Working Paper 2007-6

Multifractal analysis of spatial Income curdling: Theory

Francisco Rosales, International Potato Center (CIP)
Adolfo Posadas, International Potato Center (CIP)
and Facultad de Ciencias Fisicas DAFI, UNMSM
Roberto Quiroz, International Potato Center (CIP)

ISBN 978-92-9060-337-5
Natural Resources Management Division
Working Paper
No. 2007-6

Working Paper

Multifractal analysis of spatial Income curdling: Theory

Francisco Rosales, International Potato Center (CIP)
Adolfo Posadas, International Potato Center (CIP) and
Facultad de Ciencias Fisicas DAFI, UNMSM
Roberto Quiroz, International Potato Center (CIP)

The Natural Resources Management Division Working Paper Series comprises preliminary research results published to encourage debate and exchange of ideas. The series also includes documentation for research methods, simulation models, databases and other software. The views expressed in this series are those of the author(s) and do not necessarily reflect the official position of the International Potato Center.

Comments are invited.

This series is available on the internet at www.cipotato.org

Multifractal analysis of spatial
Income curdling: Theory

ISBN 978-92-9060-337-5

International Potato Center
P.O.Box 1558, Lima 12, Peru
cip@cgiar.org • www.cipotato.org

Produced by the CIP Communication and Public Awareness Department (CPAD)

Production Coordinator
Cecilia Lafosse

Design and Layout
Elena Taipe and contributions from Graphic Arts

Printed in Peru by Comercial Gráfica Sucre
Press run: 100
December 2007

Abstract

The spatial curdling index (C) is presented in order to study income concentration in a two-dimensional field originated by a cascade process resembling resource allocation maps. Computational simulations of spatial distributions show the relation between C and measures of income inequality such as the Gini coefficient (G) forms a concave area. Acknowledgement of this fact brings additional information to formulate geopolitical strategies for income inequality reduction, for comparison purposes in a set of countries, or across time in a single country.

Keywords: curdling, cascade processes, Gini coefficient.

JEL codes: C15, C63, D63, R12.

Acknowledgements

This research was funded by the Canadian International Development Agency through the project "Achieving Sustainable Rural Development in the Peru-Bolivia Altiplano through the Improvement of Andean Agriculture (ALTAGRO)". We want to thank Ivonne Valdizán for her support in the preparation of the manuscript and to Dr Víctor Mares for reviewing and commenting on the drafts.

Multifractal analysis of spatial Income curdling: Theory

INTRODUCTION

This document is part one of a report on a study conducted to improve cluster measures on resource allocation maps developed in the Natural Resources Management (NRM) Division of the International Potato Center (CIP). The overall intention of developing this tool is to have an adequate characterization of spatially distributed random variables to avoid oversimplifying hypotheses in crop, climate and economic models.

In this first report the theoretical part of cluster measures and some of the analytical results of studying income distribution are presented. Specifically, a theoretical framework is built for policy analysis that accounts for spatial curdling when analyzing income distribution (Champernowne, 1953; Lydall, 1959; Mandelbrot, 1960). Different from regular entropy indicators such as Theil and Atkinson indices (Atkinson, 1970), or the Gini coefficient (Champernowne, 1974; Gini, 1912), this indicator considers the system's spatial structure.

The measure of spatial curdling will be presented as an index derived from multifractal theory. Therefore a review of some concepts from multifractal formalism (Schertzer et al, 1997; Feder, 1988; Mandelbrot, 1975; Parisi and Frisch, 1985), will be needed. This approach will be useful to characterize spatial income sets resembling resource allocation maps (Appleby, 1996), and assumed to be characterized by a multiplicative cascade process like the one first presented in Kolmogorov's works on turbulence (Kolmogorov, 1941a; Kolmogorov, 1941b). Furthermore, to prove the consistency of the formulations, computational simulations of spatial income measures are executed in two-dimensional fields.

The paper is organized as follows: section 1 reviews some basic concepts of cascade processes in order to establish the basis for the spatial curdling index (C); section 2 presents a comparison between the Gini coefficient (G) and the spatial curdling index (C) by means of a two-dimensional diagram of inequality in the plane C-G; section 3 presents a discussion of the results, and section 4 lists some concluding remarks.

1. CASCADE PROCESSES

This section sets the formalism of multifractals in order to characterize two-dimensional fields. It begins with a brief description of its origins in turbulence theory (Turiel et al., 2006), and continues with the exemplification of deterministic (Feder, 1988) and random (Mandelbrot, 1974) spatial distributions generated by a binomial multiplicative process.

1.1 Turbulence intuition

The origins of multifractal theory can be tracked back to the works by Kolmogorov, (Kolmogorov, 1941a; Kolmogorov, 1941b). Under conditions of fully developed turbulence, variables as the velocity or the local dissipation of energy vary sharply from a location to another and cannot be regarded as deterministic but as random quantities. Let, for instance, $M_\delta(x)$ be the local dissipation of energy at the point x over a neighborhood of radius δ, whose expression is given by

$$M_\delta(x) = \frac{1}{B_\delta(x)} \int_{B_\delta(x)} dx' \sum_{ij} \left[\partial_i v_j(x') + \partial_j v_i(x')\right]^2 , \tag{1.1}$$

where u_i are the components of the velocity vector and $B_\delta(x)$ stands for the ball of radius δ centered around x. Kolmogorov's intuition was that the energy is transmitted from larger scales (Δ) to smaller scales (δ) by means of an injection process defined by a variable $\eta_{\delta\Delta}$ which only depends on the index δ/Δ, as

$$M_\delta = \eta_{\delta\Delta} M_\Delta . \tag{1.2}$$

In Kolmogorov's work the energy injection variable $\eta_{\delta\Delta}$ has a fixed value $\eta_{\delta\Delta} = (\delta/\Delta)^{-\alpha}$, from which it can immediately be deduced that the order-p moments of M_δ can be related with those of M_Δ in a very simple way, namely

$$\left\langle M_\delta^p \right\rangle = \left(\frac{\delta}{\Delta}\right)^{-\alpha p} \left\langle M_\Delta^p \right\rangle \sim \delta^{-\alpha p} ,$$

where all dependence on δ of the order-p moment of M is concentrated in the power law $\delta^{\alpha p}$, which is similar to what experimental results show. That is

$$\left\langle M_{\delta}^{p} \right\rangle \sim \delta^{\tau(p)}, \tag{1.3}$$

a property which is known as self-similarity. Unfortunately, the exponents $\tau(p)$ obtained in the experiments have not a linear dependence on p. In general, $\tau(p)$ as a function of p is a concave curve, in opposition to the usual linear scaling. To describe this "anomalous scaling" Kolmogorov's decomposition can still be applied but now $\eta_{\delta\Delta}$ in equation (1.2) has to be interpreted as a random variable, independent of Δ. Moreover, in order to produce a consistent scheme, the variables $\eta_{\delta\Delta}$ have to be infinitely divisible (Brzezniak and Zastawniak, 1999), because the process to go down from scale Δ to scale δ can be verified directly or in several stages forming a cascade process [see figure (1.1)].

The property of self-similarity led researchers to propose a generation model based on the existence of local scale-invariant laws. First, it is assumed that at any point x,

$$M(x+\delta) - M(x) = \delta^{\alpha(x)}, \tag{1.4}$$

holds. That is, all dependency on the scale parameter δ is conveyed by the power law factor $\delta^{\alpha(x)}$. The exponent $\alpha(x)$, which is a function of the point x under study, is called singularity exponent of point x. Then, singularity exponents can be arranged in special sets called singularity components S_α defined as

$$S_\alpha = \{x : \alpha(x) = \alpha\} \tag{1.5}$$

In order to close the model, it is required that the singularity components are of fractal character. The singularity spectrum associated to the multifractal hierarchy of fractal components is the function $f(\alpha)$ defined by the Hausdorff dimension of each component S_α, namely

$$f(\alpha) = \dim_H(S_\alpha) \tag{1.6}$$

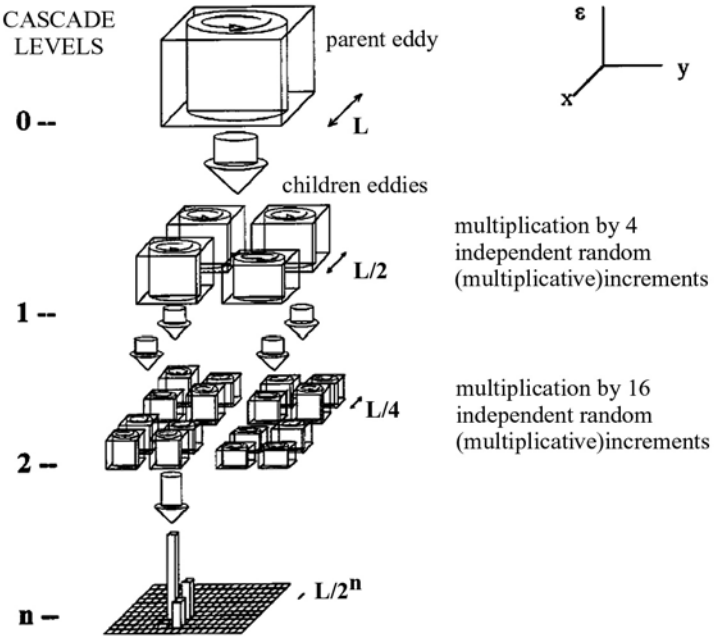


Figure 1.1 Multiplicative cascade scheme (Schertzer, Lovejoy et al., 1997).

Under some assumptions on the homogeneity and isotropy of the statistics of local singularities (Parisi and Frisch, 1985), it is possible to derive a relation between self-similarity exponents $\tau(p)$ and the singularity spectrum $f(\alpha)$. Furthermore, it has been proven that self-similar exponents $\tau(p)$ can be computed from a Legendre transformation of the singularity spectrum $f(\alpha)$

$$\tau(p) = \inf_{\alpha} \{\alpha p + d - f(\alpha)\} \tag{1.7}$$

Through equation (1.7) it is possible to relate statistics and geometry. It is then evident that the singularity spectrum contains all the information about self-similarity, hence about the multiplicative process, which means it describes the statistics of changes in scale. In this sense, precise knowledge of the singularity spectrum will provide valuable information on the inner structure of the system.

The following subsection shows deterministic and random examples of a binomial multiplicative process that will be useful to understand the concepts previously reviewed.

1.2 Deterministic Multiplicative Processes

Let's start with a one-dimensional example of the binomial multiplicative process. Let a mass $N < \infty$ be distributed over a line interval $S = [0,1]$. In order to characterize this distribution it is possible to divide S into cells of resolution $\delta = 2^{-n}$ so that $N = 2^n$ line intervals are sufficient to cover S, and the parameter n can be seen as the number of generations in the binary subdivision of S. If one uses indexes $i = 0,1,...,N-1$ to name generated line segments, the distribution of the mass over S can be specified at resolution δ by the mass v_i in the $i = th$ cell. The fraction of the global mass $\mu_i = v_i/v$ is a good measure of the content in cell i. Namely

$$W_\delta = \{\mu_i\}_{i=0}^{N-1} \tag{1.8}$$

gives a complete description of the distribution at resolution δ. Similarly, a measure $M(T)$ of a subregion T can be written as

$$M_\delta(T) = \sum_{i \in T} \mu_i \tag{1.9}$$

which is equivalent to equation (1.1).

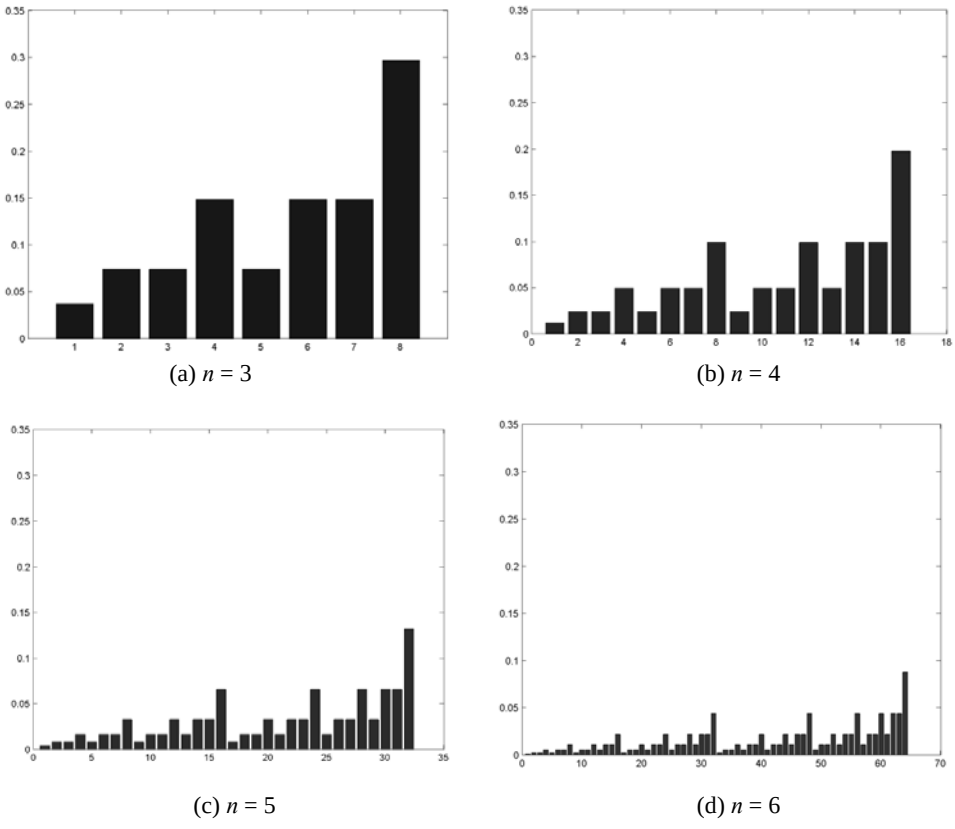
Let's consider an example. First divide S into 2 parts of equal length $\delta = 2^{-1}$, i.e., $n = 1$. The left half is given a fraction a of the mass so $\mu_0 = a$, and the right hand is given a measure $\mu_1 = 1 - a$. Increase the resolution δ to 2^{-2}. The multiplicative process divides the mass in each part in the same way, so one will have

$$W_2 = \{\mu_i\}_{i=0}^{2^2-1} = \{\mu_0\mu_0, \mu_0\mu_1, \mu_1\mu_0, \mu_1\mu_1\}.$$

Every time a new generation is made it releases a set of shorter segments carrying a smaller part of the total measure. Figure (1.2) shows a 1D numerical example when (a, b) = (1/3, 2/3), and figure (1.3) shows a similar situation for a 2D image with weights (a, b, c, d) = (0.1, 0.2, 0.3, 0.4). In the first case the simulation is performed for n = 3, 4, 5 and 6; and in the second for n = 2, 3, 4 and 5.

As one may expect, at the nth generation, there is one cell concentrating most of the measure. In fact for the 1D numerical example, if $a > 1/2$, this cell will concentrate $(1-a)^n$. In general with $\varepsilon = k/n$ interpreted as a portion of a's in each of the generated branches of the binary tree shown in figure (1.1), and k = 0, 1, . . . n, one can write the number of sets satisfying proportion ε by

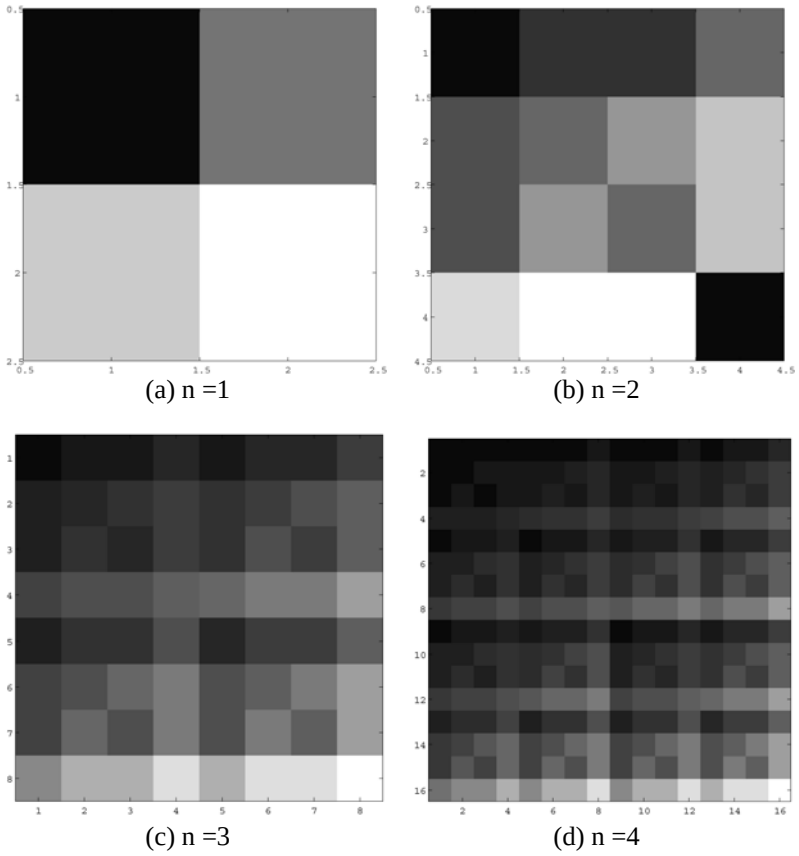
$$N_n(\varepsilon) = C_{\varepsilon n}^{n} = \frac{n!}{(\varepsilon n)!((1-\varepsilon)n)!} \tag{1.10}$$



(a) $n = 3$ (b) $n = 4$

(c) $n = 5$ (d) $n = 6$

Figure 1.2 1D Binomial multiplicative cascade example.



(a) n =1 (b) n =2

(c) n =3 (d) n =4

Figure 1.3
2D Binomial multiplicative cascade example.

where measures are all the same and equal to $\mu_\varepsilon = \Delta^n(\varepsilon)$ with

$$\Delta(\varepsilon) = \mu_0^{\varepsilon}\mu_1^{(1-\varepsilon)} = a^{\varepsilon}(1-a)^{1-\varepsilon} \tag{1.11}$$

The total measure of the segments representing the population is

$$M_\delta(S) = \sum_{i=0}^{2^n-1}\mu_i \sum_{\varepsilon=0}^{1} N_n(\varepsilon)\Delta^n(\varepsilon) = 1 \tag{1.12}$$

The cells describing the distribution of the mass cover the line completely and contain all of the measure, which is the whole mass.

1.2.1 SET OF SINGULARITY COMPONENTS

In the nth generation $N_n(\varepsilon)$ line segments have length 2^{-n} and the same measure μ_ε. This segments form a fractal set of points $S_n(\varepsilon)$ of the unit interval S = [0, 1]. To see this we cover the set with intervals of length δ and form the *d*-measure $\overline{M}_\delta(S_\varepsilon)$ (Feder, 1988),

$$\overline{M}_\delta(S_\varepsilon) = \sum_{S_\varepsilon} \delta^d = N_n(\varepsilon)\delta^d \tag{1.13}$$

and determine the fractal dimension $D(\varepsilon)$ of this set by studying how the M_d behaves as $\delta \to 0$, and $n \to \infty$.

Using Stirling's formula for n! in equation (1.10), one finds an approximate expression for $N_n(\varepsilon)$:

$$N_n(\varepsilon) \cong \frac{e^{-n(\varepsilon \ln \varepsilon + (1-\varepsilon)\ln(1-\varepsilon))}}{\sqrt{2\Pi n\varepsilon(1-\varepsilon)}}. \tag{1.14}$$

Noting that $n = -\dfrac{\ln\delta}{\ln 2}$, one can rewrite measure $M_d(S_\varepsilon)$ as

$$\overline{M}_\delta(S_\varepsilon) \sim \delta^{-f(\varepsilon)}\delta^d \tag{1.15}$$

with exponent $f(\varepsilon)$ defined by

$$f(\varepsilon) = -\frac{\varepsilon \ln \varepsilon + (1-\varepsilon)\ln(1-\varepsilon)}{\ln 2} \tag{1.16}$$

It follows that the d-measure $\overline{M}_\delta$ for the set S_ε remains finite as $\delta \to 0$ only for $f(\varepsilon)$, and therefore the fractal dimension $D(\varepsilon)$ of each set S_ε is $f(\varepsilon)$. This is one reason for the term "multifractal".

1.2.2 SINGULARITY SPECTRUM $f(\alpha)$

An alternative way of showing dimensions distribution is to plot the f as function of singularity exponents α instead of using the parameter ε. Namely considering equation (1.4).

$$\mu_{\varepsilon} = M_{\delta}\left(x(\varepsilon)+\delta\right) - M(\varepsilon) = \delta^{\alpha} \tag{1.17}$$

where the parameter α characterizes singularities of the measure M at each point x where proportion ε occurs, i.e., $M_{\delta}(x(\varepsilon))$.

It follows from equations (1.11), (1.17) and $\delta = 2^{-n}$ that the local measure of the mass has a singularity exponent of

$$\alpha(\varepsilon) = \frac{\ln \mu_{\varepsilon}}{\ln \delta} = -\frac{\varepsilon \ln a + (1-\varepsilon)\ln(1-a)}{\ln 2} \tag{1.18}$$

which clearly is a linear function of ε with

$$\begin{aligned} \alpha_{\min} &= -\ln(1-a)/\ln 2, \quad for\ \varepsilon = 0, \\ \alpha_{\max} &= -\ln a/\ln 2, \qquad for\ \varepsilon = 1, \end{aligned} \tag{1.19}$$

The measure $M_{\delta}(x)$ has singularities characterized by a set of singularity exponents α with intensities showed in $f(\alpha(\varepsilon))$ which is obtained when the value of ε in equation (1.18) is replaced in equation (1.16), showing the same as equation (1.6) when evaluated in the set given by (1.5).

The generated $f(\alpha)$ curve has a few features that will be discussed further. For this matter it would be useful to see the form of its derivative

$$\frac{df(\alpha)}{d\alpha} = \frac{\ln \varepsilon - \ln(1-\varepsilon)}{\ln a - \ln(1-a)} \tag{1.20}$$

Support Dimension

The maximum is $f(\alpha_0) = 1$, with $\varepsilon = 1/2$

$$f_{\max} = f(\alpha_0) = 1, \tag{1.21}$$
$$\alpha_0 = -\frac{\ln a + \ln(1-a)}{2\ln 2}$$

It is a general result that the maximum value of the fractal dimension of the subset S_α equals the fractal dimension of the support of the measure, which is 1 since the measure is defined over the whole unit interval.

Curdling Dimension

Another special point on the $f(\alpha)$ curve occurs at $df(\alpha)/d\alpha = 1$, that is $\varepsilon = a$, and

$$f(\alpha_S) = \alpha_S = S, \tag{1.22}$$

the fractal dimension of the set S_{α_S} which is recognized as the curdling of the binomial multiplicative process.

1.2.3 Self-similarity Exponents $\tau(p)$

A generalized version of measure $\overline{M}_\delta(\delta)$ in equation (1.13) can be obtained considering order-p moments

$$\left\langle \overline{M}_\delta^p \right\rangle = \sum_{i=1}^{N} \mu_i^p \delta^d = N(\delta, p)\delta^d \tag{1.23}$$

so $\tau(p)$ becomes the parameter that allows finite measure of the set $\left\langle \overline{M}_\delta^p \right\rangle$ when $\delta \to \infty$, as expressed in equation (1.3). The measure is characterized by a whole sequence of exponents $\tau(p)$ that controls how the moments of the probabilities $\{\mu_i\}$ scale with δ. It follows from equation (1.23) that the weighted number of boxes $N(\delta, p)$ has the form

$$N(\delta, p) = \sum_i \mu_i^p \sim \delta^{-\tau(\rho)} \tag{1.24}$$

and the mass exponent is given by

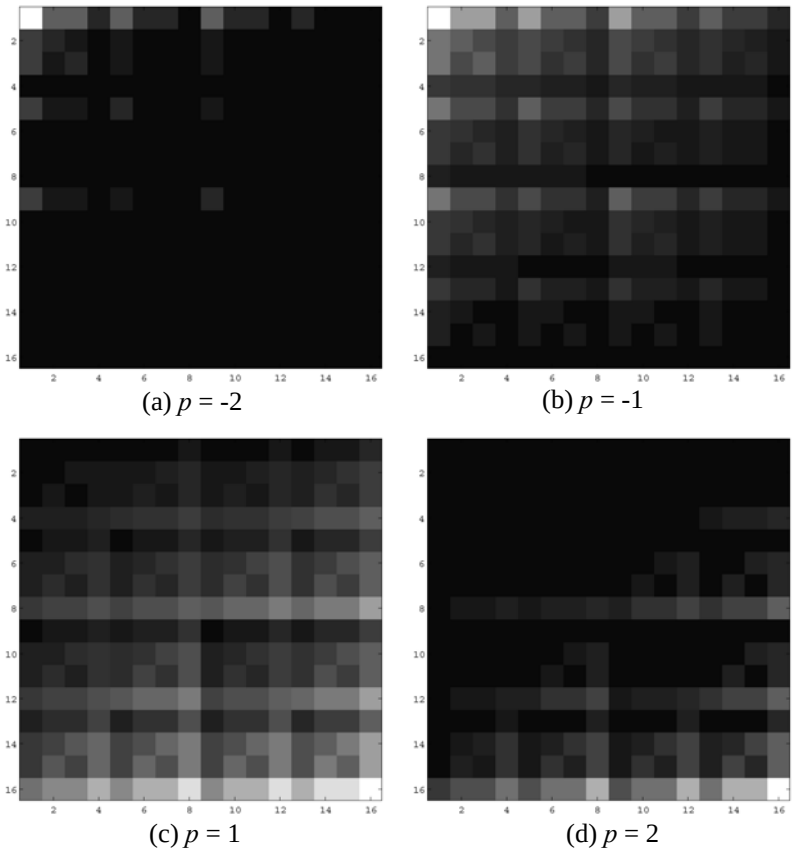
$$\tau(p) = -\lim_{\delta \to 0} \frac{\ln N(\delta,p)}{\ln \delta} \tag{1.25}$$

One first notes that when $p = 0$, $\mu_i^{p=0} = 1$ therefore $N(\delta, p = 0) = N(\delta)$ is simply the number of boxes needed to cover the set, and $\tau(p) = D$ becomes the fractal dimension of the set. On the other hand, in equation (1.23) higher order moments, i.e., $p >> 0$ can be interpreted as favoring cells with high mass concentration, which is helpful for analyzing curdling measures. In the same fashion, negative order moments, i.e., $p << 0$ favors cells with small mass concentration. Figure (1.4) shows how measure changes when order-p moments are applied.

Other interesting properties of the $\tau(p)$ generated curve are better viewed considering its derivative

$$\frac{d\tau(p)}{dp} = -\lim_{\delta \to 0} \left(\sum_i \mu_i^p \ln \mu_i \Big/ \left(\sum_i \mu_i^p \right) \ln \delta \right), \tag{1.26}$$

and two particular cases for identifying singularity exponents range by employing either μ_- or μ_+, that is, the minimum or the maximum value of μ_i in the sum when limits are taken in equation (1.26) for $p \to -\infty$ and $p \to \infty$:



(a) p = -2 (b) p = -1

(c) p = 1 (d) p = 2

Figure 1.4 Order-p moments of a 2D measure.

$$\left.\frac{d\tau(p)}{dp}\right|_{p\to-\infty} = -\lim_{\delta\to 0}\frac{\ln\mu_{-}}{\ln\delta} = -\alpha_{\max} \tag{1.27}$$

$$\left.\frac{d\tau(p)}{dp}\right|_{\rho\to+\infty} = -\lim_{\delta\to 0}\frac{\ln\mu_{+}}{\ln\delta} = -\alpha_{\min}, \tag{1.28}$$

In fact, the generalization of this result shows a close relation between τ and α by $\frac{d\tau(p)}{dp} = \alpha$, by means of a Legendre transformation.

Further more, for p = 1 one finds that $\frac{d\tau(p)}{dp}$ has an interesting value:

$$\left.\frac{d\tau(p)}{dp}\right|_{p=1} = -\lim_{\delta\to 0}\frac{\sum_i \mu_i \ln\mu_i}{\ln\delta} = \lim_{\delta\to 0}\frac{S(\delta)}{\ln\delta}, \tag{1.29}$$

where $S(\delta)$ is the curdling of the partition of the measure $M = \{ \mu_i \}_{i=0}^{N-1}$ over boxes of size δ, which may be written as

$$S(\delta) = -\sum_i \mu_i \ln \mu_i \sim -\alpha_1 \ln \delta \tag{1.30}$$

The exponent $\alpha_1 = -\frac{d\tau(p)}{dp}\Big|_{p=1} = f_s$ is also the fractal dimension of the set onto which the measures concentrate and describes the scaling with the box size δ of the curdling of the measure. Table (1.1) summarizes some important values in $f(\alpha)$ and $\tau(p)$ curves.

p	$\tau(p)$	$\alpha = -d\tau(p)/dp$	$f = p\alpha + \tau(p)$
$p \to -\infty$	$\sim -p\alpha_{max}$	$\to \alpha_{max} = -\ln\mu_{-}/\ln\delta$	$\to 0$
$p = 0$	D	α_0	$f_{max} = D$
$p = 1$	0	$\alpha_1 = -S(\delta)/\ln\delta$	$f_S = \alpha_1 = S$
$p \to +\infty$	$\sim -p\alpha_{min}$	$\to \alpha_{min} = -\ln\mu_{+}/\ln\delta$	$\to 0$

Table 1.1 Critical values in multifractal formalism.

1.2.4 THE RELATION BETWEEN $f(\alpha)$ AND $\tau(p)$

Since the set S is fractal with fractal dimension D, the subsets have fractal dimensions $f(\alpha) \le D$, where the number of segments of length δ needed to cover sets S_α with $\alpha \in [\alpha, \alpha + d\alpha]$ is

$$N(\alpha, \delta) = \rho(\alpha) d\alpha \, \delta^{-f(\alpha)} \tag{1.31}$$

Here $\rho(\alpha)$ is the number of sets from S_α to $S_{\alpha+d\alpha}$. For these sets the measure μ_α in a cell of size δ has the power law dependence in equation (1.17) on the length scale δ so that we may write $\mu_\alpha = \delta^\alpha$, and therefore the measure $\overline{M}_\delta$ for the set *S* given in equation (1.23) may be written

$$\left\langle \overline{M}_\delta^{\,p} \right\rangle = \int \rho(\alpha) d\alpha \, \delta^{-f(\alpha)} \delta^{\alpha p} \delta^d = \int \rho(\alpha) d\alpha \, \delta^{\alpha p - f(\alpha) + d} \tag{1.32}$$

The integral in (1.32) is dominated by the terms where the integrand has a maximum value. In other words for

$$\frac{d}{d\alpha}\{\alpha p - f(\alpha)\}\Big|_{\alpha+\alpha(p)} = 0, \tag{1.33}$$

i.e., when equation (1.7) holds. The integral in equation (1.32) is asymptotically given by

$$\left\langle \overline{M}_{\delta}^{p} \right\rangle \sim \delta^{\alpha p - f(\alpha)+d} \tag{1.34}$$

Here $\left\langle \overline{M}_{\delta}^{p} \right\rangle$ remains finite in the limit $\delta \to 0$ if d equals the mass exponent $\tau(p)$ given by

$$\tau(p) = f(\alpha(p)) - \alpha(p)p \tag{1.35}$$

where $\alpha(p)$ is the solution of equation (1.33). Thus the mass exponent is given in terms of singularity exponent $\alpha(p)$ for the mass and the fractal dimension $f(\alpha(p))$ of the set that supports this exponent.

On the other hand, if one knows mass exponents $\tau(p)$, determine singularity exponents α and f by means of equations (1.33) and (1.35), which gives

$$\begin{aligned} \alpha(p) &= -\frac{d}{dp}\tau(p) \\ f(\alpha(p)) &= \alpha(p)p + \tau(p) \end{aligned} \tag{1.36}$$

These two equations give a parametric representation of the $f(\alpha)$ curve, i.e., the fractal dimension, $f(\alpha)$ of the support of singularities in the measure with singularity exponents α. The $f(\alpha)$ curve characterizes the measure and is equivalent to the self-similarity exponents $\tau(p)$. The pair of equations (1.36) in effect constitute a Legendre transformation from independent variables τ and p to independent variables f and α.

1.3 Random Multiplicative Processes

A random representation of multiplicative processes is needed in order to model real world phenomena. The first proposal made in this direction was the well known *Kolmogorov's third hypothesis* (Kolmogorov, 1941a; Kolmogorov, 1941b), which stated that mass distribution at each point [see equation (1.4)] follows a lognormal probability law. The first proposal to justify lognormality was performed in Yalom's work. This attempt based the proof on a probabilistic argument involving a self-similar cascade sometimes called "microcanonical". Figure (1.5) shows an example of a microcanonical measure, and presents the resulting singularity spectrum when compared with the one generated by a deterministic binomial process (a, b, c, d) = (0.1, 0.2, 0.3, 0.4).

Unfortunately, it has been some time since this lognormality justification was questioned. For example, works by Mandelbrot (Mandelbrot, 1972; Mandelbrot, 1974) conclude that even though Yalom's construction process is valid, it might be better to use a "canonical" approach. That is, to keep the intuition that a cascade process divides an eddy exactly into subeddies but combining splitting with some kind of diffusion, in such a way that conservation of dissipation only holds on the average. Figure (1.6) presents an example of a canonical measure when a uniform probability law is considered, and its resulting singularity spectrum is compared with the one generated by a deterministic binomial process (a, b, c, d) = (0.1, 0.2, 0.3, 0.4).

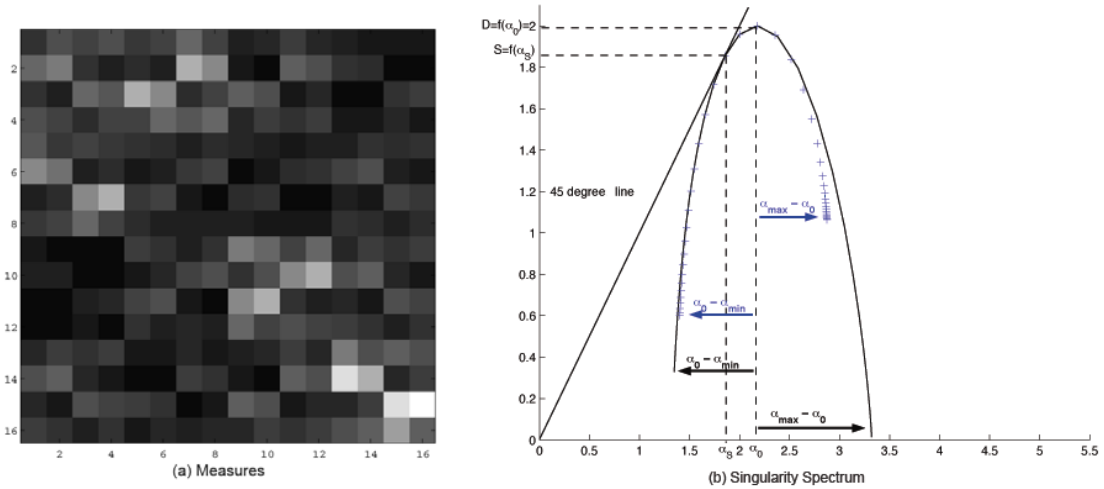


Figure 1.5 Random microcanonical process example.

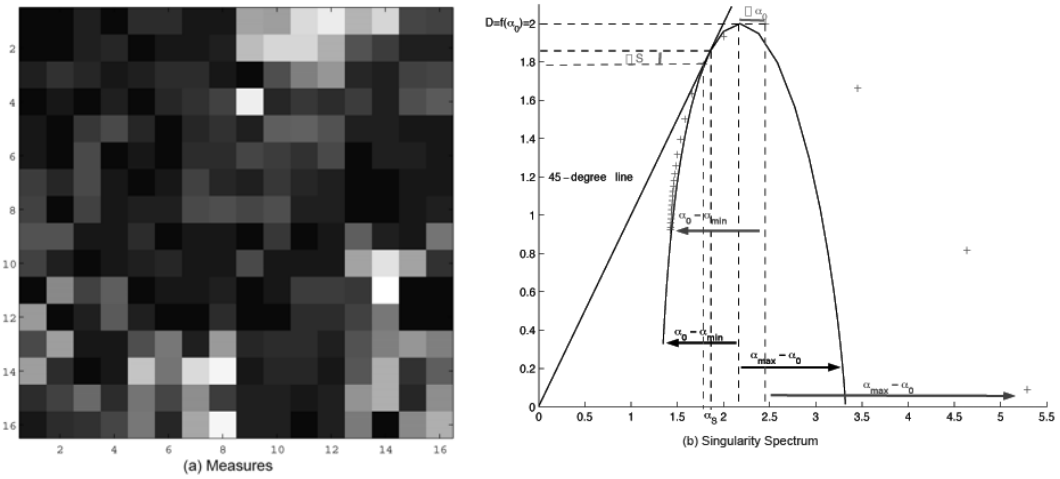


Figure 1.6 Example of random canonical process.

2. MEASURES OF INCOME DISTRIBUTION

This section compares measures of income inequality such as the Gini coefficient (G) with one alternative multifractal index derived from the formalism explained previously. Namely it will be showed that the spatial curdling index (C) captures features that can be of value for policy makers that are not taken into account by measures such as G.

2.1 The Gini Coefficient G

The Gini coefficient (Gini, 1912), is a measure of income inequality defined as an index with values between 0 (perfect income equality) and 1 (perfect income inequality): the numerator is the area between the Lorenz curve of the distribution (Lorenz, 1905) and the uniform distribution line; the denominator is the area under the uniform distribution line.

Computation of the Gini coefficient is achieved building an index on the Lorenz curve diagram showed in figure (2.1). If the area between the line of perfect equality and Lorenz curve is A, and the area under the Lorenz curve is B, then the Gini coefficient is A/(A + B). Since A + B = 1/2, the Gini coefficient, G = 2A =1 -2B. If the Lorenz curve is represented by the function L(x), the value of B can be found by integration as

$$G = 1 - \int_0^1 L(x)dx \qquad (2.1)$$

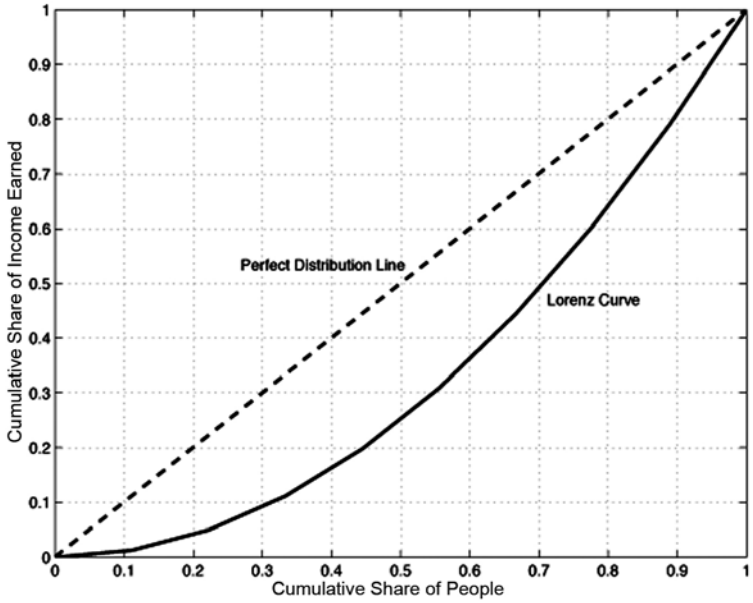


Figure 2.1
Lorenz Curve diagram.

The Cumulative Share of people with Lower Income

As can be noted, this coefficient omits spatial distribution. This occurs because income is characterized only by distribution law and consequently the households' locations across space are lost. This can be viewed as part of what is known as the "anonymity" principle of the Gini coefficient (Gini, 1912). In the following subsection the use of the spatial curdling index in order to account for spatial income curdling is suggested.

2.2 Spatial Curdling Index C

This subsection will show the use of multifractal formalism to characterize spatial income concentration by index C. The purpose of this is to account for spatial concentration features of income to complement traditional measures of income inequality.

In previous subsections it has been shown that the curdling dimension captures the measure of fractal sets where curdling occurs. That is when the singularity spectrum $f(\alpha)$ is tangent to the 45-degree line, [see equation (1.22)]. Loosely speaking, this can be understood as a measure of income curdling in multifractal sets. Under this interpretation, one can see that systems with low concentration have a smaller curdling dimension $f(\alpha_S)$ than more highly concentrated ones. Holding this idea, one can write

$$C = 1 - \frac{f(\alpha_S)}{2}, \qquad (2.2)$$

as a normalized measure of income curdling in a two-dimensional system ranging from 0 to 1. As C approaches 1, it is possible to state that a large part of income curdles spatially; and as C gets near 0 one can say income is found across space and does not concentrate. Likewise in extreme cases such as C = 1, one finds perfect curdle (a nation's entire income concentrated in one individual only); C = 1/2, imperfect curdle (income concentrated in a well-defined section across space); and C = 0, no curdle at all[1] (all individuals have the exact same income). In consequence, interpretation of extreme values of C and G are actually aligned.

Figure (2.2) show examples of the situations formerly mentioned and a binomial multiplicative cascade with (a, b, c, d) = (0.1, 0.2, 0.3, 0.4). These cases are then located in the multifractal spectrum diagram [see figure (2.3)] of the binomial cascade.

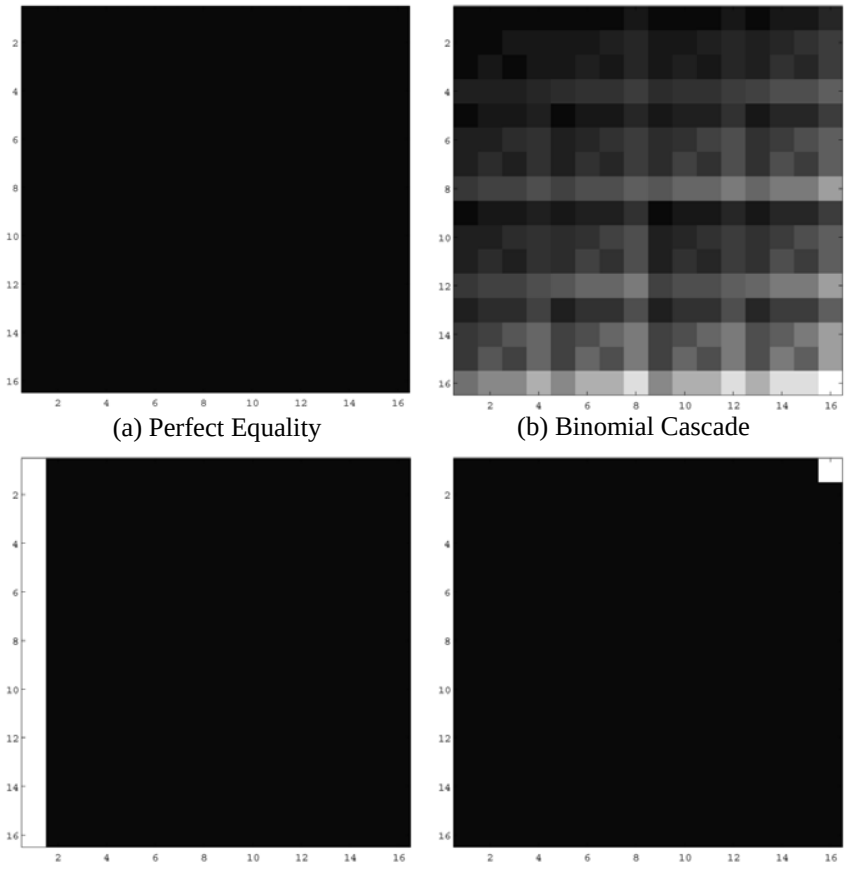
Having clarified the location of critical values of C in a singularity spectrum diagram, one will use computational simulations of alternative measures to show graphically that the spatial curdling index C accounts for spatial income concentration. Figure (2.4) accomplishes this by showing the measure generated by uniform distribution (a), and its changes when the intensity of its concentration varies from low to high concentration [(b), (c) and (d)].

In figure (2.5) singularity spectrums of the measures in figure (2.4) are plotted, showing higher curdling dimensions [when $f(\alpha) = \alpha$] when the concentration of the measure increases.

2.3 The Relation between G and C

In order to visualize the relation between the Gini coefficient G and the spatial curdling index C, figure (2.3) shows the locations of examples previously discussed in the C - G plane. As shown in figure (2.3), the relation between these two indicators is far from being linear [even though both indicators coincide in the extreme values of perfect equality (perfect curdle) and perfect inequality (no curdle at all)].

[1] Actually the last three situations do not correspond to multifractal or monofractal sets, but one can still calculate their curdling dimensions by applying equation (1.22), knowing that they would be equal to their respective Euclidian dimensions.



(a) Perfect Equality (b) Binomial Cascade

(c) Imperfect Equality (d) Perfect Inequality

Figure 2.2
Examples of C measures of extreme values.

A deeper insight into this phenomenon will show that not only is the relation non linear, but that there is an area relating these two indicators instead of a function. This happens because for a given Gini coefficient one can find spatial distributions with many different curdling patterns[2]. In order to find the shape of the area relating G to C, one simulates all the possible spatial distributions generated from different populations with several Gini coefficient values [See figure (2.7)].

[2] One should note that this does not happen in extreme cases G = C = 0 and G = C = 1, nor in any monofractal set for that matter.

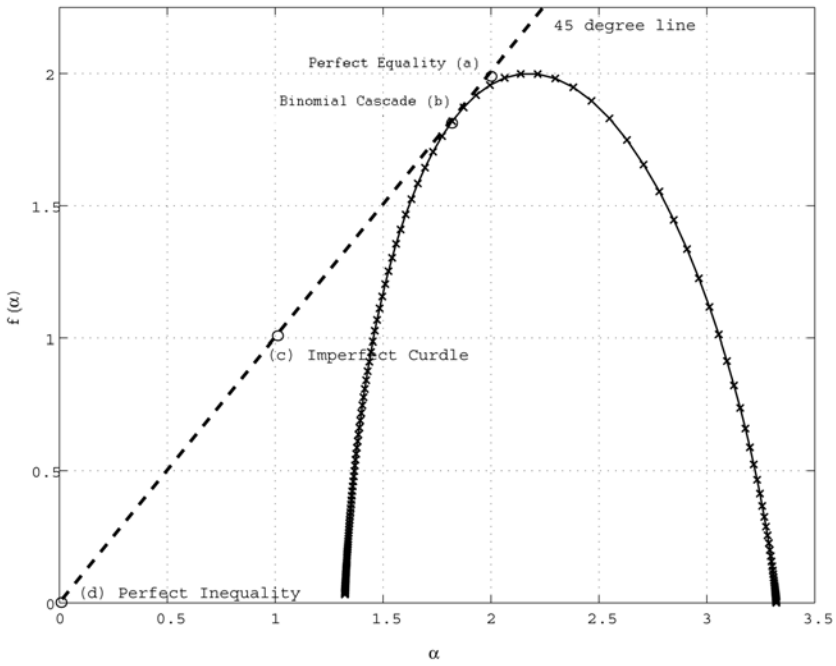


Figure 2.3 Extreme values of C in the singularity spectrum diagram.

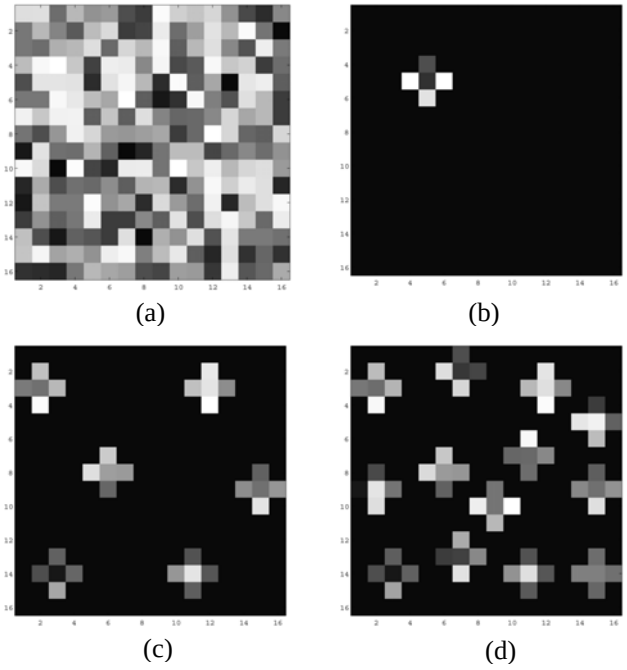


Figure 2.4 Examples of measures with different concentration intensities.

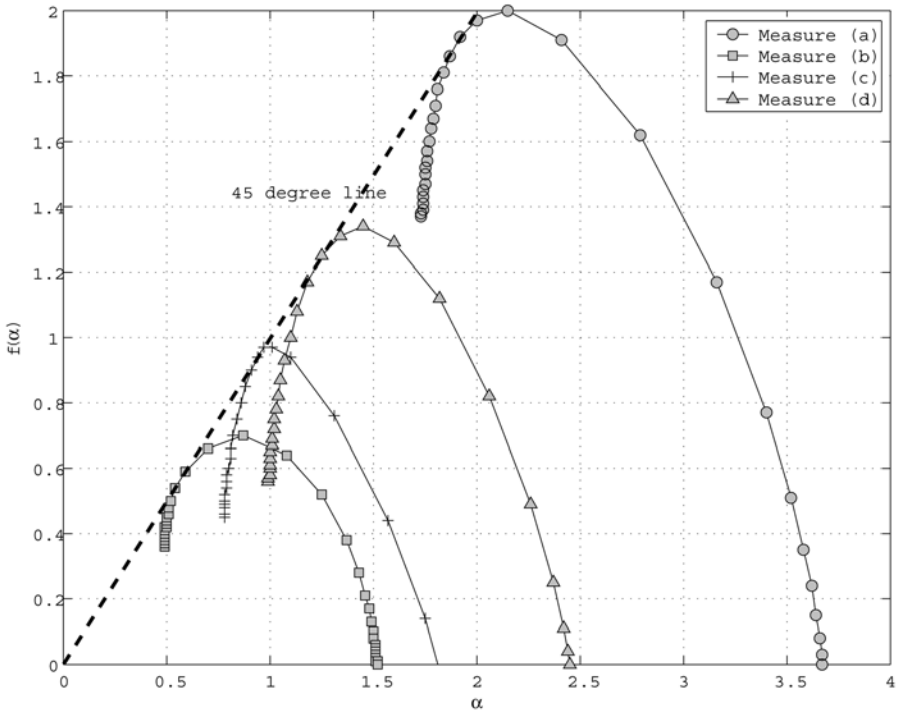


Figure 2.5 Singularity spectrums of measure examples.

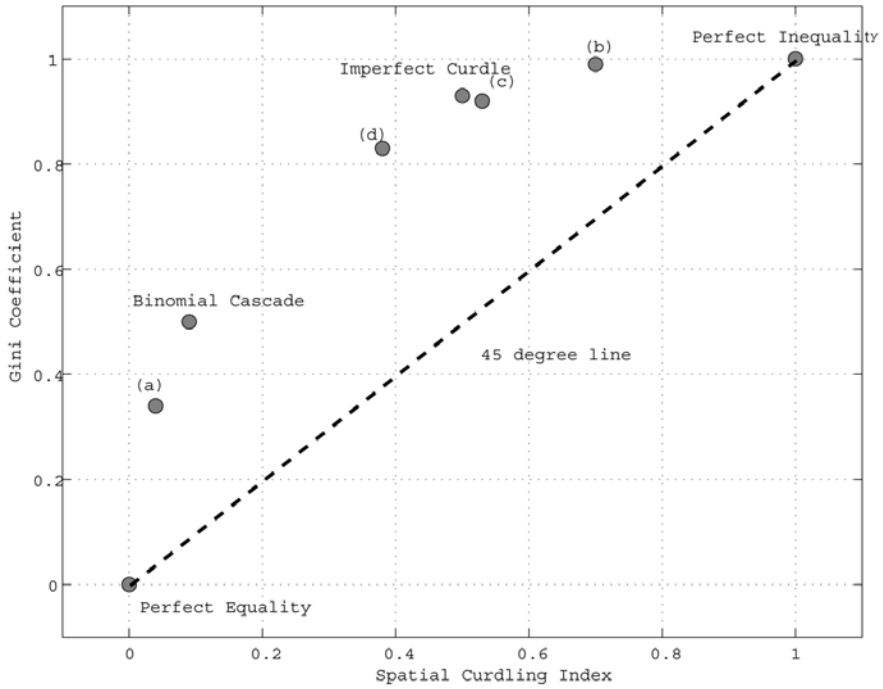


Figure 2.6 Comparison between G and C.

Table of Contents

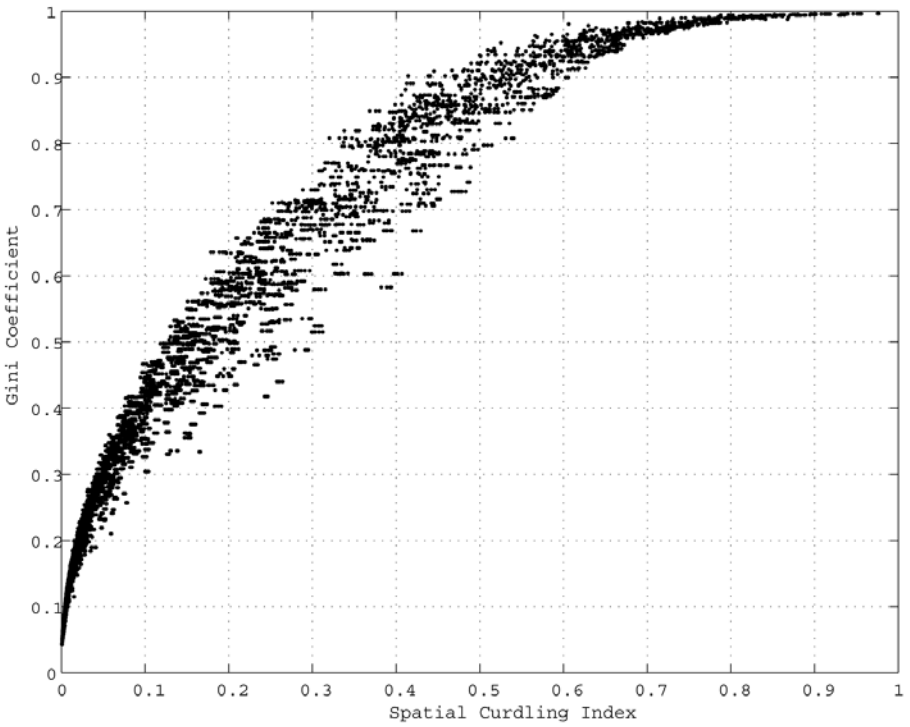


Figure 2.7 Generation of the G-C concave area.

Once stated the relation between G and C, one should ask which would be the more efficient strategy to get to the perfect equality point (0, 0) starting from x = (0.3385, 0.7070) in the G - C plane like the one shown in figure (2.8). In order to answer this question one should first locate the horizontal boundary values in the C-G area for a distribution with G = 0.7070 and the corresponding vertical boundaries for C, like the ones indicated by the black arrows coming from x. Then one can see that any combination reducing G, C or both will cause x to approximate the ideal perfect equality (0, 0). By the same token, a measure characterizing both dimensions of inequality in a country might, for example, be the vector modulus of point x.

Therefore, this type of analysis allows comparing two countries with similar Gini coefficients, by showing the behavior of its spatial distribution. That is, if income concentrates in a single area, or if it is located in several places in the country. Similarly, one can differentiate between countries with similar spatial distribution patterns by looking at its Gini index.

In the following section we will discuss the implications of the results and enumerate the most important conclusions of the study.

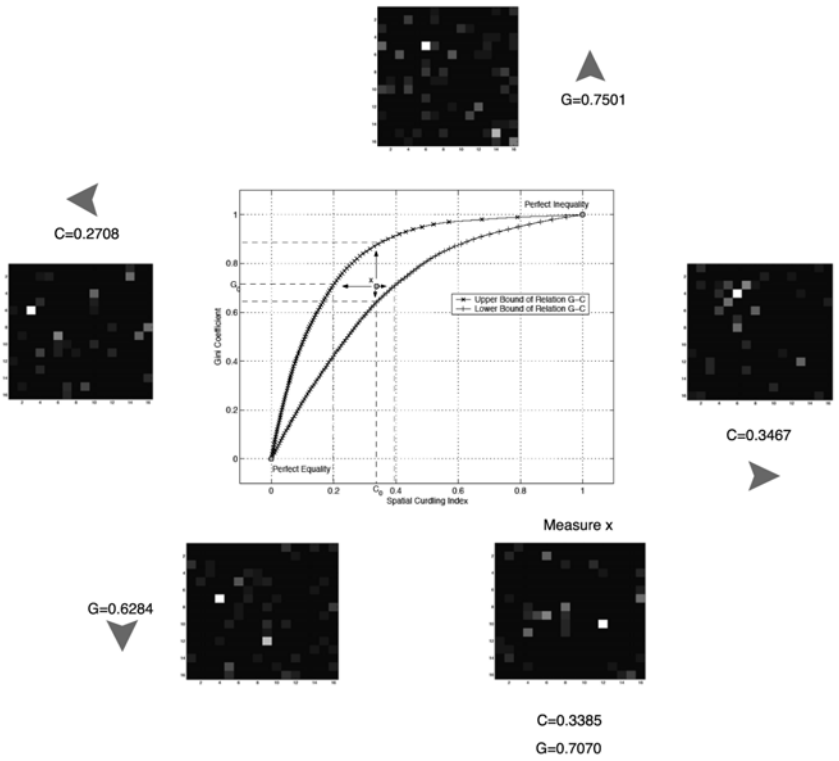


Figure 2.8 Application example.

DISCUSSION

We have presented the spatial curdling index C as a measure of spatial clustering in income inequality. Nonetheless the applications of C are not restricted to this specific issue. In fact, one can apply the same index as the indicator of spatial clustering of any spatially distributed physical, biological, economic or social random variable.

It should be clearly established that we recommend the application of index C (and any other statistical or mathematical technique) to be performed parallel to a structured explanation of the underlying phenomenon. In the next part of this study a practical application including both sides of the story will be developed.

CONCLUDING REMARKS

The spatial curdling index (C) presented in this study adds a new dimension to the traditional measures of inequality by accounting for income curdling when analyzing income distribution. Computational simulations of spatial distributions show the relation between C and traditional measures such as the Gini coefficient (G) forms a concave area, which indicates policy decisions should account for the spatial structure of a country when formulating goals for income inequality reduction, comparing income distribution in a set of countries, or the evolution of income in a single country.

BIBLIOGRAPHY

Appleby, S. 1996. Multifractal characterization of the distribution pattern of human population. *Geographical Analysis* 28(2): 147-160.

Atkinson, A. 1970. On the measurement of inequality. *Journal of Economic Theory* 2(3): 244-263.

Brzezniak, Z. and T. Zastawniak. 1999. *Basic Stochastic Processes*. London: Springer-Verlag.

Champernowne, G. 1953. A model of income distribution. *The Economic Journal* 63(250): 318-351.

Champernowne, G. 1974. A comparison of measures of inequality of income distribution. *The Economic Journal* 84(336): 787-816.

Feder, J. 1988. *Fractals*. New York: Plenum Press.

Gini, C. 1912. Variabilitá e mutabilitá. Reprinted In: *Memorie di Metodologia Statistica,* ed. E. Pizetti and T. Salvemini. Rome Libreria Eredi Virgilio Veschi, 1955.

Kolmogorov, A. 1941a. Dissipation of energy in the locally isotropic turbulence. *Dokl. Akad. Nauk SSSR* 32(1): 19-21.

Kolmogorov, A. 1941b. Local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Dokl. Akad. Nauk SSSR* 30(4): 299-303.

Lorenz, M. O. 1905. Methods of Measuring the Concentration of Wealth. *Publications of the American Statistical Association* 9(70): 209-219.

Lydall, H. F. 1959. The Distribution of Employment Incomes. *Econometrica* 27(1): 110-115.

Mandelbrot, B. 1960. The pareto-levy law and the distribution of income. *International Economic Review* 1(2):79-106.

Mandelbrot, B. 1972. Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermitent turbulence. In: *Statistical Models and Turbulence*, ed. M. Rosenblatt and C. Van Atta. *Lecture Notes in Physics* 12:333-351. New York: Springer.

Mandelbrot, B. 1974. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics* 62(2): 331-358.

Mandelbrot, B. 1975. Geometry of homogeneous scalar turbulence: iso-surface fractal dimensions 5/2 and 8/3. *Journal of Fluid Mechanics* 72(2): 401-416.

Parisi, G., and U. Frisch. 1985. On the singularity structure of fully developed turbulence. In: *Turbulence and Predictability* In: *Geophysical Fluid Dynamics*, ed. M. Ghil, R. Benzi, G. Parisi. Proceedings of the International School of Physics E. Fermi, North Holland, Amsterdam, 84-87.

Schertzer, D., S. Lovejoy, F. Schmitt, Y. Chiguirinskaya, and D. Marsan. 1997. Multifractal cascade dynamics and turbulent intermittency. *Fractals* 5(3): 427-471.

Turiel, A., C. J. Pérez-Vicente, and J. Grazzini. 2006. Numerical methods for the estimation of multifractal singularity spectra on sampled data: A comparative study. *Journal of Computational Physics* 216(1): 362-390.

CIP's Mission

The International Potato Center (CIP) seeks to reduce poverty and achieve food security on a sustained basis in developing countries through scientific research and related activities on potato, sweetpotato, and other root and tuber crops, and on the improved management of natural resources in potato and sweetpotato-based systems.

The CIP Vision

The International Potato Center (CIP) will contribute to reducing poverty and hunger; improving human health; developing resilient, sustainable rural and urban livelihood systems; and improving access to the benefits of new and appropriate knowledge and technologies. CIP will address these challenges by convening and conducting research and supporting partnerships on root and tuber crops and on natural resources management in mountain systems and other less-favored areas where CIP can contribute to the achievement of healthy and sustainable human development.
www.cipotato.org

CIP is supported by a group of governments, private foundations, and international and regional organizations known as the Consultative Group on International Agricultural Research (CGIAR).
www.cgiar.org

International Potato Center
Apartado 1558 Lima 12, Perú • Tel 51 1 349 6017 • Fax 51 1 349 5326 • email cip@cgiar.org