

OPTIMIZATION OF GINI COEFFICIENT AFFECTED BY IMPERFECT INPUT DATA

Petr David¹

¹*Mendel University in Brno, Czech Republic*



EUROPEAN JOURNAL
OF BUSINESS SCIENCE
AND TECHNOLOGY

Volume 5 Issue 1
ISSN 2694-7161
www.ejobsat.com

ABSTRACT

Most indicators used for determining the distributional effects of taxes as well as the inequality in the income distribution are based on the Gini coefficient and the Lorenz curve to a substantial extent, although the potential application of the Gini coefficient itself is much larger. However, the Lorenz curve and in particular the Gini coefficient need not present precise information on income or the distribution of wealth in a society. The Gini coefficient values may be affected by the form of the input data. We have ascertained that the level of Gini coefficient distortion depends on the number of households included in the research given that the income distribution in the sample is unequal. In addition, we define the form of the Gini coefficient in light of the form of the input data.

KEY WORDS

Gini coefficient, Lorenz curve, distribution, income, taxation

JEL CODES

H22, H23

1 INTRODUCTION

The numeric indicator of the income or wealth distribution in a society is called the Gini coefficient and was first formulated by Corrado Gini in 1912 (Gini, 1912). The curve that served as the basis for Gini's calculations had been developed by Max Otto Lorenz in 1905 (Lorenz, 1905). This curve represents the cumulative income of households in the context

of the cumulative number of households in the economy in question.

The Gini coefficient is, in combination with the Lorenz curve, widely used as an instrument for the measurement of the impact of tax policy on the distribution of income or wealth in a society. However, the Lorenz curve, and especially the Gini coefficient, do not provide precise

information about certain specific situations (David, 2017). The Gini coefficient, for example, is unable to take into account the specifics of a particular form of income distribution. This imperfection can be compensated for with the Lorenz curve, which shows the exact shape of the distribution of income or wealth in the society. However, Gini coefficient values may also be negatively affected by the number of households, provided that income is unequally distributed among the population. In this case, the Lorenz curve cannot be used to correct for the deficiencies of the Gini coefficient.

This paper focusses on the parameters, which influence the Gini coefficient (e.g. the number of households and the unequal distribution of income) in an economy. We aim to formalise the Gini coefficient based on the deviations from expected Gini coefficient values, which result from the real-world nature of the input data, in order to exclude the before mentioned input data parameters. We base our calculations on the definition of the deviation from the expected value of the Gini coefficient.

2 THEORETICAL BACKGROUND

2.1 Essence and Significance of the Gini Coefficient and Lorenz Curve

The impacts of tax policy can be examined through either the direct and indirect tax incidence or as a property of wealth (Sen, 1976). The significance of indirect tax incidence is apparent from the studies by Besley and Rosen (1999), Boeters et al. (2010) and David (2012). The measurement of direct tax incidence is equally important. The Gini coefficient and the Lorenz curve are two tools for measuring the effects of the tax progressivity of direct taxes.

One of the factors which affects the analysis of tax progressivity is the selected period. When measuring tax progressivity, taxes can be assessed from the perspective of a specific year or of a life cycle. The importance of the lifetime approach is stressed by Poterba (1989) and Fullerton and Rogers (1994). For a concise overview, of the lifetime approach to tax progressivity, we refer the reader to Metcalf and Fullerton (2002). Empirical research in this area was conducted by Caspersen and Metcalf (1994) and Metcalf (1994). Theoretical studies draw attention to the substantial differences between the annual and lifetime approaches, annual data are more commonly used. This is also the case for the Lorenz curve and Gini coefficient.

Global tax progressivity is almost exclusively graphed as the Lorenz curve. This method

analyses the income distribution and the cumulative proportion of the total national income set off against the cumulative proportion of taxpayers. The Lorenz curve measures the impact of tax rate changes on the redistribution of the real disposable income of households in an economy. In case a Lorenz curve lies below another Lorenz curve, the former distribution leads to more equality than the latter (Fellman, 1976). The Lorenz curve is therefore often used to compare the income distributions before and after-tax rate changes. Where the curve approximates the line of equality, a progressive tax system is in place.

