Journal of Cleaner Production 199 (2018) 668–686

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

The Gini coefficient structure and its application for the evaluation of regional balance development in China

Hui Shu [a, b, *], Pingping Xiong [c, b]

[a] School of Science, Nanjing University of Posts and Telecommunications, Nanjing, 210023, China
[b] China Manufacturing Development Research Institute, Nanjing University of Information Science and Technology, Nanjing, 210044, China
[c] School of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing, 210044, China

ARTICLE INFO

Article history:
Received 2 March 2018
Received in revised form
23 July 2018
Accepted 23 July 2018
Available online 30 July 2018

Keywords:
Multigroup overall Gini coefficient
Regional balanced development
Economy and ecology
Spatial heterogeneity

ABSTRACT

The evaluation of balanced regional development should consider not only the economic and ecological benefits but also regional differences. A nongrouped Gini index is constructed based on a continuous distribution function to evaluate the regionally balanced development of the economy and the environment, and the applicability of the method is analyzed given the data characteristics. Combining the difference in the regional economic base and ecological carrying capacity, a multigroup overall Gini index (MGO-Gini) with regional spatial heterogeneity is further developed, along with its decomposition structure. This work evaluates and analyses the regional balance in China's development with regard to regional industrial output, energy consumption and pollution emissions from 2000 to 2015 using both the nongrouped and the MGO-Gini. The results show that ①the Gini index constructed in this paper can be well adapted to fat-tailed data, which is a typical feature of China's regional economic and ecological benefits data. ②The regionally balanced development of China does not obviously promote industrial output, whereas the regional balance of energy consumption shows the same upward trend when evaluated with both the nongrouped and the MGO-Gini from 2000 to 2015. In recent years, after a period of improvement, the regional balance of pollution emissions has experienced a downward trend. There is disparity in "cleaner production efficiency" between regions of China as well as an upward trend in this disparity in recent years. ③China's interregional balance has been enhanced in terms of energy consumption and pollution emissions considering regional spatial heterogeneity, whereas the balance of economic output has been further reduced. ④The balance in regional energy consumption plays a supporting role in balanced development among regions. The imbalance within developed areas is the main source of regional economic output and energy consumption imbalances, and the imbalance in developing areas is the main source of the regional pollution imbalance in China.

1. Introduction

A prominent economic and ecological spatial imbalance exists between the various regions of China. GDP in the eight provinces and municipalities located in the eastern region of China accounted for 48% of the total for the entire country during the years 2000–2015, while the energy consumption of that region in the same period (measured by standard coal equivalent) accounted for only 8% of the total energy consumption. The proportion of industrial pollution as SO_2 emissions in the eight eastern provinces and municipalities also decreased from 30% to 23% between 2000 and 2015. The overall imbalance between regional economic output and ecological efficiency indicators creates difficulties and uncertainty for policy implementation in China. Therefore, an investigation into the features of the regional imbalances found in economic and ecological indices is critical for the pursuit of better decision-making.

The great differences between the economic and ecological benefits achieved from interregional development are influenced by many factors. Compared to the central and western regions of China, the eastern coastal regions have collected a large proportion of China's labor force. In terms of industrial technology level, the eastern coastal regions have an absolute advantage and started enjoying various preferential policies at an earlier date. The internal

* Corresponding author. School of Science, Nanjing University of Posts and Telecommunications, Nanjing, 210023, China.
E-mail address: shuh@njupt.edu.cn (H. Shu).

https://doi.org/10.1016/j.jclepro.2018.07.224
0959-6526/

and external conditions of regional production activities offer great advantages. Therefore, if we only evaluate the balance in regional development considering the economic and ecological benefits of production activities, we cannot reflect the differences in development across different regions.

Regional industrial production activities should be carried out within the scope of regional resources and the environment. The theory of "doughnut economics", which explores the connotation of a green economy, states that ecological security and social demand are the two principles of human economic activities (Lonergan, 2000). To assess the regional balance in development levels, it is necessary to evaluate the economic development base of each region, the carrying capacity of resources and the environment, and so on; however, different regions are characterized by great differences in the distribution of industrial outputs, resource consumption, and environmental pollution data. The diversification and differentiation of data types in regional balance evaluation, and the diversification of indicators, all indicate the need for higher requirements for evaluation methods.

2. Literature review

The Gini coefficient is an index tool used for the quantitative determination of balance based on the Lorenz curve. The traditional Gini coefficient has been widely used to evaluate the regionally balanced development of an economy and its ecology (Arne et al., 2005; GrovesKirkby et al., 2009; Tao et al., 2010; Fei et al., 2011). The Gini coefficient calculation method in these studies is the same as that used for the welfare balance. However, there are major differences in the distribution patterns between residential income data and regional economic output, pollution emissions and energy consumption data. The sample size of the population and regional area are at different quantitative levels. The differences in these characteristics play an important role in the calculation method used to obtain the Gini coefficient. China's regional development has long been characterized by excessive spatial agglomeration, and there is highly unequal distribution of GDP per unit of air pollutants and CO_2 emissions between eastern and western regions (Hanwei et al., 2016), leading to an elongation of the heavy tails of different regional industrial output and pollution emission data (Michel and Abdoul, 2016). How will the difference between data distribution patterns affect the evaluation results of the Gini coefficient? Is the traditional Gini coefficient method applicable to the regional output, emissions and energy consumption data of China? A discussion of these questions is lacking.

In a nonparametric approach, one computes the index from the empirical distribution of the available data, which can lead to unreliable inference in the presence of heavy tails (Bahadur and Savage, 1956). Andrea et al. (2018) study the problems related to the estimation of the Gini index in the presence of a "fat tail", which is another name for the "heavy-tailed" data, and they show that in such a case, the Gini coefficient cannot be reliably estimated using conventional nonparametric methods because a downward bias emerges under a fat tail. What is impressive is that most of the Gini coefficient algorithms in regional balance or income balance assessment studies use nonparametric methods (Jin et al., 2016; Jiandong et al., 2016a,b; Baiqiong et al., 2017; Andrew et al., 2017; Si'ao et al., 2017; Jiandong et al., 2017a). In the face of data with significant fat-tailed characteristics, the Gini coefficient calculation method needs to be improved or even replaced.

Another type of Gini coefficient calculation method is based on the parametric equation of the Lorenz curve and the continuous distribution function of the index (Kendall and Alan, 1977; Dorfman, 1979; Yizhaki, 1982; Lambert, 1989). The statistical distribution of the index needs to be provided when calculating the Gini coefficient using the parametric method, and the derivation and analysis are more difficult. Yonghong (2006) derives a new formula for calculating the Gini coefficient based on the parametric equation of the Lorenz curve using the income distribution function. However, the application of this type of parametric Gini method to regional equilibrium evaluation is rare, and the calculation formula based on the distribution of regional economic output data or pollution emission data has not been studied.

In addition, when the differences in regional economic and ecological conditions are taken into account in the balanced evaluation of regional development, the overall Gini coefficient is needed. The overall Gini coefficient has been widely studied in the research field of welfare balance (Bourguignon, 1979; Shorrocks, 1980, 1984; Jacques, 1989; Lambert and Richard, 1993; Dagum, 1997; Shujie, 1999; Eliazar and Cohen, 2014). The multigroup overall income Gini coefficient can measure the imbalance of mixed data types. Most of the existing overall Gini coefficients are based on the nonparametric method, and studies using the overall Gini method based on the continuous distribution of regional economic and ecological benefits indices are still scarce.

An in-depth research field for the overall Gini coefficient addresses its decomposition structure. The decomposition of the overall Gini coefficient can effectively reveal the source of the welfare imbalance. The existing decomposition methods for the overall Gini coefficient almost all come from the research field of the income Gini coefficient (Lambert and Richard, 1993; Dagum, 1997). The decomposition method with a nonparametric overall Gini coefficient is most common in the field of welfare imbalance source research (Modalsli, 2018). Most studies on the unbalanced structure of regional economic and ecological benefits directly apply the decomposition technology of the income Gini (Emilio and Juan, 2013; Lorena et al., 2016; Hanwei et al., 2016; Jiandong et al., 2016a,b; Caizhi et al., 2016; Si'ao et al., 2017; Jiandong et al., 2017a,b; PongLung et al., 2017). However, almost none of these applied studies discriminate the type of sample data.

The inadaptability of the nonparametric Gini method has attracted the attention of some scholars. Michel and Abdoul (2016) develop a decomposable inequality index that can be implemented in the framework of a finite mixture of lognormal distributions so that overall inequality can be decomposed into within-subgroup and between-subgroup components. Michel and Abdoul (2016) find that the lognormal distribution is suitable for describing the income distribution in theoretical terms. Ebert (2010) characterizes the class of weakly decomposable (aggregable) inequality measures that can be decomposed into the sum of the usual within-group term and a between-group term. Yen-Sheng (2015) conduct a computational experiment to study the difference between global and local inequality assessed by the Gini index. Druckman and Jackson (2008) describe an indicator of inequality in resource use called the AR-Gini. The AR-Gini is an area-based measure of resource inequality that estimates inequalities between neighborhoods with regard to the consumption of specific consumer goods. In addition, the expansion of the dimensions of the Gini index is also an important research field, for example, the multicriteria or multidimensional Gini index (Xiaocun and Fenglai, 2017; Motahareh et al., 2018).

All of these studies focus on the applicability of the Gini coefficient and further extend its application scope. This article also follows this path. First, we try to derive the Gini index formula based on a continuous distribution and assess its suitability for regional economic output data and pollution emission data types. Second, we try to derive a calculation method for the overall Gini coefficient with multigroup mixed data. Third, we will derive the decomposition structure of the overall Gini coefficient.

In the third section, we analyze the differences between the