Measuring inequalities from the Law of Pareto's type

Dr. Mamadou-Youry SALL
Professor
Unit of Formation and Research in
Economic Sciences and Management at
Gaston Berger University, Saint-Louis, Senegal
BP 234
E-mail: sallmy@ufr-seg.org; http://www.ufr-seg.org

Measuring inequalities from the Law of Pareto's type

Abstract:

The study, from the statistical point of view, of empirical distributions, necessarily leads to the theoretical ones which generate these last and permit to understand them.
In Social sciences, there are relatively few theoretical laws elected. The most known are the law of Laplace-Gaüss, the exponential law or the one of Gibrat (Lognormal). Here, we want to see how much the Pareto's law contributes to the study of certain social phenomena like the inequality in the distributions of socio-economic wealth.

The law of Vilfriedo Pareto is still little used in this area. Yet, there is evidence (BROT Mandel, 1973; BARBUT MARC, 1989) that the distribution of many natural phenomena or socio-economic can be represented by such laws. They fit better when we have to discuss issues relating to the allocation of goods and services. Indeed, it appeared that the functions of Pareto's type provide a good representation of phenomena characterized by asymmetric distribution. We must remember that Pareto discovered this function when he was exploring the distribution of wealth in his country.

Here, by considering the results found about the Pareto's law, we want to study the distribution of income regarding their concentration and components. We insist on the graphic representation to compare the distribution of income according to their sources and geographic regions

Key words: Indictors; measurement; income, wealth; inequality; Pareto; fitting

A) INTRODUCTION

An indicator, in any field, generally does not have clear interpretation than when it intervenes like parameter of a theoretical distribution or "law" of reality observed. It is the case, for example, of the average or the standard deviation which are the parameters of Laplace-Gauss distributions.
To say that, the study, from the statistical point of view, of empirical distributions, necessarily leads to those of the theoretical ones which generate these last and permit to understand them.
In Social sciences, there are relatively few theoretical laws elected. The most known are the law of Laplace-Gaüss, the exponential law or the one of Gibrat (Lognormal). Here, we want to see how much the Pareto's law contributes to the study of certain social phenomena like the inequality in the distributions of socio-economic wealth.

The law of Vilfriedo Pareto is still little used in this area. Yet, there is evidence (BROT Mandel, 1973; BARBUT MARC, 1989) that the distribution of many natural phenomena or socio-economic can be represented by such laws. They fit better when we have to discuss issues relating to the allocation of goods and services. Indeed, it appeared that the functions of Pareto's type provide a good representation of phenomena characterized by asymmetric distribution. We must remember that Pareto discovered this function when he was exploring the distribution of wealth in his country.

But, it is really not easy to understand how to utilise this model in the social sciences from the work of Pareto only. Indeed, PARETO uses a method of fitting the observed data to this law which is neither that of least square nor that of maximum likelihood. As a result, it is difficult to understand how he could find the parameters of his law. But, from the results of other works (BARBUT MARC, 1989) one can avoid this difficulty and use a usual method of fitting.

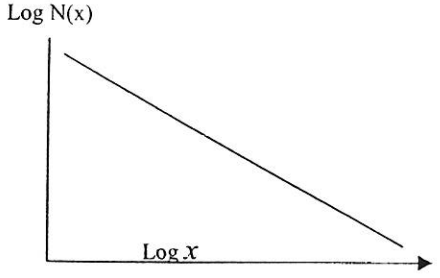
Here, by considering the results found, we want to study the distribution of income regarding their concentration and components. We insist on the graphic representation to compare the distribution of income according to their sources and geographic regions

B) PARETO's distribution

Pareto had to study the distribution of income. He found the reality of this one too chaotic. From where, he has the idea to transform it mathematically. And, like every good engineer, he used the logarithmic function. At the end, he was surprised to find that

Log N(x) decrease for each increasing X

In other words, the income distribution obeys always a law of the form:



$$Log\ N(x) = -\alpha\ log(x) + cte + \varepsilon(x)$$

$$1 < \alpha < 2 \qquad (1)$$

N(x) representing the number of taxpayers having an income higher than *X*, that is to say :

$$N(x) = P_r(X>x) = K x^{-\alpha} \qquad (2)$$

But, one notes that Pareto didn't take into account the zero (origin) of the measurement, which would give the following form:

$$Log\ N(x) = -\alpha \log(x+c)+cte+\varepsilon(x) \qquad 1<\alpha<2 \qquad (3)$$

$$P_r(X>x) = P(x) = K (x+c)^{-\alpha} \qquad (4)$$

This gives as density function:

$$f(x)=\frac{\alpha(x_o+c)^{\alpha}}{(x+c)^{\alpha+1}} \qquad \alpha>1 \quad x\geq x_o \qquad (5)$$

