A new ordered family of Lorenz curves with an application to measuring income inequality and poverty in rural China

Zuxiang WANG [a], Russell SMYTH [b,*], Yew-Kwang NG [c]

[a] *Department of Economics, Wuhan University, Wuhan 430072, China*
[b] *Department of Economcs, Monash University, Caulfield East, Vic 3145, Australia*
[c] *Department of Economcs, Monash University, Clayton, Vic 3800, Australia*

ARTICLE INFO

ABSTRACT

The most common data source on income distribution in China is grouped data. When income data is in grouped form, some acceptable Lorenz model is needed to approximate the underlying Lorenz curve. This paper presents a new family of Lorenz curves and applies the main model in our proposed family of Lorenz curves to income data for rural China over the period 1980 to 2006. We find that the income share of the rural population at the low end of the income scale has been shrinking, income inequality in rural China has increased over time and that income inequality has impeded attempts to reduce poverty. However, the welfare of the rural population is still improving in terms of the generalized Lorenz dominance criterion.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

China has undergone monumental economic transition since 1979, which has resulted in a high rate of economic growth. The distribution of income has also gone through large-scale change. There have been several studies of income inequality in rural China (see eg. Benjamin, Brandt, & Giles, 2005; Griffen & Zhao, 1993; Hu, Wang, & Kang, 1997; Rozelle, 1996; Wan & Zhou, 2005; Yao, 1997). These studies point to a general trend of an increase in income inequality since the late 1970s. Rising rural income inequality threatens China's ability to maintain sustainable growth and potentially impinges on political and social stability (Wan & Zhou, 2005, p.107). The latter has been of particular concern to the Chinese government with income distribution a central platform of constructing a harmonious society as first enunciated by the Hu-Wen administration during the 2005 National People's Congress. Income inequality in China also has implications that extend beyond its national boundaries. As noted by Chotikapanich, Rao, and Tang (2007, pp. 127–128): "As China accounts for about a quarter of the world's population, changes in income and income inequality in China have important implications [for] global income inequality … This means that any advancement in the measurement of income inequality within China is not only important for understanding the economic development and well-being of people inside the 'Middle Kingdom', but also important in the global context".

Studies which have examined income inequality in China have used one of the two main sources of data on income distribution in China; namely, household survey data or grouped data. While rich in detail, the release of household survey data has been sporadic and restricted to a few provinces at a time. Thus, studies of income inequality in China that have used household survey data have been forced to focus on specific geographic locales and isolated years (see eg. Gustafsson & Li, 2002; Meng, 2004). In contrast to household surveys, grouped data is more readily available and has a wider coverage. However, since the income data is in grouped form, some acceptable Lorenz model is needed to approximate the underlying Lorenz curve. Several parametric Lorenz

---

* Corresponding author.
  E-mail address: Russell.Smyth@BusEco.monash.edu.au (R. Smyth).

models to fit grouped data have been developed (see Basmann, Hayes, Slottje, & Johnson, 1990; Cheong, 2002; Chotikapanich, 1993; Gupta, 1984; Holm, 1993; Kakwani, 1980; Kakwani & Podder, 1973, 1976; Ogwang & Rao, 1996, 2000; Ortega, Martin, Fernandez, Ladoux, & Garcia, 1991; Rao & Tam, 1987; Rasche, Gaffney, Koo, & Obst, 1980; Rossi, 1985; Ryn & Slottje, 1996; Sarabia, 1997; Sarabia, Castillo, Pascual, & Sarabia, 2005; Sarabia, Castillo, & Slottje, 1999, 2001; Schader & Schmid, 1994). A limitation of some models that have been used to approximate the Lorenz curve is that they do not in fact satisfy the definition of the Lorenz curve (see Kakwani, 1980; Kakwani & Podder, 1976). Another shortcoming of some Lorenz models, as noted by Cheong (2002), is that they tend to underestimate or overestimate the Gini index by generating values outside the Gastwirth (1972) lower or upper bounds for the Gini index. Recent research has concentrated on finding models which are satisfactory both in theory and in practice. These models need to be able to satisfy the definition of the Lorenz curve as well as exhibit good performance for a variety of data sources (see, eg, Ogwang & Rao, 2000; Ryn & Slottje, 1996; Sarabia et al., 1999).