The Gini coefficient is a measure of income inequality (Gini, 1912). It is calculated on the basis of the discrepancy in the Lorenz curve. The coefficient may also be used for to compare different alternatives of the income redistribution within one country, as well as internationality. The Gini coefficient assumes values between 0 and 1. The resulting value of “zero” indicates an entirely equal distribution of incomes. On the other hand, the resulting value of “one” can be seen as a completely unequal distribution. The lower the Gini coefficient, the more equally income is distributed among various groups of taxpayers. The Gini coefficient can only be used for monitoring tax progressivity through comparison of income before and after tax.

2.2 Relation between the Gini Coefficient and Other Indicators of Tax Progressivity

Besides the Gini coefficient and Lorenz curve, global tax progressivity can also be measured with the Musgrave and Thin method, the Kakwani method, the Suits method, the Reynolds-Smolensky index, the Robin Hood index, entropy methods, the Atkinson index and the index of tax progressivity mentioned by Široký and Maková (2009). Most of the above-mentioned alternative indicators are directly or indirectly based on the Lorenz curve and the Gini coefficient and are therefore depended on the type of input data.

Musgrave and Thin (1948) use the Lorenz curve as a function related to the distribution of incomes before and after tax. The Musgrave and Thin index derives from the ratio of the area below the curve of the function after tax set off against the area below the curve before tax. The index is structured on the basis of the Gini coefficient before and after taxes. Although the efficacy of the Musgrave and Thin index is indisputable, it was criticized by Bracewell-Milnes (1979) on the grounds that it is unable to identify the deviation from proportionality.

Kakwani's method is also based on the Lorenz curve (Kakwani, 1977). The Kakwani index is derived from the ratio of the area between the Lorenz curve function and the function of the concentration curve below the line of equality. To determine the concentration coefficient, it is necessary to carefully classify households according to income levels (Hoffmann, 2012). The Kakwani index may be understood as a function of the Gini coefficients of the concentration curve and the Lorenz curve. Bracewell-Milnes (1979) points out that Kakwani's difference between the progressivity and the amount of the tax is less relevant compared to the difference between progressivity and the inequality of income before tax. To defend his index, Kakwani (1979) states that it focuses on the progressivity of taxes incidence and public expenditure. In addition, according to Kakwani, helps to analyze the effects of various types of taxes as well as government expenditure.

The Suits index measures the dependence of the cumulative proportion of tax liability on the cumulative proportion of incomes before tax (Suits, 1977). This index is not directly based on the Lorenz curve or the Gini coefficient; however, it was designed on the same principles. The line of equal distribution represents the situation in which tax is proportional. Where progressive tax is applied, the function deviates to the right of the straight line of equality. The area between the line of equality and the function concerned indicates the degree of deviation. Although the Suits index and the Kakwani index are similar, they lead to different tax progressivity assessments in cases, in which the source data are incomplete. This fact reveals the weakness of the Kakwani index and, by extension, the Gini coefficient.

Another indicator of global tax progressivity is the Reynolds-Smolensky index (Reynolds and Smolensky, 1977). Similarly, to the Musgrave and Thin index, its focus is the redistributive effects of taxes. The index is defined as the difference between the inequality of after-tax incomes and inequality of pre-tax incomes. The index may also be expressed through the Kakwani index if the total average tax rate and the Kakwani index are known. The Reynolds-Smolensky index is influenced by the tax rate and tax progressivity based on the Kakwani index.

The extend of redistribution that would lead to complete equality of income in an economy can be expressed with the Hoover index, also known as the Robin Hood index (Hoover, 1936). The index is related to the Gini coefficient and is based on the same principle. It measures the distance between the Lorenz curve and the line of equality. The resulting index value shows the percentage of overall income that has to be redistributed in order to achieve equality in the income distribution, provided that taxpayers' incomes as well as their mean value are known. The index can also be used for the measurement of progressivity through the quantification of the difference between its values before and after taxation.

Other important methods of tax progressivity measurement do not have any direct or

indirect link to the Gini coefficient and we will therefore only provide a brief overview. These are primarily entropy methods, which estimate the difference between entropy indexes on the basis of the distribution of income before and after tax (Zandvakili, 1991). The most frequently used entropy index is the Theil index of inequality (Theil, 1967). In addition, there are indicators of the mean logarithmic deviation, squared coefficient of variation, and generalized entropy (Kesselman and Cheung, 2004). The Atkinson index, which is based on

the calculation of a fair average per capita income (Atkinson, 1970), forms the foundation for two additional but largely similar progressivity indicators. The first, the Kiefer index (1984), uses the Atkinson index before and after taxation. Blackorby and Donaldson (1980) use the same indicators as Kiefer (1984), but in a slightly different way. Finally, we should mention the tax progressivity index which is composed of a quotient of volatility of tax revenues and incomes (Kakinaka and Pereira, 2006).

3 METHODOLOGY AND DATA

This theoretical overview of the measurements of the distributional effects of tax and the estimation of the distribution of income in a given economy, using the Lorenz curve and the Gini coefficient, aims to present new findings and the contexts in which they are relevant. The research is based on the generation of simple models which enable a more comprehensive study of the income distribution in an economy, without the necessity to collect concrete data that may prevent general conclusions.

The application of models generally presupposes the use of data concerning household incomes. This approach to the measurement of tax policy impacts was criticized by Fullerton and Rogers (1991), who prefer to approach the problem from the perspective of the incomes of individuals. However, if such models consistently used household data, such values may be considered representative of the incomes of individuals, with little risk of distorting results or their interpretation.

Standard mathematical methods are used to define model situations through equations, which are adjusted and solved according to needs. The definitions of functional relationships and limits of sequences of values are considered as well.

This research does not involve practical applications and does not use, and therefore does not require, empirical data. This work merely aims to redefine the formula of the Gini coefficient for further use, while eliminating its hitherto imperfections, i.e. the distortions its output due to the range of the examined sample and the degree of equality in the income distribution in an economy. The adjustments proposed by alternative methods for examining the distributional effects of taxation and the income distribution in an economy, point at shortcomings in the Gini coefficient and related indexes, which derive from imperfect input data with regard to the number of included households and the inequality in the income distribution in an economy.

4 RESULTS OF ELIMINATION OF THE GINI COEFFICIENT DISTORTION

Based on our previous points, it can be stated that the importance of the Gini coefficient and the Lorenz curve goes beyond their frequent application. Many, if not most, of the alternative indicators of tax progressivity and the

inequality of income distribution are based on the Lorenz curve and the Gini coefficient to a certain extent, and therefore are equally influenced by input data. However, the possibilities of the Gini coefficient use are much broader.

Evidence for this for instance derives from the measurements of the intersectoral digital divide, which result from differences in the use of information systems (Fidan, 2016). Therefore, to insure that the results reflect the real state of affairs and not just the optimal situation based on optimal input data, the form of the Gini coefficient and Lorenz curve must be defined as precisely as possible.

The Lorenz curve is usually depicted in a uniform manner. However, there is no way of removing the distortion caused by the limited number of households included, and the unequal distribution of income among those households from this graphic representation of the income distribution. This imperfection should, however, not lead to refraining from the application of the Lorenz curve. As Dixon et al. (1987) point out, certain imperfections of indicators should not lead to their automatic rejection. Therefore, it suffices to keep these imperfections in mind when working with the Lorenz curve. Moreover, the curve provides additional information about the distribution of income among households that is not apparent from any other indicators, which determine only numeric values of the income distribution in the society.

The original Gini coefficient can be quantified in several different ways. Lerman and Yitzhaki (1989) define it as a covariance of the income function. The calculations can, in addition, be based on the median difference of income values. We may also mention the well-known Brown formula (Brown, 1994). There are some creatively formed mutations of the Gini index (see sub-chapter two), like the well-known Brown formula (Brown, 1994), which stress the significance of the Gini coefficient. In this paper, we will limit ourselves to the variation of the Gini coefficient G , which is based on the application of basic mathematical operations (Foldvary, 2006). This variation is an optimal form of the logical solution of simple models of income distribution in an economy, where n is the number of households and I is the households' income:

$$G = 1 + \frac{1}{n} - \frac{2}{n^2 \cdot \bar{I}} \cdot (I_1 + 2I_2 + \dots + nI_n) \quad (1)$$