Making *C=0, we will have* $f(x)=\frac{\alpha\, x_o^{\alpha}}{x^{\alpha+1}}$ which is the primary form of Pareto's law

Its moment of k order:

$$m_k = \int_{x_o}^{\infty} t^k dF(t) \qquad (6)$$

Exists if and only if $\alpha > k$

Hence, the average of the Pareto's Law exists if and only if $\alpha > 1$, and will be

$$m_1 = M(x_0) = \frac{\alpha}{\alpha - 1} x_0 + \frac{c}{\alpha - 1} = \beta\, x_o + \mu \qquad (7)$$

$$\beta = \frac{\alpha}{\alpha - 1} \qquad \mu = \frac{c}{\alpha - 1}$$

And the standard deviation $\square^{\square}$ exists if and only if $\square > 2$:

$$\sigma^2 = \frac{\alpha (x_o + c)^2}{(\alpha - 2)(\alpha - 1)^2} \qquad (8)$$

But, the quartiles $\mu_q(x)$ being defined as follow:

$$P(\mu_q(x)) = qP(x) \qquad 0<q<1 \qquad (9)$$

$$\Rightarrow \mu_q(x) = q^{-\frac{1}{\alpha}} x + c\,(q^{-\frac{1}{\alpha}} - 1) \qquad (10)$$

The median of the goods $\geq x$ *is obtained* by taking $q = \frac{1}{2}$

If one can't get the standard deviation; when $0 < \alpha \leq 2$, the *mean gap* e always exists:

$$e = \frac{\int_{x_o}^{\infty} \left| x - \mu_{1/2} \right| dF(x)}{P(x)} = \frac{\alpha}{\alpha - 1}(x_o + c)(2^{\frac{1}{\alpha}} - 1) \qquad (11)$$

The function of concentration from this law will be:

$$Q(p) = (1 + \frac{c}{M})p^{\frac{1}{\beta}} - (\frac{c}{M})p \qquad (12)$$

Meaning that *Q(p)* is a barycentre of p and $p^{\frac{1}{\beta}}$ weighted by $(-\frac{c}{M})$ and $(1+\frac{c}{M})$,

N.B. The power function will result from Pareto's law by taking:

$x_m = -c$ et $\gamma = -\alpha$ **this gives**

$$f(x) = \begin{cases} \gamma \dfrac{(x_m - x)^{\gamma}}{(x_m - x_o)^{\gamma+1}} & \forall x \in [x_o, x_m] \\ 0 & \text{if not} \end{cases} \qquad (13)$$

C) Difficulties in Pareto's method:

Let us point out that, since Pareto one estimates his law by transforming it to a logarithmic function that is certainly due to the genesis of this one.

From the following form

$$P(x) = \left(\frac{c}{x}\right)^{\alpha} \qquad (14)$$

The Paret's method would give

$$\ln P(x) = -\alpha \ln(x) + c' + \varepsilon(x) \qquad (15)$$

Since this last has a linear appearance, one is allowed to apply the method of least squares to estimate it. But, this way is not without violating some principles which governs the use of this method. More, the cumuli of *P(x)* can distort the principle of the independence of the $\varepsilon(x)$ errors and its stability (homoscedasticity). Add to this problem of principle, the difficulties that we will meet if the law counts more than two parameters as the general form. One does not know how Pareto had made to estimate the third parameter. In all cases, the parameter α, he regarded as the indicator of the inequality of the incomes wasn't one. It is the dual of this (β) which measures this phenomenon.

D) How to fit the laws of Pareto's type

Let F be a distribution function of Pareto's type with □> 1, its conditional average will be a linear function

$$M(x) = \frac{\int_{x_0}^{\infty} t dF(t)}{P(x)} = \beta x + \mu \qquad \beta \geq 1 \quad \mu = c(\beta - 1) \qquad (16)$$

Conversely, if a conditional average of a law is linear:

$$M(x) = ax+b \qquad (17)$$

It is demonstrated (BARBUT M., 1989) that, this one is

- **Pareto's type** *if* $a > 1$ (18)
- ***Exponential*** **one if** $a = 1$ *and* $b \neq 0$ (19)
- ***Power*** **one** **if** $0 < a < 1$ *et* $b \geq 0$ (20)

Interpretation: the conditional average *M(x)* expresses clearly the variation of goods between an individual and those which are richer than him. On average, they have β times of goods than him.

As we see, carrying out adjustments (fitting) with the laws mentioned is very simple. Here, we can apply the method of least squares without violating any principle. This method of fitting the Pareto's law, though simple is so robust.

E) COMPARISON OF the LAWS ACCORDING TO THEIR ASYMPTOTIC BEHAVIOR

Let us suppose *x* very large, and *ln(x)* = *u*,

The Pareto's law :

$$f(x) = \frac{k}{x^{\alpha+1}} \qquad \alpha > 0 \qquad (21)$$

$$\Rightarrow |\ln f(x)| = (\alpha + 1)u + k' \qquad (22)$$

The exponential:

$$f(x) = ae^{\frac{-x}{\mu}} \qquad \mu > 0 \quad a > 1 \qquad (23)$$

$$\Rightarrow |\ln f(x)| = a' + \frac{e^u}{\mu}$$

Log normal law (Gibrat) :

$$f(x) = \frac{1}{x} e^{-b(\ln x)^2} \qquad x > 0 \quad b > 0 \tag{24}$$

$$\Rightarrow |\ln f(x)| = u + bu^2$$

We see, from the comparison of these three laws, that Pareto is the less uneven (concentrate) distribution if one considers the higher level of goods, the law of Gibrat follows and then the exponential one.

G) Application: inequality measure of a distribution

One will apply these results to study the households' income distribution using data from Senegal. We differentiate agricultural incomes from the others. After calculating the functional expression of inequalities by the mean of estimated parameters of this law, we will add the indices of GINI (*G*) and Coefficients of variation (*V*):

$$G = 1 - 2\int_0^1 Q(p)dp = \frac{\beta - 1}{\beta + 1}\left(1 + \frac{c}{M}\right) \tag{25}$$

$$V = \left(1 + \frac{c}{M}\right)\sqrt{\frac{(\beta - 1)^2}{2\beta - \beta^2}} \tag{26}$$

It is obvious that □is the coefficient of the inequality of each distributions generated by a Pareto's law, more □ move away from the unit, plus the distribution is uneven

Table I : Regional incomes distribution according to the number of households in 91/92								
Region	N°	RVNU* 1	%	NMGE 2	%	Rev/men 3	MTS 4	Cfdd 5
Dakar	1	326 158 658	49,23%	198 893	23,75%	1639,87	1639,87	1,00
St.louis	4	64 821 624	9,78%	82 022	9,79%	790,30	1391,81	1,76
Thiès	7	70 965 958	10,71%	111 005	13,26%	639,30	1178,67	1,84
Louga	8	32 811 134	4,95%	54 585	6,52%	601,10	1108,07	1,84
Diourbel	3	39 541 108	5,97%	72 242	8,63%	547,34	1029,98	1,88
Tamba	5	21 838 798	3,30%	48 177	5,75%	453,30	980,97	2,16
Kaolack	6	40 315 277	6,09%	90 916	10,86%	443,43	906,68	2,04
Ziguinchor	2	23 030 517	3,48%	53 335	6,37%	431,81	871,07	2,02
Kolda	10	24 747 856	3,74%	67 737	8,09%	365,35	827,09	2,26
Fatick	9	18 296 947	2,76%	58 495	6,99%	312,80	791,17	2,53
Senegal		662 527 877	100,0%	837 407	100,0%	791,17		
Extreme ratio :						5,24		
supreme share :			49,2%	for	23,8%			

Lecture : Column 1 : Incomes per region, column 2 : number of households per region; column3 : Ratio of column 1 to column 2 ; column 4 : Estimated M(x) ; Column 5 : ratio of column4 to column3

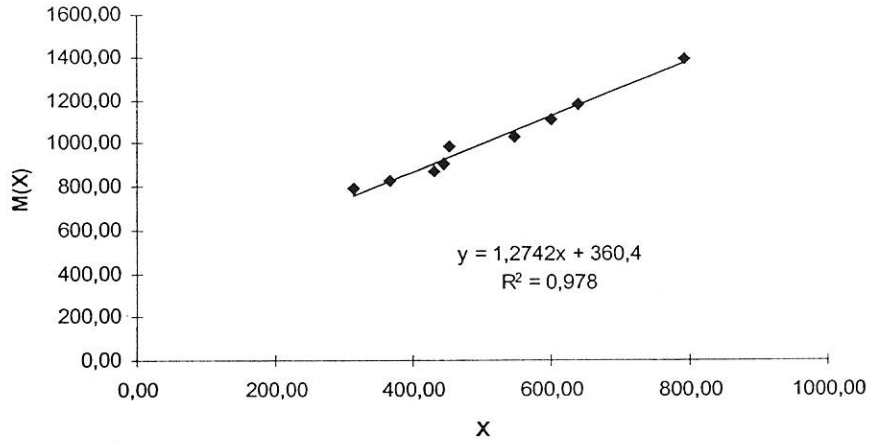
Household = Group of persons living together, sharing all or part of their resources and recognizing the authority of one among them: the household chief.

