Mathematical Modelling of Lorenz curves

Johan Fellman
Hanken School of Economics, Helsinki, Finland,
fellman@hanken.fi

Consider a distribution function $F(x)$ and the corresponding frequency function $f(x)$ defined for non-negative variables $x$. Let the mean be $\mu = \int\limits_0^{x_{\max}} xf(x)dx$ and the Lorenz curve $L(p) = \frac{1}{\mu}\int\limits_0^{x_p} xf(x)dx$. The Lorenz curve ia a convex increasing curve satisfying the conditions $L(0) = 0$ and $L(1) = 1$. The Gini coefficient is $G = 1-2\int\limits_0^1 L(p)dp$, that is, the ratio of the area between the diagonal and the Lorenz curve and the area of the whole triangle under the diagonal. Primary income data yields the most exact estimates of the Gini coefficient.

Various attempts have been made to obtain accurate estimates of the Gini coefficient based on the Lorenz curves. The trapezium rule is simple, but yields a negative bias for the Gini coefficient. Simpson´s rule is better fitted to the Lorenz curve, but this rule demands an even number of subintervals of the same length. When one applies Simpson´s rule, one has to consider Lorenz curves with deciles. Lagrange polynomials of second degree can be considered as a generalisation of Simpson´s rule because they do not demand equidistant points. In addition, one can use Golden´s method (2008). No method is uniformly optimal, but the trapezium rule is almost always inferior and Simpson´s rule is superior. Golden´s method is usually of medium quality (Fellman, 2012a).

There has been a number of studies in which the scientists have built models for Lorenz curves. The step from Lorenz curve to distribution function is more difficult than the step from distribution function to Lorenz curve. There is a difference between advanced and simple Lorenz models. Advanced models with several parameters yield a better fit to data, but are difficult to connect to exact income distributions. Simple one-parameter models can more easily be associated with the corresponding income distribution, but when statistical analyses are performed the goodness of fit is often poor (Fellman, 2012b).

In this study, simple models are considered more in detail and the results are compared with results obtained by numerical methods without any assumptions concerning the Lorenz model (Fellman, 2012a).

References

Fellman, J. (2012a). Estimation of Gini coefficients using Lorenz curves. Journal of Statistical and Econometric Methods. 1(2):31-38.

Fellman, J. (2012b). Modelling Lorenz curves. Journal of Statistical and Econometric Methods. 1(3): 2012: 53-62.

Golden, J. (2008). A simple geometric approach to approximating the Gini coefficient. J. Economic Education, 39(1):68-77.