For the sake of simplicity, we will use the standard form of the Lorenz curve, with the x -axis depicting the cumulative proportion of households and the y -axis, depicting cumulative income in ascending order, with a maximum value of “one”, or 100%. Consequently, the average income \bar{I} equals the proportion of the number of households n :

$$\bar{I} = \frac{1}{n} \quad (2)$$

For the examination of the Gini coefficient, we will use a simple model with two households. One household has a 100% of the income and the other household has no income. This is an absolutely unequal distribution of incomes in an economy and the value of the Gini coefficient should be “one”. In this model it follows that the total proportion of household incomes equals 1:

$$I_1 + 2I_2 = 1 \quad (3)$$

By necessity, the sum of the income of all households equals 1:

$$I_1 + 2I_2 + \dots + nI_n = 1 \quad (4)$$

After the modification of the basic formula, while formulas (3) a (4) apply, we can get the Gini coefficient in following form:

$$G = 1 - \frac{1}{n} \quad (5)$$

After substituting the number of households with the value “two”, the resulting value of the Gini coefficient is surprisingly 0.5, and not the expected “one”. Let us add other households to the model and leave the situation at the most unequal level; three households of which one receives all the incomes two have no income at all. The result of the Gini coefficient, in this case, is 0.66 (calculated as $1 - \frac{1}{n}$, that is $1 - \frac{1}{3}$). If we add another household with zero income, the coefficient value will increase to 0.75. With the growing number of households, the value of the Gini coefficient approximates the expected value of 1. If we have an unequal distribution of income, the Gini coefficient may be expressed as a function of the number of households $G(n)$:

$$G(n) = 1 - \frac{1}{n} \quad (6)$$

If, with a growing number of households, the value of the Gini coefficient approximates the value of “one”, this relationship can be expressed with the help of a $\lim G(n)$ and subsequently be modified:

$$\begin{aligned} \lim_{n \rightarrow \infty} G(n) &= \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right) \\ &= 1 - \lim_{n \rightarrow \infty} \frac{1}{n} = 1 \end{aligned} \quad (7)$$

If the number of households reaches infinity, the index value will equal the expected value of “one”. Therefore, the following relationship applies:

$$\text{if } n \rightarrow \infty, \text{ then } \frac{1}{n} \rightarrow 0. \quad (8)$$

The deviation from the Gini coefficient value can be quantified for this model with absolute inequality of distribution of incomes in an economy. If equation (5) applies, the deviation from the Gini coefficient D_G reaches the following level:

$$D_G = 1 - G = \frac{1}{n} \quad (9)$$

Although the identified deviation applies, it is relatively trivial and divorced from reality for two reasons. Firstly, the small number of households in our example does not represent any real economy. And secondly, absolute inequality, although theoretically plausible, is highly unlikely in reality. The deviation from the expected value decreases with the growing number of households, and the deviation may be considered negligible with a sufficient number of households. Unfortunately, in practice we must use the available data, which often classify households in percentiles, or, more commonly, deciles. Consequently, despite the originally broad sample, the sample yields only ten datapoints. In practice, the Gini coefficient deviation may reach the value of 0.1 if we have a sample of ten datapoints and the hypothetical situation of an absolutely unequal distribution of income in the economy in question, may arise. This is a significant problem and the question thus arises; what is the dependence of this deviation on the inequality of income distribution in a given economy? A completely

equal situation for two households, or any number of households, requires the coefficient result to be “zero”. If the basic formula (1) is applied, we find that the result always equals “zero”. Consequently, it holds true that, if household income is distributed absolutely equally, the deviation from the expected value is “zero” regardless of the number of households.

On this basis we can assume that any deviation from the expected Gini coefficient value will indicate inequality in the income distribution in a given economy. Let us introduce a model with a certain level, although not absolute, inequality, where the concurrent validity of the conditions is represented by symbol \wedge .

$$\begin{aligned} I_1, I_2 \neq 0 &\wedge I_1, I_2 \neq 1 \\ &\wedge I_1 + I_2 = 1 \\ &\wedge I_1 > I_2 \end{aligned} \quad (10)$$

Let us modify formula (1) for the purposes of this model:

$$G = 1 + \frac{1}{2} - (I_1 + 2I_2) \quad (11)$$

This equation may be further modified, if the conditions under (10) apply, i.e. the sum of incomes of the two households in the model equals 1:

$$G = 1 + \frac{1}{2} - (1 + I_2) \implies G = \frac{1}{2} - I_2 \quad (12)$$

At this point we should focus on the identification of the deviation from the Gini coefficient value. Unfortunately, we are not able to define the expected Gini coefficient value in a model with an unequal distribution of incomes. Consequently, we are unable to precisely quantify the deviation from the expected value. What we do know is that the deviation lies between “zero” and 0.5. The maximum deviation $\max D_G$ is therefore 0.5. We also know that an increase in inequality raises the maximum deviation. The value of the deviation in this situation may be quantified in several manners. We will choose a simple method, which consists of the quantification of deviations from the income average, which is 0.5:

$$D_G = \max D_G \cdot (|I_1 - 0.5| + |I_2 - 0.5|) \quad (13)$$

After inserting the income distribution 1/0, the deviation is at its maximum: 0.5. In case the income distribution is, for instance, 0.7/0.3, the deviation is only 0.2. If the equality of the income distribution between two households is further increased to 0.6/0.4, the deviation will be 0.1 and a distribution of 0.5/0.5 yields a deviation of 0.

Finally, the model must be generalized to include any number of households and their unequally distributed income. The average value is a quotient of the value “one” and the number of households n . The deviation from the Gini coefficient depends on the average deviation from the average value of household income:

$$D_G = \frac{|I_1 - \frac{1}{n}| + |I_2 - \frac{1}{n}| + \dots + |I_n - \frac{1}{n}|}{n} \quad (14)$$

Now we can make our final modification to the Gini coefficient G^* , based on formulas (1) and (14):

$$G^* = 1 + \frac{1}{n} - \frac{2}{n^2 \cdot \bar{I}} \cdot (I_1 + 2I_2 + \dots + nI_n) + D_G \quad (15)$$

This equation must be broken into partial items and subsequently modified to get the final form of the Gini coefficient which allows for a deviation from its expected value.

$$G^* = 1 + \frac{1}{n} - \frac{2}{n} \cdot (I_2 + 2I_3 + \dots + (n-1)I_n) + \frac{|I_1 - \frac{1}{n}| + |I_2 - \frac{1}{n}| + \dots + |I_n - \frac{1}{n}|}{n} \quad (16)$$

The final resulting value of the Gini coefficient can be further modified through standard mathematical operations. However, this will not lead to an apparent simplification of the coefficient allowing for a deviation from the expected value. Formula (16) is the resulting form of the modified Gini coefficient G^* , which allows for imperfections and inconsistencies in initial datasets.

It should, however, be noted that, after the adjustment of the original calculation of the Gini coefficient, values still range from 0 to 1. This means that the extreme, i.e. maximum values and their interpretation have not been affected by the adjustment.

5 DISCUSSION AND CONCLUSIONS

A number of methods may be used for the measurement of the distributional effects of taxes and of distribution of wealth in an economy. There are many entropy indexes and other indicators unrelated to the Lorenz curve and Gini coefficient, for instance the Theil index, Atkinson index etc. The Lorenz curve, Gini coefficient and indexes which are based on them, are widely used for the measurement of distributional effects of taxes and of the distribution of wealth in a society. These indexes have indisputable advantages, but also certain imperfections. Some of them have been the subject of previous papers, e.g. Bracewell-Milnes (1979). In addition, the differences of the Suits index (Suits, 1977) and the Kakwani index (Kakwani, 1979) are commonly recognized.

The Gini coefficient, Lorenz curve and many other indicators (Musgrave and Thin index, Kakwani index, Suits index, Reynolds-Smolensky index, Hoover index) also suffer from certain shortcomings, which result from incomplete input data. This includes the problem of a limited number of households entering the research, while the incomes in a society are unequally distributed. These imperfections are can be resolved in the case of the Gini coefficient, whereas the Lorenz curve must be interpreted with a certain distortion in mind.