* Incomes in thousands of West African money in 1991 (Actually 1$US = 551,755 FCFA).

Table I shows the distribution of income according to the number of households. It is obvious there that, half of the richness (supreme share) is held by less than one quarter (23.8%) of the households. Only those of Dakar reach the average of the income (1.639.870 Fcfa), they gain five times (5.24) more than those of Fatick. On average, the differences between households according to regions of Senegal, are at least, 30% (Gini= 31.4%). This is reflected by the Lorenz curves

An indicator, in any field, generally does not have clear interpretation than when it intervenes like parameter of a theoretical distribution or "law" of reality observed. It is the case, for example, of the average or the standard deviation which are the parameters of Laplace-Gauss's distributions.

Graphic I : Ditribution of income

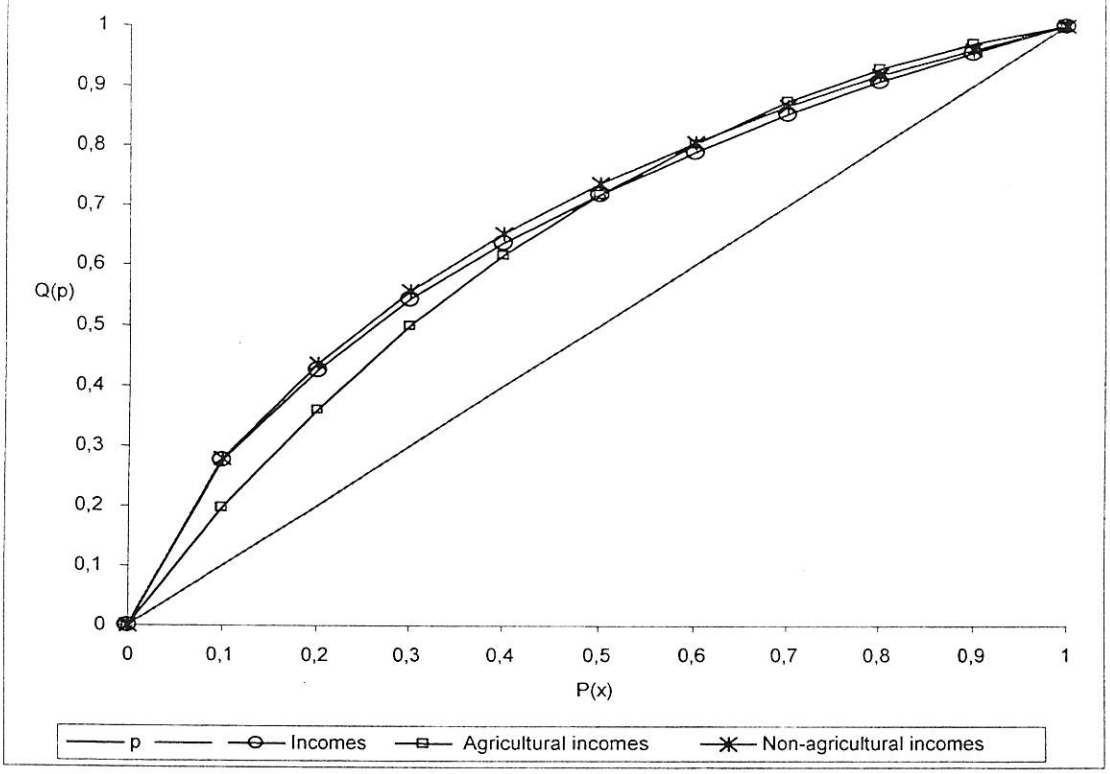


$$M(x) = 1.27\,x + 360.4 + \varepsilon(x)$$

$$P(x) = \left(\frac{1652.45}{1314.38 + x} \right)^{4.65}$$

$$Q(p) = -1.66\,p + 2.66\,p^{0.78}$$

Graphic I B: The income distribution according to the number of households and sources.91/92



Tables II and III constitute the agricultural and non agricultural income distributions. It results from this decomposition that the distribution of agricultural income is less uneven. It reduces the total disparity of the income distribution.

Table II : Regional distribution of agricultural income according to the number of households in 91/92

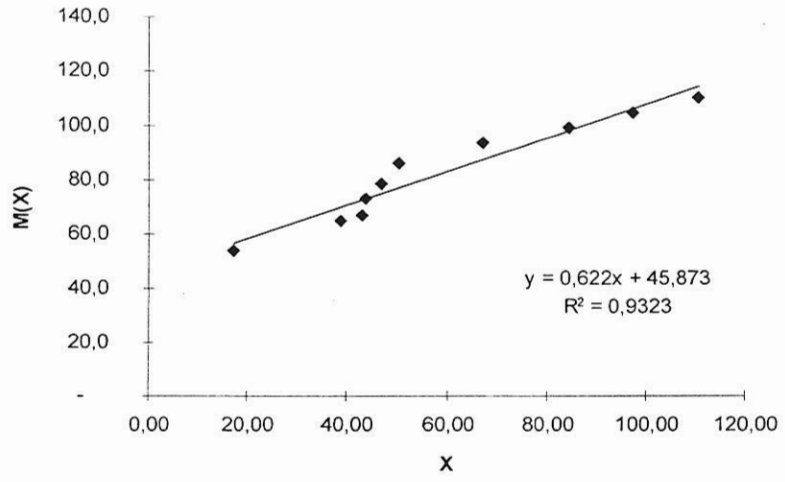
Region	N°	RVNUG* 1	%	NMGE 2	%	Rvnug/men 3	MTS 4	Cfdd 5
Kaolack	6	10040844	22,39%	90916	10,86%	110,44	110,4	1,00
Kolda	10	6599631	14,72%	67737	8,09%	97,43	104,9	1,08
Fatick	9	4936570	11,01%	58495	6,99%	84,39	99,4	1,18
Tamba	5	3243363	7,23%	48177	5,75%	67,32	93,5	1,39
Louga	8	2747612	6,13%	54585	6,52%	50,34	86,2	1,71
Diourbel	3	3396362	7,57%	72242	8,63%	47,01	79,0	1,68
St.louis	4	3581505	7,99%	82022	9,79%	43,67	72,9	1,67
Thiès	7	4777221	10,65%	111005	13,26%	43,04	67,2	1,56
Ziguinchor	2	2066392	4,61%	53335	6,37%	38,74	64,8	1,67
Dakar	1	3454398	7,70%	198893	23,75%	17,37	53,6	3,08
Sénégal		44843898	100,0%	837407	100,0%	53,5509		
Rapport extrême :						6,36		
Part suprême :			55,3%	pour	31,7%			

Lecture : **Column 1** : Agricultural Incomes per region, **column 2** : number of households per region; **column3** : Ratio of column **1** to column **2** ; **column 4** : Estimated ***M(x)*** ; **Column 5** : ratio of column**4** to column**3**

Household = Group of persons living together, sharing all or part of their resources and recognizing the authority of one among them: the household chief.

* Incomes in thousands of West African money in 1991 (Actually 1$US = 551,755 FCFA).

Graphic II: Ditribution of agricultural income



$$M(x) = 0.62\,x + 45.87 + \varepsilon(x)$$

$$P(x) = \left(\frac{121.37 - x}{109.03}\right)^{1.65}$$

$$Q(p) = 2.27\,p - 1.27\,p^{1.61}$$

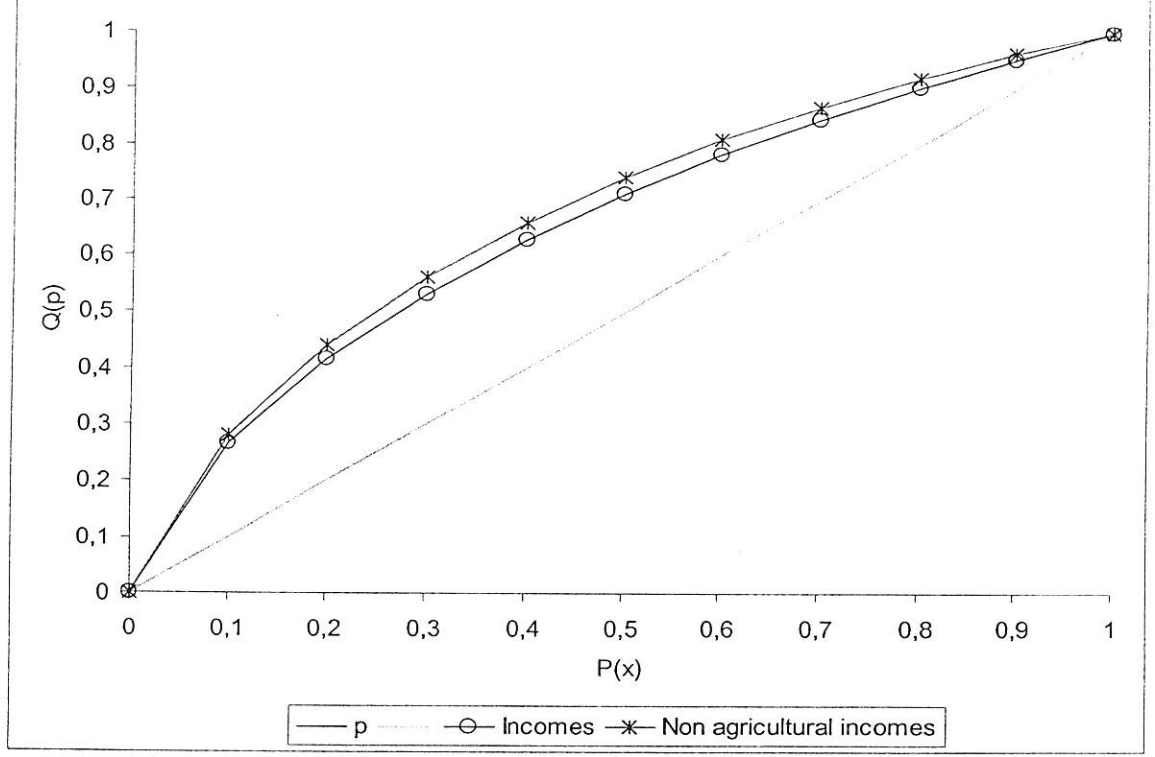


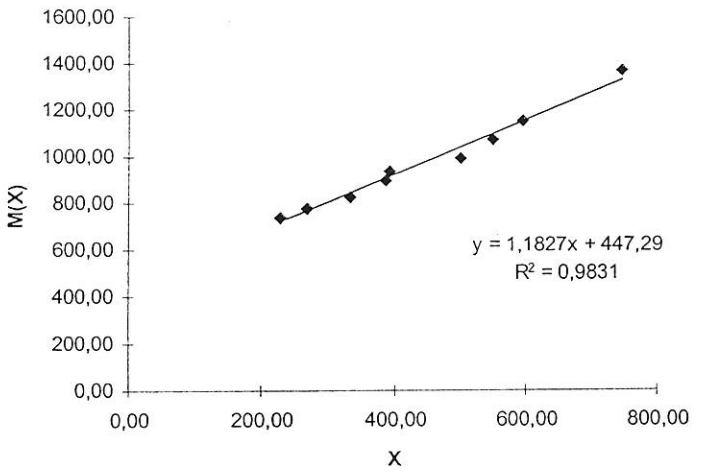
Table III: Regional distribution of non-agricultural income according to the number of households. 90/91

Region	N°	RVNUNG 1	%	NMGE 2	%	Revng/men 3	MTS 4	Cfdd 5
Dakar	1	322 704 260	52,24%	198893	23,75%	1622,5	1622,50	1,00
St.louis	4	61 240 119	9,91%	82022	9,79%	746,63	1366,76	1,83
Thiès	7	66 188 737	10,72%	111005	13,26%	596,27	1148,53	1,93
Louga	8	30 063 522	4,87%	54585	6,52%	550,77	1075,46	1,95
Diourbel	3	36 144 746	5,85%	72242	8,63%	500,33	995,36	1,99
Ziguinchor	2	20 964 125	3,39%	53335	6,37%	393,07	939,21	2,39
Tamba	5	18 595 435	3,01%	48177	5,75%	385,98	896,24	2,32
Kaolack	6	30 274 433	4,90%	90916	10,86%	332,99	824,24	2,48
Kolda	10	18 148 225	2,94%	67737	8,09%	267,92	775,86	2,90
Fatick	9	13 360 377	2,16%	58495	6,99%	228,40	737,62	3,23
Senegal		617 683 979	100,0%	837407	100,0%	737,61502		
Rapport extrême :							7,1	
Part suprême :				62,16%	pour	33,5%		

Lecture : Column 1 : Agricultural Incomes per region, column 2 : number of households per region; column3 : Ratio of column 1 to column 2 ; column 4 : Estimated M(x) ; Column 5 : ratio of column4 to column3

Household = Group of persons living together, sharing all or part of their resources and recognizing the authority of one among

Graphic III: Distribution of non agricultural income



$$M(x) = 1.18x + 447.29 + \varepsilon(x)$$

$$P(x) = \left(\frac{2693.36}{2447.88 + x} \right)^{6.47}$$

$$Q(p) = -3.32\,p + 4.32\,p^{0.85}$$

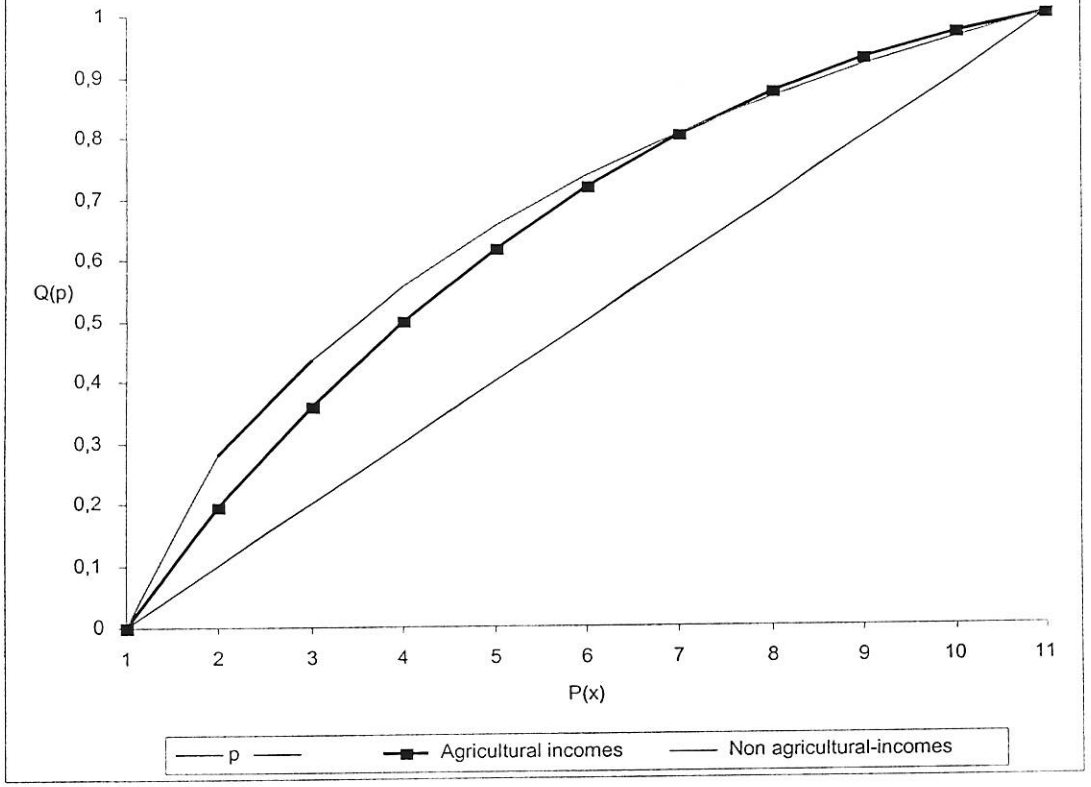


Tableau IV : Theoretical and empirical indices of the income inequality			
Indices and coefficients	Agricultural incomes	Non agricultural incomes	All Incomes
Gini : G			
Theoretical	0,295	0,362	0,321
Empirical	0,313	0,358	0,314
Coefficient of variation *V* :			
Theoretical	0,341	0,445	0,431
Empirical	0,316	0,406	0,369

Tableau IV permits to verify if the observed distribution correspond to the theoretical one generated by the law. Thus we can compare the two indices (*Gini* and Coefficient of variation) calculated from the two distributions.

Conclusion

As we have seen, the use of Pareto's type laws brings much to the study and the understanding of socio-economic phenomenon, notably their distribution. They permit to represent and interpret better these phenomenons. The comparison of distributions becomes more acceptable with this family of functions. As we seen, one can conclude from the application that the

the distribution of the richness. Especially for the countries stripped in resources. It is necessary to reduce well the inequality distribution of the goods and public services, to combat the poverty which, in more of being a material situation, is also a state of mind, the feeling of injustice.

Atkinson A. B. (1970) "On the measurement of inequality" in *Journal of Economic Theory* 2:244-63,

AIGNER Dennis J. And Arthur S. Goldberger (1970) "Estimation of Pareto's law from grouped observations»; Journal *of the American Statistical Association* 65:712-23

Alker H.and B.M.Russett, (1964) "On measuring inequality" *Behaviorial Sciencs* 9:207-18,

ALLISON P.D. "Measures of inequality", American *sociological review* 43, pp.865-880 19

BARBUT M. (1988)."De bons et de moins bons usages des distributions paretiennes en analyse de données" ; *Histoire et mesure*, III, pp. 111-128

BARBUT.M (1989)." Distribution de type Paretien et représentation des inégalités" ; *Mathématiques informatiques et sciences humaines*, 106, pp.53-69

BARBUT M. (1990)."Introduction à l'analyse mathématique des inégalités" ; *Archive de psychologie*, 58, pp. 91-114

BARTELS, C.P.A. and P.Nijkamp (1976) "An empirical welfare approach to regional income distributions". *Socio-Economic-Planning Sciences* 10:117-28

CHAMPERNOWNE, D. G. (1974) "A comparison of measure of inequality of income distribution" ; *The economic journal*, 84, pp. 787-816

Coulter P.B. (1989). *Measuring inequality*; West view Press, San Francisco.

Creedy john. (1977) " The principle of the transfers and the variance of the logarithms " *Oxford Bulletin of Economics and statistics* 39 : 152-8

Dalton Hugh (1920) "The measurement of inequality of incomes" *Economic Journal* 30 : 348-61

EYTAN M. (1962) "Quelques propriétés de la loi de Pareto et leurs incidences en sciences humaines" ; *Mathématiques et sciences humaines*, 8, pp. 21-26

FIELDS G.S., J.C.H.Fei (1978) "On inequality comparisons" ; *Econometrica*,46, pp. 303-316

FOSTER James E. (1985) "Inequality measurement", *Proceeding of symposia in applied mathematics*, vol. 83,

GASTWIRTH J. L. (1971) "A general definition of the Lorenz curve" ; *Econometrica* ,39,1971, pp.1037-1039

GASTWIRTH J. L (1972) "The estimation of the Lorenz curve and the Gini index" ; *The Review of Economics and Statistics* 54:306-16

KAKAWANI Nanak (1990) *Large sample distribution of several inequality measures with application to Côte d'ivoire*, Living standard measurement study, Working paper, n° 61, The world bank, Washington D.C.,

KAKAWANI Nanak and N.Podder. (1973). "On the estimation of Lorenz Curves from grouped observations" *International Economic Review* 14: 278-92

MANDELBROT BENOIT, "Formes nouvelles du hasard dans les sciences", *Économie Appliquée, vol. 26, 1973, 307-319*

PARETO. W. (1896) *La courbe de la répartition de la richesse.* Lausanne, Faculté de Droit /Librairie Viret-Genton.

PYATT G. (1976)."On the interpretation and disaggregation of Gini coefficient", *Economic journal*, 86, 1976, pp 243-255

Ra0 C.R. (1965). Linear statistical inference and its applications New York : John Wiley & Sons

RAY James Lee and J.David Singer (1973) "Measuring the concentration of power in the international system " *Sociological Method and Research* 1 : 403-37 ,

SALL Mamadou-Y. (1997) *Mesure de l'inégalité dans l'Éducation : le Cas du Sénégal* Atelier National de Reproduction de Thèses (A.N.R.T.) Université-Lille3, Lille .

SALL Mamadou-Y. *(200)* "Evaluating the cost of wastage rates: The case of Gaston BERGER University of Senegal" *Higher education policy, , 16, (333-349)*

SEN Amartya K. (1973), *On economic inequality*, New York, Norton

SENEGAL (1984) *Situation économique du Sénégal 1984*, Direction de la statistique, Ministère d l'économie et des finances

SENEGAL (1988) *Situation économique 1988*, Direction de la prévision et de la statistic Ministère de l'économie et des finances

TRACHEEN Ahmed *Développement et dynamique de l'inégalité*, Paris, Economica

ZAPF Wolfgang (1975) "Les systèmes d'indicateurs sociaux : approches et problèmes" dans *Revue internationale des sciences sociales*, XXVII n°3 pp.507-530 Unesco

ان الهدف الرئيسي للمجلة الاحصائية المصرية هو نشر اضافات الاحصاء سواء في النواحي النظرية او التطبيقية .

تقبل المجلة الاحصائية المصرية للنشر الابحاث التي تحتوي علي وصف الاساليب او طرق احصائية جديدة يكون من المتوقع نفعها في مختلف المجالات العلمية ، كما تقبل المجلة توصيف مختصر للمشاكل غير المحلوله بالاضافه الي الملاحظات الفنية والخطابات الموجهة للنشر.

الاشتركات

١- داخل جمهورية مصر العربية ٢٠ جنية مصرية في العام (عدد يونيو - وديسمبر) . اما بالنسبة للاعداد التي صدرت في الاعوام السابقة فيمكن ارسال المتوفر منها مقابل ١٠ جنيهات مصرية للمجلد (المجلد يشمل العددين الصادرين في العام) .

٢- خارج جمهورية مصر العربية ١٠٠ دولار (مائة دولار او ما يعادلها) في العام للفرد ١٢٠ دولار للمؤسسات اما الاعداد التي صدرت في الاعوام السابقة فيمكن ارسال المتوفر منها مقابل ٥٠ دولار (خمسون دولار او ما يعادلها) للمجلد .

الاشتركات وطلب الاعداد السابقة وتغيير العناوين وكافة المراسلات تكون علي العنوان التالي:

السيد الاستاذ الدكتور / عميد معهد الدراسات والبحوث الاحصائية – جامعة القاهرة.
ص . ب : ١٠١٧ – ٥ شارع احمد زويل – الاورمان – الجيزة
فاكس :٧٤٨٢٥٣٣