Atkinson's conclusions are framed in the utility functions context, the limits of which are known and highlighted by Sen himself. According to other scholars, the introduction of a utility function has replaced the choice of a measure of inequality, with another difficult problem to solve: the choice of a utility function. Furthermore, Atkinson's distinction between objective and normative measures is not justified, as evidenced, for example, by Giorgi (1984), Giorgi and Pallini (1985) and Muliere (1987).

A criticism that is put forward to the Gini index is that it cannot be decomposed in an additive way (see, e.g., Giorgi, 2011), that is the decomposition is not of analysis of variance type (within + between). In fact, the breakdown of the Gini index

also includes a third component (overlapping) of residual type (i.e. within + between + overlapping). According to Dagum (1977), Lambert and Decoster (2005) and Sen (1999) the latter type of decomposition should be preferred to that of the "analysis of variance" type because it clarifies some of the aspects related to inequality between the groups, adapting rather well to the complex socio-economic reality in which we live. In this regard, Yitzhaki (1994) interprets the overlapping component as a measure of stratification between socio-economic groups.

The problem of decomposing the Gini index can be overcome by the Shapley decomposition. This method removes, one at a time, the contribution of all the possible combinations of each factor, so that the sum of these contributions gives the exact value of the index of inequality considered (Deutsch and Silber, 2007; Shorrocks, 1999)

Finally, when the Lorenz curves corresponding to different income distributions intersect and enclose concentration areas of the same size, the Gini index has the same value. Some researchers, among others Giurovich (1959), Hagerbaumer (1977) and Patimo (1977), proposed to overcome this stalemate with further information. The first considers an asymmetry index of the Lorenz curve (*LC*); the second integrates *G* with an index that measures the relative position of the poor and, the third suggests using together with *G* another index based on *LC* divided into two parts which takes into account both the difference in shape of these parts and the size of the difference (see, e.g., Giorgi 1992).

In conclusion we can say that the concentration ratio, after more than one hundred years since its appearance in the literature, is far from being enveloped by the mists of oblivion. Indeed, as highlighted above, the extensions, the new interpretations and above all the new applications that were unthinkable not only in 1914 but also twenty years ago make the Gini index extremely topical.

References

- ATKINSON, A.B. 1970. On the measurement of inequality, *J. Economic Theory*, Vol. 2, pp. 244-263.
- BERNASCO W.- STEENBEEK W. 2017. More places than crimes: Implications for evaluating the law of crime concentration place, *Journal of Quantitative Criminology*, Vol. 33, pp. 451-467.
- BERNASCO W., DE GRAF T., ROWENDAL J., STEENBECK W. 2017. Social interactions and crime revisited: An investigation using individual offender data in Dutch neighborhoods, *The Review of Economics and Statistics*, Vol.99, No. 4 , pp. 622-636

- DAL BIANCO, C., GARROUSTE, C., PACCAGNELLA O. 2013. Early-life circumstances and cognitive functioning dynamics in later life. In: A. Börsch-Suban, M. Brandt, H. Litwin, & G. Weber (Eds) *Active aging and solidarity between generation in Europe* (pp. 209-223). De Gruiter, ISBN 978-3-11-029546-7
- DAGUM C. 1977. A new model of personal income distribution: specification and estimation, *Economie Appliquée*, Vol. 30, pp. 413-437.
- DALTON H. 1920. The measurement of the inequality of incomes. *Economic Journal*, Vol. 30, pp. 348-361.
- DEUTSCH J., SILBER J. 2007. Decomposing income inequality by populations subgroups: a generalization. In J.A Bishop and Y. Amiel (Eds), *Inequality and poverty* (pp. 237-253), Berlin: Springer.
- FREES E.W., MEYERS G., CUMMINGS A.D. 2011. Summarizing insurance scores using a Gini index. *J. Am. Stat. Assoc.*, Vol. 106, pp. 1085-1098.
- FREES E.W., MEYERS G., CUMMINGS A.D. 2014. Insurance ratemaking and a Gini index. *J. Risk Insur.*, Vol. 81, pp. 335-366.
- GINI C. 1909. Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza. *Giornale degli Economisti*, Vol. 20, pp. 27-83.
- GINI C. 1912. Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche, *Studi Economico-Giuridici della Facoltà di Giurisprudenza della Regia Università di Cagliari*, 3, Bologna, Italy: Cuppini.
- GINI C. 1914. Sulla misura della concentrazione e della variabilità dei caratteri, *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 73, 1203-1248. English translation in *Metron* (2005), pp. 3-38.
- GINI C. 1921. Measurement of inequality of incomes. *Economic Journal*, Vol. 31, pp. 124-126.
- GINI C. 1926. The contributions of Italy to modern statistical methods. *J. Royal Statistical Society*, Vol. 89, pp. 703-724
- GIORGI G.M. 1984. A methodological survey of recent studies for the measurement of inequality of welfare carried out by some Italian statisticians. *Economic Notes*, Vol. 13, pp. 146-157.
- GIORGI G.M. 1992. *Il rapporto di concentrazione di Gini: genesi, evoluzione ed una bibliografia commentata*, Siena, Italy: Libreria Editrice Ticci.
- GIORGI G.M. 1999. Income inequality measurement: the statistical approach. In J. Silber (Ed), *Handbook on Income Inequality Measurement* (pp. 245-260). Boston: Kluwer Academic Publishers.
- GIORGI G.M. 2011. The Gini index decomposition: An Evolutionary Study. In J. Deutsch and J. Silber (Eds.), *The measurement of individual well-being and group inequalities. Essays in memory of Z.M. Berrebi*, (pp. 185-218). Abingdon (GB): Routledge. 85, 325-339.

- GIORGI G.M. 2020. Gini coefficient. In P. A. Atkinson, A. Cernat, S. Delamont, J. W. Sakshaug, R. Williams (Eds.), SAGE Research Methods Foundations. [DOI: 10.4135/] *in press*.
- GIORGI G.M., GIGLIARANO C. 2017. The Gini concentration index: a review of the inference literature. *Journal of Economic Survey*, Vol. 31, pp. 1130-1148.
- GIORGI G.M., GUBBIOTTI S. 2017. Celebrating the Memory of Corrado Gini: a personality out of the ordinary. *International Statistical Review*, Vol. 85, pp. 325-339.
- GIORGI G.M.- PALLINI A. 1985. Di alcune misure di disuguaglianza del benessere e della loro determinazione mediante trasformazione di variabile. *Note Economiche*, Vol. 18, pp. 182-194.
- GIUROVICH G. 1959. Su un indice approssimato di asimmetria della curva di concentrazione. *Statistica*, Vol. 19, pp. 463-482
- HAGERBAUMER J.B. 1977. The Gini concentration ratio and the minor concentration ratio: A two-parameter Index of inequality. *Review of Economics and Statistics*, Vol. 59, pp. 377-379
- HASISI B., PERRY S., ILAN Y., WOLFOWICZ M. 2019. Concentrated and close to home: the spatial clustering and distance decay of lone terrorist vehicular attacks, *Journal of Quantitative Criminology*, Doi: 10.1007/s10940-019-09414-z
- LAMBERT P.J., DECOSTER A. 2005. The Gini coefficient reveals more. *Metron*, Vol. 63, pp. 373-400
- LORENZ M.O. 1905. Methods of measuring the concentration of wealth. *Quarterly Publications of the American Statistical Association*, Vol. 9, pp. 209-219.
- MAZZONNA F., PERACCHI F. 2017. Unhealthy retirement? *Journal of Human Resources*, Vol. 52, pp. 128-151.
- MULIERE P. 1987. Misure della concentrazione: alcune osservazioni sull'impostazione assiomatica. In M. Zenga (Ed), *La distribuzione personale del reddito: problemi di formazione, di ripartizione e di misurazione* (pp. 190-213), Milan, Italy: Vita e Pensiero.
- O'HAGAN S., MUELAS M.W., DAY P.J., LUNDERBERG E., KELL D.B. 2018. GeneGini: assessment via the Gini coefficient of reference "housekeeping genes and diverse human transporter expression profiles". *Cell Systems*, Vol. 6, pp. 230-234.
- OLIVEIRA J., ANDREOLI F., LEIST A.K., CHAUVEL L. 2018. Inequality in old age cognition across the world. *Economics and Human Biology*, Vol. 29, pp. 179-188.
- PARETO V. 1895. La legge della domanda. *Giornale Degli Economisti*, Vol. 10, pp. 59-68.
- PARETO V. 1897. Aggiunta allo studio sulla curva delle entrate. *Giornale degli Economisti*, Vol. 14, pp. 15-26.

- PATIMO A. 1977. Su un indice specificativo del rapporto di concentrazione. *Rivista Italiana di Economia, Demografia e Statistica*, Vol. 31, pp. 333-350
- PRETOLANI D. 2014. Apportionments with minimum Gini index of disproportionality: a quadratic knapsack approach. *Ann. Oper. Res.*, Vol. 215, pp. 257-267.
- REN X., LÜ L., LIU R., ZHANG J. 2014. Avoiding congestion in recommender systems. *New Journal of Physics*, Vol. 16, pp. 1-19
- ROUSSEAU R. 1998. Evenness as a descriptive parameter for Department or Faculty evaluation studies. In E. de Smet (Ed) *Werkgemeenschap Informatiewetenschap* (pp. 135-145), Antwerp, Belgium.
- SEN A. 1972. Utilitarianism and inequality. *Economic and Political Weekly*, Vol. 7, pp. 343-344.
- SEN A. 1973. *On economic inequality*, Oxford (GB): Clarendon Press
- Sen, A. 1999. Foreword. In J. Silber (Ed), *Handbook on Income Inequality Measurement*, (pp. xvii-xxvi). Boston: Kluwer Academic Publishers
- SHKOLNIKOV V.M.- ANDREEV E.E.- BEGUN A.Z. 2003. Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples. *Demographic Research*, Vol. 8, pp. 305-358.
- SHORROCKS A.F. 1999. Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. *Unpublished Paper*, University of Essex, Department of Economics.
- TORRES-GARCIA B.L., CASTRO-PASTRANA L.I., RODRIGUEZ-RODRIGUEZ S., ESTRADA-MARIN L., CEDILLO-CRAVALLO B., GUZMAN-GARCIA O., RUÍZ-ARGÜELLES F. 2017. Gini index-based maximum concentration and area under the curve split points for analysing adverse event occurrence in bioequivalences studies. *Farmaceutical Medicine*, Vol. 6, pp. 51-56.
- TU J., SUI H., FENG W., SUN K., XU C., HAN Q. 2017. Detecting building façade damage from oblique aerial images using local symmetry feature and the Gini index. *Remote Sensing Letters*, Vol. 8, pp. 676-685.
- WU D., ZENG G., MENG L., ZOU W., LI L. 2018. Gini coefficient-based task allocation for multi-robot systems with limited energy resources. *IEEE/CA A Journal of Automatica Sinica*, Vol. 5, pp. 155-168.
- YITZHAKI S. 1994. Economic distance and overlapping of distributions. *Journal of Econometrics*, Vol. 61, pp. 147-159.
- ZHANG Q., GU X., SINGH V.P., XU C.Y., KONG D., XIAO M., CHEN X. 2015. Homogenization of precipitation and flow regimes across China: changing properties, causes and implications. *Journal of Hydrology*, Vol. 530, pp. 462-475.
- ZHENG M., JUMPENG Y., CHENG S., ZHIYU Y., ZHENGLU Y., YUNTAO P. YISHAN W. 2008. Using Lorenz curve and Gini coefficient to reflect the

inequality degree of S&T publication: An examination of the institutional distribution of publications in China and other Countries, *4th International Conference on Webometrics, Informetrics and Scientometrics and 9th COLLNET Meeting*, Humboldt-Universität zu Berlin, Berlin.

SUMMARY

The Gini concentration ratio: Back to the future

The topicality of the Gini index, over 100 years after its appearance in the literature, is underlined by showing how it is currently used in fields other than economics, such as criminology, hydrology, insurance, biomedicine, geosciences and to evaluate scientific productivity of universities and departments.