We have ascertained that the degree of the Gini coefficient distortion grows with a decreasing number of households and increasing inequality in a society. The maximum degree of the Gini coefficient distortion, 0.5, was iden-

tified in an extreme model of two households, in which income was distributed absolutely unequally. Conversely, a model with an infinite number of households as well as a model with absolute equality shows no distortion and reached the value of 0. Although this situation is not feasible in practice (as well as the opposite extreme), it may be approximated to a certain extent. It may be concluded from the above that the value of the Gini coefficient will be systematically overestimated in cases with a limited number of households and a growing inequality of income. On the other hand, it must be said that the change in the Gini coefficient after tax in comparison with the situation before tax or the situation with the original tax is substantial. The deviations of the two compared Gini coefficients are compensated for, in the indicator of the Gini coefficient change. Therefore, the distortion of the difference of the Gini coefficients, which is caused by the input data, is reduced. However, we cannot completely eliminate the distortion, which would only be possible in case the distributional effect of taxation on the household income distribution is zero.

The distortions in the Gini coefficient may be mitigated by modifying the standard formula (1) used for its calculation. The formula for

calculating the Gini coefficient (16) involves, following Foldvary (2006), the sum of the standard formula (1) and of the deviation from the Gini coefficient in formula (14). This is one way to eliminate some of the imperfections, which result from secondary data that include only data on certain quantiles of the distribution of household income in the economy, when the real distribution of incomes among individual households is unequal. The application of this finding may not be restricted to the examination of the distributional effects of taxation, and could be useful in other domains, as well. The original range of the Gini coefficient has not been changed through the application of the mentioned procedure.

The validity of the proposed procedure is supported by its logic rationale, and its impact will be quantified in the future work of the author. We assume that the impact on realistic measurements will not be significant and that the proposed modification of the formula will predominantly contribute to formal correctness in cases for which the input data are not perfect. The only significant impact of our reformulation of the Gini coefficient could occur in the case of extremely unequal income distribution and a very low number of households or taxpayers, a plausible, but highly unlikely scenario.

6 REFERENCES

- ATKINSON, A. B. 1970. On the Measurement of Inequality. *Journal of Economic Theory*, 2 (3), 244–263.
- BESLEY, T. J. and ROSEN, H. S. 1999. Sales Taxes and Prices: An Empirical Analysis. *National Tax Journal*, 52 (2), 157–178.
- BLACKORBY, C. and DONALDSON, D. 1980. Ethical Indices for the Measurement of Poverty. *Econometrica*, 48 (4), 1053–1060.
- BOETERS, S., BÖHRINGER, C., BÜTTNER, T. and KRAUS, M. 2010. Economic Effects of VAT Reforms in Germany. *Applied Economics*, 42 (17), 2165–2182.
- BRACEWELL-MILNES, B. 1979. Measurement of Tax Progressivity: A Comment. *The Economic Journal*, 89 (355), 648–651.
- BROWN, M. C. 1994. Using Gini-Style Indices to Evaluate the Spatial Patterns of Health Practitioners; Theoretical Considerations and an Application Based on the Alberta Data. *Social Science & Medicine*, 38 (9), 1243–1256.
- CASPERSEN, E. and METCALF, G. E. 1994. Is a Value Added Tax Regressive? Annual Versus Lifetime Incidence Measures. *National Tax Journal*, 47 (4), 731–746.
- DAVID, P. 2012. Effects of Increased Value Added Tax Burden Levied on Agricultural Products and Foodstuffs in the Czech Republic and Slovakia. *Journal of Economics*, 60 (5), 509–527.
- DAVID, P. 2017. A Piece to the Puzzle of Lorenz Curve and Gini Coefficient – Instruments for Evaluation of Fairness. In *Enterprise and Competitive Environment: Conference Proceedings*. Brno: Mendelova univerzita v Brně, pp. 216–223.