The objective of this paper is to develop a new Lorenz curve model, building on a basic model, first proposed by Sarabia et al. (1999), and apply this model to examine rural income inequality and poverty with grouped data in China. Sarabia et al. (1999) suggest a basic model of the form $\tilde{L}(p) = p^\alpha L(p)^\eta$. We first show that this model can be generalized and proceed to apply the generalized model to income data for rural China over the period 1980 to 2006. This allows us to examine the relationship between rural income inequality and poverty over a quarter century of market reforms. To this point: "One striking feature of the current discussion of inequality in China is the absence of well-documented facts about inequality and temporal changes in the structure of the income distribution" (Benjamin et al., 2005, p.769). This reflects the fact that while there are several studies of income inequality in China, with few exceptions most have provided only a snapshot without a time profile. Our results show that the income share of low income earners in rural China is getting smaller, income inequality in rural China has increased over time and that rural income inequality has impeded attempts to reduce poverty. However, in spite of worsening rural income inequality, we also find that the welfare of the rural population is still improving in terms of the generalized Lorenz dominance criterion.

The paper is set out as follows. Section 2 presents a new Lorenz curve model, building on a basic model, first proposed by Sarabia et al. (1999). Section 3 outlines the indices used to measure income inequality and poverty in rural China and generalized Lorenz dominance used to evaluate changes in social welfare over time. Section 4 applies the new Lorenz curve model to present new estimates of income inequality and poverty in rural China. The final section is the conclusion.

2. A new Lorenz curve model

A Lorenz curve is a function relation $L(p)$ which represents the income share earned by the bottom p percent of low income earners. Assume $L(p)$ is defined and continuous in the interval [0,1] with second derivative $L''(p)$. It is called a Lorenz curve if: $L(0) = 0, L(1) = 1$, $L'(0^+) \geq 0$ and $L''(p) \geq 0$ for all p [0,1].

Sarabia et al. (1999) give two theorems to describe the condition for the basic form of $\tilde{L}(p) = p^\alpha L(p)^\eta$ to be a Lorenz curve which can be briefly written as a single result:

Theorem 1. *If $L(p)$ is a Lorenz curve and $\eta \geq 1$, then*

$$\tilde{L}(p) = p^\alpha L(p)^\eta$$

is a Lorenz curve if one of the following conditions holds

   (1) $\alpha = 0$,
   (2) $\alpha \geq 1$,
   (3) $\alpha \in [0,1]$ *and* $L'''(p) \geq 0$ *for all* $p \in [0,1]$.

Thus, from a special parametric Lorenz model $L(p)$, three new models can be established. From the Lorenz curve associated with the classical Pareto distribution:

$$S_0(p) = 1 - (1-p)^\beta \tag{1}$$

Sarabia et al. (1999) introduce the generalized Pareto family of Lorenz curves which encompasses the model in Eq. (1) and the following three models:

$$S_1(p) = p^\alpha \left[1 - (1-p)^\beta \right], \tag{2}$$

$$S_2(p) = \left[1 - (1-p)^\beta \right]^\eta, \tag{3}$$

$$S_3(p) = p^\alpha \left[1 - (1-p)^\beta \right]^\eta. \tag{4}$$

S_1 and S_2 are well-known Lorenz models suggested by Ortega et al. (1991) and Rasche et al. (1980) respectively. S_3 is the main model of the family discussed further below. Sarabia et al. (1999) claim that, because $S_0(p)$ is a Lorenz curve with $S_0'''(p) \geq 0$, $S_3(p)$ is a Lorenz curve if Theorem 2 can be satisfied: