Using the Gini index to measure the inequality in infrastructure services provided within an urban region

Abhik Kumar Das* and Jai Asundi

Center for Study of Science, Technology and Policy,
Dr. Raja Ramanna Complex,
Raj Bhavan Circle, High Grounds,
Bangalore – 560001, India
E-mail: abhik@cstep.in
E-mail: asundi@cstep.in
*Corresponding author

Abstract: In this article, we suggest the use of the Gini index to measure the inequality in infrastructure services provided within an urban region. We develop a functional form for the Lorenz curve, the basis of measurement of the Gini index. The sparse nature of data available to measure the distribution of infrastructure services within an urban region makes estimating the Lorenz curve challenging. However, the proposed functional form for the Lorenz curve requires a sparse amount of data (only 25th and 50th percentile) and it is simple to estimate and compute the coefficients of the equation as well as the Gini index. Our methodology is shown to be relatively accurate in estimating the Gini index for national level income data compared to other formulations found in the literature. In this paper, with the help of collected data, we demonstrate the use of the Gini index to characterise urban regions in terms of their accessibility of public infrastructure services.

Keywords: Gini index; inequality; infrastructure service.

Reference to this paper should be made as follows: Das, A.K. and Asundi, J. (2012) 'Using the Gini index to measure the inequality in infrastructure services provided within an urban region', *Int. J. Critical Infrastructures*, Vol. 8, Nos. 2/3, pp.178–186.

Biographical notes: Abhik Kumar Das is with Centre for Study of Science, Technology and Policy, Bangalore, India.

Jai Asundi is with Centre for Study of Science, Technology and Policy, Bangalore, India.

This paper is a revised and expanded version of a paper entitled 'Using the Gini index to measure the inequality in infrastructure services provided within an urban region' presented at The 4th Annual International Conference on Next Generation Infrastructures, Virginia Beach, Virginia, 16–18 November 2011.

1 Introduction

Infrastructure services within any urban region are unequally distributed (Cliff and Ord, 1981; Davies, 1968; Hodge and Gatrell, 1976; Talen and Anselin, 1998). Given the paucity of data available, it is very difficult to quantify the variance or even summarily characterise the inequality in access to the infrastructure services. We propose the use of the Gini index to characterise an urban region in terms of its accessibility to public infrastructure services.

The Gini index is derived from the Lorenz curve. A Lorenz curve is traditionally used as a method to represent the distribution of income and wealth in a population. It is created by plotting cumulative proportion of income and the cumulative proportion of income-receiving unit. If $\pi(x)$ is the proportion of units that receive income up to x and $\eta(x)$ is the proportion of the total income received by the same unit, the Lorenz curve can be represented as $\eta = \eta(\pi)$. The necessary and sufficient condition for the function $\eta(\pi)$ to describe the Lorenz curve is:

a $\eta(0) = 0$

b $\eta(1) = 1$

c $0 \leq \eta \leq \pi \leq 1$

d $\dfrac{d\eta}{d\pi} > 0$ and $\dfrac{d^2\eta}{d\pi^2} > 0.$

Similar to income distribution, a Lorenz curve can be constructed to characterise the inequality of the distribution of public infrastructure services within an urban region. As per our example (shown later) the Lorenz curve represents the percentile of population with access to the percentile of infrastructure services.

The Lorenz curve may be estimated from data grouped by connecting the observed pair points (π_i, η_i), $i = 1, 2, \ldots$ using some interpolation technique or it may be presumed to follow a particular parametric functional form and fit to tabulated data (Gastwirth and Glauberman, 1976). The most popular functional Lorenz curves are the forms proposed by Kakwani and Podder (1976), Rasche (1980), Gupta (1984), Rao and Tam (1987), Ortega et al. (1991), and Chotikapanich (1993). These alternative functional forms satisfy the above necessary and sufficient conditions and depend on the available data and computational complexity. Since any functional form uses error minimisation, the extraction of parameters becomes inconvenient when the functional form becomes implicit like Kakwani and Podder (1976). Sometimes the calculation of Gini index becomes complex, as some of the functions require the evaluation of the beta function (Kakwani and Podder, 1976; Ortega et al., 1991) or hyper-geometric function (Rao and Tam, 1987). When the available data is sparse in nature, the error-minimisation techniques may underestimate or overestimate the Gini index.

To address these issues, we propose a simple explicit functional form of the Lorenz curve which is computationally convenient to calculate the Gini index. The parameters used in the proposed functional form can be calculated easily and if the available data is low (e.g., only lowest 20% and highest 20% share of access), the functional form satisfactorily calculates the Gini index and predicts the Lorenz curve without using interpolation or error minimisation techniques for national level income data. The same method is used to measure the inequality of infrastructure services within an urban region. The different popular analytical forms of Lorenz curve are represented in Section 2. Section 3 shows the proposed explicit functional form and validates the proposed model using the national income data to compare to other functional forms. To use the proposed method in measuring the inequality of infrastructure services, we have taken Bangalore, India as a study area and measured the inequality of the service accessibility for different zones of Bangalore. Due to the presence of sparse data, the proposed method to calculate Gini index is used, which gives satisfactory results as shown in Section 4. Section 5 concludes the paper by highlighting the contribution.

2 Popular functional forms

Kakwani and Podder (1976) proposed a functional form for the Lorenz curve as follows:

$$m = \gamma n^{\alpha}\left(\sqrt{2} - n\right)^{\beta} \tag{1a}$$

where α, β, $\gamma > 0$,

$$m = \frac{1}{\sqrt{2}}(\pi + \eta) \text{ and } n = \frac{1}{\sqrt{2}}(\pi - \eta) \tag{1b}$$

This functional form is implicit and requires the extraction of three unknown parameters α, β and γ using the method of ordinary least square error minimisation on the function in logarithmic form. The Gini index of this curve can be formulated as,

$$C = 2\gamma\left(\sqrt{2}\right)^{1+\alpha+\beta} B(1+\alpha,\ 1+\beta) \tag{1c}$$

where $B(1 + \alpha, 1 + \beta)$ is the beta function.

Rasche et al. (1980) proposed an alternate functional form for the Lorenz curve which is:

$$\eta(\pi) = \left[1 - (1-\pi)^{\alpha}\right]^{\frac{1}{\beta}} \tag{2a}$$

where $0 \le \alpha, \beta \le 1$.

This functional form is explicit using two unknown parameters which can be estimated using non-linear least square. The Gini index can be evaluated as:

$$C = 1 - \frac{2}{\alpha} B\left(\frac{1}{\alpha},\ 1 + \frac{1}{\beta}\right) \tag{2b}$$

where $B(1/\alpha, 1 + 1/\beta)$ is the beta function.

An alternate form was proposed by Gupta (1984) as:

$$\eta(\pi) = \pi\alpha^{\pi-1},\ a > 1 \tag{3a}$$

This is a single parametric form which uses linear least square on the logarithmic form of the function. The Gini index of this function can be represented as:

$$C = 1 - \frac{2}{\ln\alpha}\left(1 - \frac{1}{\ln\alpha} + \frac{1}{\alpha\ln\alpha}\right) \tag{3b}$$

Rao and Tam (1987) proposed a functional form which is a refinement of the earlier form using two unknown parameters as follows,

$$\eta(\pi) = \pi^{a} b^{\pi-1} \tag{4a}$$

which gives the Gini index as:

$$C = 1 - \frac{2e^{-\ln b}}{1+a} F(1+a,\ 2+a, \ln b) \tag{4b}$$

where $F(1 + a, 2 + a, \ln b)$ is the confluent hyper-geometric function.

Ortega et al. (1991) proposed the functional form as:

$$\eta(\pi) = \pi^{\alpha}\left[1-(1-\pi)\beta\right] \tag{5a}$$

where $\alpha \geq 0$ and $0 \leq \beta \leq 1$. This form can be viewed as refinement of function proposed by Rasche. This functional form uses two parameters which can be extracted using non-linear error minimisation. The Gini index is given by:

$$C = \frac{a-1}{a+1} 2B(1+a,\ 1+b) \tag{5b}$$

where $B(1 + a, 1 + b)$ is the beta function.

An alternate single parametric functional form was proposed by Chotikapanich (1993) using exponential expression as:

$$\eta(\pi) = \frac{e^{k\pi}-1}{e^{k}-1}, \text{ where } k > 0 \tag{6a}$$

The unknown parameter can be derived using ordinary least square minimisation and the Gini index can be represented as:

$$C = \frac{(k-2)e^{k} + (k+2)}{k\left(e^{k}-1\right)} \tag{6b}$$

3 Explicit functional approximation

The proposed functional form to approximate the Lorenz curve can be represented as:

$$\eta(\pi) = (1+m)\pi + m\pi^{n}, \tag{7}$$

where $0 \leq \pi \leq 1$, $0 \leq m$ 1 and $n > 1$.

Here it is very trivial to show that $\eta(0) = 0$ and $\eta(1) = 1$, which satisfies the first two necessary and sufficient conditions of Lorenz curve as discussed earlier.

Here,

$$\frac{d\eta}{d\pi} = (1-m) + mn\pi^{n-1} \tag{8}$$

and

$$\frac{d^2\eta}{d\pi^2} = mn(n-1)\pi^{n-2} \tag{9}$$

Since $0 \le m \le 1$ and $n > 1$, we can state that $\frac{d\eta}{d\pi} > 0$ and $\frac{d^2\eta}{d\pi^2} > 0$ for all values of π lying between 0 and 1, which satisfies the last necessary and sufficient condition. Hence, the slope of $\eta(\pi)$ is non-negative and monotonically increasing. Considering the condition of the slope, it is easy to state that $0 \le \eta \le 1$, since $0 \le \pi \le 1$. The proposed equation can be represented as $\eta(\pi) = \pi - m\pi\,(1 - \pi^{n-1})$, since $0 \le m \le 1$ and $n > 1$, $\eta(\pi) \le \pi$ for $0 \le \pi \le 1$. Hence, it is proved that the proposed functional form follows the all necessary and sufficient condition of Lorenz curve.

As shown in Figure 1, the nature of the proposed functional form depends on the values of m and n. When π is small and near to 0, which represents the lowest proportion share of income, the proposed form can be approximated as $\eta(\pi) \approx (1 - m)\pi$ and the 'linear factor' m captures the linear dependency of η with π. When π is high and near to 1, which shows the highest proportion share of income, the 'power factor' n captures the exponential effect of the Lorenz curve. An important application of the Lorenz curve is in the estimation of the Gini index, a popular measure of the income inequality. Using simple algebraic manipulation, the Gini index can be formulated as:

$$C = 1 - 2\int_0^1 \eta(\pi)\mathrm{d}\pi = m\frac{n-1}{n+1} \tag{10}$$

In the proposed model function, there are two unknown variables m, n to approximate the Lorenz curve. From the real data we can approximate these values using two different values of η at $\pi = \alpha, \beta$, where $0 < \alpha < \beta < 1$. Using simple algebraic manipulation we can state that:

$$m \approx 1 - \frac{\eta(\alpha)}{\alpha} \tag{11}$$

$$\eta \approx \frac{\log\left[\{\eta(\beta) - (1-m)\beta\}m^{-1}\right]}{\log \beta} \tag{12}$$

The choice of α, β depends on the availability of the data. It is realistic to consider those data, where the percentage share of income less than 20% and more than 20% are available, hence α =0.2 and β = 0.8, hence the equation generating m and n can be transformed into:

$$m \approx 1 - 5\eta(0.2) \tag{13}$$

$$n \approx -4.48\log\left[\{\eta(0.8) - 0.8(1-m)\}m^{-1}\right] \tag{14}$$

Using the above equations, the wide variety of Gini index is estimated and also compared with other equations which have explicit functional form and comparatively easy computation method for Gini index, as shown in Table 1. We observe that our proposed explicit functional form of the Lorenz curve and its respect Gini index value is comparable to previous estimation models.

Figure 1 This shows the proposed functional form of $\eta = \eta(\pi)$, (a) for different values of n and (b) for different values of m (see online version for colours)

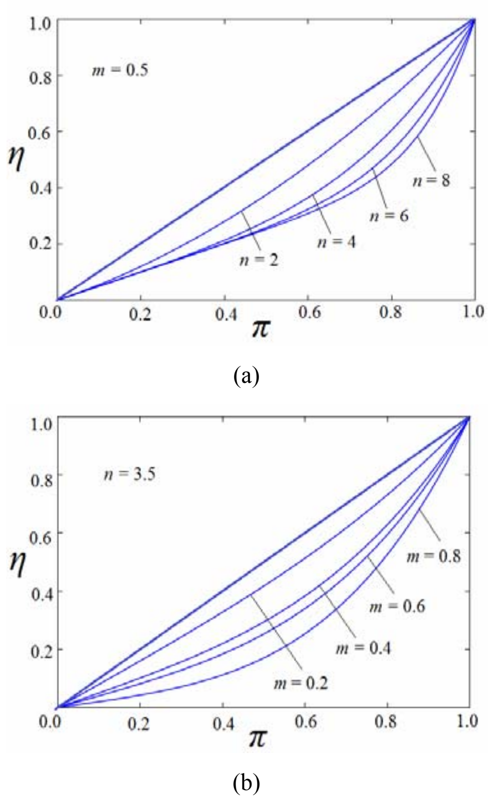

Table 1 Comparison of Gini index using different methodologies

Country	*Gini (reported)*	*η(0.2)*	*η(0.8)*	*m*	*n*	*Gini index*				
						Proposed	*Rasche*	*Gupta*	*Ortega*	*Chotikapanich*
Japan	*0.249*	0.106	0.643	0.47	3.42	*0.247*	0.248	0.266	0.246	0.272
Hungary	*0.300*	0.086	0.613	0.58	3.30	*0.299*	0.301	0.318	0.301	0.314
Poland	*0.349*	0.073	0.576	0.64	3.60	*0.349*	0.350	0.371	0.346	0.367
Trinidad and Tobago	*0.403*	0.055	0.541	0.74	3.6	*0.404*	0.399	0.423	0.399	0.416
Cameron	*0.446*	0.056	0.491	0.73	4.40	*0.445*	0.441	0.476	0.446	0.471
Zimbabwe	*0.501*	0.046	0.443	0.77	4.90	*0.500*	0.501	0.536	0.502	0.529
Brazil	*0.550*	0.030	0.413	0.86	4.70	*0.546*	0.544	0.544	0.550	0.563
Belize	*0.596*	0.021	0.371	0.89	5.10	*0.591*	0.597	0.611	0.601	0.608
Comoros	*0.643*	0.026	0.319	0.87	6.2	*0.630*	0.639	0.657	0.646	0.651
Namibia	*0.743*	0.006	0.217	0.92	8.00	*0.718*	0.748	-	0.751	0.738

Source: Data are from World Bank (2010)

4 Results

Public service accessibility is one of the key functionalities in any urban area. The proposed method to calculate the Gini index is used to measure the inequality of accessibility of public services in Bangalore, India (see Figure 2). Bangalore, India, which has been substantially affected by globalisation and rapid urbanisation over the last decade, is used as the study area of the proposed model. Due to sparse nature of available data, our proposed functional form and computation methodology is useful for the calculation of the Gini index. The study area is divided in eight zones as decided by the Directorate of Urban Land Transport, Government of Karnataka, India. The zones represent various administrative parts of the city and the peri-urban area. They are shown in Figure 2. The service accessibility data are collected as the percentage of population that has access to public services within stipulated time. It is assumed that the maximum reasonable access time is one hour. Further assuming that 15-min and 30-min are the 25% and 50% 'accessibility' measures, the Lorenz curve coefficients (*m* and *n* values) are computed as well as the Gini index values for each of the zones described in Figure 2. The results are shown in Table 2. The respective Lorenz curves are plotted in Figure 3.

Figure 2 Different zone of Bangalore, India as study area, showing the public transport (PT) coverage area, road network and bus stops (see online version for colours)

Table 2 Gini index of eight zones using proposed methodology

Zone	*Service accessibility (15-min)*	*Service accessibility (30 min)*	*m*	*n*	*Gini index*
1	0.16	0.26	0.36	1.85	0.1072
2	0.15	0.33	0.40	1.15	0.0283
3	0.12	0.25	0.52	2.00	0.1733
4	0.12	0.19	0.52	2.89	0.2529
5	0.15	0.24	0.40	2.15	0.1462
6	0.14	0.22	0.44	2.46	0.1856
7	0.18	0.30	0.28	1.22	0.0280
8	0.15	0.24	0.40	2.15	0.1462
Average	*0.15*	*0.25*	*0.40*	*2.00*	*0.1333*

Figure 3 Lorentz curve of eight zones of Bangalore, India (see online version for colours)

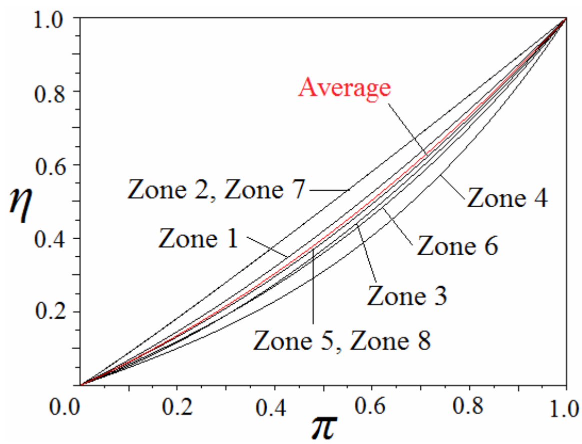

Table 3 Gini index of different example values

Service accessibility (15 min)	*Service accessibility (30 min)*	*m*	*n*	*Gini index*
0.1	0.2	0.60	2.58	0.2652
0.01	0.02	0.96	6.58	0.7069
0.001	0.002	0.99	9.96	0.8142

It is observed that the various zones described in Table 2 have differing Gini index values implying that some zones have better accessibility to public services. In this manner, public policy officials can target those zones where the Gini index is higher.

Table 3 provides a simple explanation for the interpretation of the results. If 10% of the population has access to public services within 15 minutes of their location and 20% of the population has access to public services within 30 minutes of their location, then we obtain a Gini index of 0.2652 for that locality. Similarly, if only 1% of the population has access within 15 minutes and 2% of the population has access within 30 minutes of their respective locations, then the Gini index is 0.8142.

Consistent with the traditional interpretation of inequality in distribution, a high Gini index implies a greater disparity (and effectively larger access times) for a majority of the population.

5 Conclusions

In this paper, we have proposed the use of the Gini index to characterise urban areas in terms of their accessibility. This measure can be used by urban planners to identify areas in cities that lack appropriate infrastructure services. We propose a functional approximation for the Lorenz curve and how it can be applied to this context. We validate the functional form by comparing it with existing models. We show that in spite of using sparse data the results are comparable to existing models. For this reason, our model is suitable for characterising public infrastructure services. Defining appropriate threshold Gini index values for public services will be a topic of study in our future work.

References

Chotikapanich, D. (1993) 'A comparison of alternative functional forms for the Lorenz curve', *Economics Letters*, Vol. 41, No. 2, pp.129–138.

Cliff, A. and Ord, J.K. (1981) *Spatial Processes: Models and Applications*, Pion, London.

Davies, B.P. (1968) *Social Needs and Resources in Local Services*, Michael Joseph, London.

Gastwirth, J.L. and Glauberman, M. (1976) 'On the interpolation of the and from grouped data', *Econometrica*, Vol. 44, No. 3, pp.479–483.

Gupta, M.R. (1984) 'Functional form for estimating the Lorenz curve', *Econometrica*, Vol. 52, No. 5, pp.1313–1314.

Hodge, D. and Gatrell, A. (1976) 'Spatial constraint and the location of urban public facilities', *Environment and Planning A*, Vol. 8, No. 2, pp.215–230.

Kakwani, N.C. and Podder, N. (1976) 'Efficient estimation of the s and the associated inequality measures from grouped observations', *Econometrica*, Vol. 44, No. 1, pp.137–148.

Ortega, P. et al. (1991) 'A new functional form for estimating Lorenz curves', *Review of Income and Wealth*, Vol. 37, No. 4, pp.447–452.

Rao, U.L.G. and Tam, A.Y. (1987) 'An empirical study of selection and estimation of alternative models of the Lorenz curve', *Journal of Applied Mathematics*, Vol. 14, No. 3, pp.275–280.

Rasche, R.H. et al. (1980) 'Functional forms for estimating the Lorenz curve', *Econometrica*, Vol. 48, No. 4, pp.1061–1062.

Talen, E. and Anselin, L. (1998) 'Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds', *Environment and Planning A*, Vol. 30, No. 4, pp.595–613.

World Bank (2010) *World Development Indicators*, available at http://www.data.worldbank.org/.