- DIXON, P. M., WEINER, J., MITCHELL-OLDS, T. and WOODLEY, R. 1987. Bootstrapping the Gini Coefficient of Inequality. *Ecology*, 68 (5), 1548–1551.
- FELLMAN, J. 1976. The Effect of Transformations on Lorenz Curves. *Econometrica*, 44 (4), 823–824.
- FIDAN, H. 2016. Measurement of the Intersectoral Digital Divide with the Gini Coefficients: Case Study Turkey and Lithuania. *Engineering Economics*, 27 (4), 439–451.
- FOLDVARY, F. E. 2006. The Measurement of Inequality, Concentration and Diversification. *The Indian Economic Journal*, 54 (3), 179–188.
- FULLERTON, D. and ROGERS, D. L. 1991. *Lifetime vs. Annual Perspectives on Tax Incidence*. NBER Working Paper No. 3750 [online]. Available at: <https://doi.org/10.3386/w3750>. [Accessed 2018, February 17].
- FULLERTON, D. and ROGERS, D. L. 1994. *Distributional Effects on a Lifetime Basis*. NBER Working Paper No. 4862 [online]. Available at: <https://doi.org/10.3386/w4862>. [Accessed 2018, December 18].
- GINI, C. 1912. *Variabilità e mutabilità; contributo allo studio delle distribuzioni e delle relazioni statistiche*. Bologna: Tipografia di Paolo Cuppini.
- HOFFMANN, R. 2012. How to Measure the Progressivity of an Income Component. *Applied Economics Letters*, 20 (4), 328–331.
- HOOVER, E. M. 1936. The Measurement of Industrial Localization. *The Review of Economics and Statistics*, 18 (4), 162–171.
- KAKINAKA, M. and PEREIRA, R. M. 2006. *A New Measurement of Tax Progressivity*. GSIR Working Papers. Economic Development & Policy Series EDP067 [online]. Available at: http://nirr.lib.niigatau.ac.jp/bitstream/10623/31266/1/EMS_2006_09.pdf. [Accessed 2018, November 12].
- KAKWANI, N. C. 1977. Measurement of Tax Progressivity: An International Comparison. *Economic Journal*, 87 (345), 71–80.
- KAKWANI, N. C. 1979. Measurement of Tax Progressivity: A Reply. *Economic Journal*, 89 (355), 653–657.
- KESSELMAN, J. R. and CHEUNG, R. 2004. Tax Incidence, Progressivity, and Inequality in Canada. *Canadian Tax Journal*, 52 (3), 709–789.
- KIEFER, D. W. 1984. Distributional Tax Progressivity Indexes. *National Tax Journal*, 37 (4), 497–513.
- LERMAN, R. I. and YITZHAKI, S. 1989. Improving the Accuracy of Estimates of Gini Coefficients. *Journal of Econometrics*, 42 (1), 43–47.
- LORENZ, M. O. 1905. Methods of Measuring the Concentration of Wealth. *Journal of the American Statistical Association*, 9 (70), 209–219.
- METCALF, G. E. 1994. *Lifecycle vs. Annual Perspectives on the Incidence of a Value Added Tax*. NBER Working Paper No. 4619 [online]. Available at: <https://doi.org/10.3386/w4619>. [Accessed 2019, January 8].
- METCALF, G. E. and FULLERTON, D. 2002. *The Distribution of Tax Burdens: An Introduction*. NBER Working Paper No. 8978 [online]. Available at: <https://doi.org/10.3386/w8978>. [Accessed 2019, January 8].
- MUSGRAVE, R. A. and THIN, T. 1948. Income Tax Progression, 1929–48. *Journal of Political Economy*, 56 (6), 498–514.
- POTERBA, J. M. 1989. *Lifetime Incidence and the Distributional Burden of Excise Taxes*. NBER Working Paper No. 2833 [online]. Available at: <https://doi.org/10.3386/w2833>. [Accessed 2018, December 11].
- REYNOLDS, M. O. and SMOLENSKY, E. 1977. *Public Expenditures, Taxes, and the Distribution of Income: the United States, 1950, 1961, 1970*. New York: Academic Press.
- SEN, A. 1976. Poverty: An Ordinal Approach to Measurement. *Econometrica*, 44 (2), 219–231.
- SUITS, D. B. 1977. Measurement of Tax Progressivity. *American Economic Review*, 67 (4), 747–752.
- ŠIROKÝ, J. and MAKOVÁ, K. 2009. Independence between the Efficient Tax Rate and Tax Progressiveness in the Czech Republic During 1993–2007. *Journal of Economics*, 57 (7), 653–666.
- THEIL, H. 1967. *Economics and Information Theory*. Amsterdam: North Holland Publishing.
- ZANDVAKILI, S. 1991. *Redistribution Through Taxation: An International Comparison*. The Jerome Levy Economics Institute Working Paper No. 46. [online]. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=176788. [Accessed 2019, March 18].

AUTHOR'S ADDRESS

Petr David, Department of Accounting and Taxes, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: david@mendelu.cz