



HAL
open science

La loi de Zipf dans la science régionale : entre anciennes controverses et nouvelles perspectives

Alexandra Schaffar

► **To cite this version:**

Alexandra Schaffar. La loi de Zipf dans la science régionale : entre anciennes controverses et nouvelles perspectives. *Cybergeo: Revue européenne de géographie / European journal of geography*, 2009, 450, pp.en ligne. 10.4000/cybergeo.22200 . hal-00924606

HAL Id: hal-00924606

<https://hal.science/hal-00924606>

Submitted on 15 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



La loi de Zipf dans la science régionale : entre anciennes controverses et nouvelles perspectives

Zipf's law in regional science: old controversies and new developments

Alexandra Schaffar



Édition électronique

URL : <https://journals.openedition.org/cybergeog/22200>

DOI : 10.4000/cybergeog.22200

ISSN : 1278-3366

Éditeur

UMR 8504 Géographie-cités

Ce document vous est offert par Bibliothèques de l'Université de La Réunion



Référence électronique

Alexandra Schaffar, « La loi de Zipf dans la science régionale : entre anciennes controverses et nouvelles perspectives », *Cybergeog: European Journal of Geography* [En ligne], Systèmes, Modélisation, Géostatistiques, document 450, mis en ligne le 27 avril 2009, consulté le 15 novembre 2022. URL : <http://journals.openedition.org/cybergeog/22200> ; DOI : <https://doi.org/10.4000/cybergeog.22200>



Creative Commons - Attribution 4.0 International - CC BY 4.0

<https://creativecommons.org/licenses/by/4.0/>

La loi de Zipf dans la science régionale : entre anciennes controverses et nouvelles perspectives

Zipf's law in regional science: old controversies and new developments

Alexandra Schaffar

LIM / Université de La Réunion
UFR Sciences de l'Homme et de l'Environnement, Université de La Réunion,
117, Avenue Général Ailleret, 97430 Tampon, La Réunion.
LEAD / Université de Toulon

schaffar@univ-reunion.fr

Résumé :

La loi de Zipf conditionne une grande partie des travaux scientifiques sur les hiérarchies urbaines. Par un large rappel des principales contributions théoriques et empiriques sur ce sujet depuis une trentaine d'années, l'objectif de ce papier est de mettre en évidence les controverses qui accompagnent l'utilisation des modèles rang-taille en science régionale. Celles-ci sont essentiellement d'ordre méthodologique, liées au bornage de l'échantillon des villes et à la méthode d'estimation du coefficient de hiérarchisation. Elles mettent, cependant, en lumière les risques d'erreur d'interprétation liés à l'utilisation de ces modèles. L'introduction plus ou moins récente des modèles rang-taille dans une approche dynamique permet d'ouvrir des nouvelles perspectives de recherche, en faisant de la loi de Zipf, l'aboutissement d'un processus de croissance urbaine complexe.

Mots-clés : loi de Zipf, modèles rang-taille, croissance urbaine, hiérarchies urbaines.

Abstract:

Zipf's law is one of the main features in regional sciences, when it comes in studying urban systems' organization. By recalling on the most representative studies on this issue, over the last thirty years, this paper aims to focus on the main controversies that characterize the rank-size models, that is, the use of a truncated city-size distribution and the method of estimating the Pareto exponent. The recent introduction of the rank-size models in a more dynamic approach opens up new perspectives in understanding the Zipf law, as the latter becomes the outcome of an urban growth process.

Key-words: Zipf law, rank-size distributions, urban growth, urban hierarchies.

1 Introduction

En 1949, en observant les lois puissance qui caractérisent certains phénomènes économiques et sociaux, Schumpeter écrivait de façon prémonitoire : *« Il y a vraiment très peu de chercheurs qui saisissent les possibilités scientifiques qu'ouvre l'existence de telles régularités statistiques. De surcroît, aucun d'entre eux ne semble réaliser que leur interprétation peut être à l'origine d'une théorie économique et sociale complètement novatrice »* (Schumpeter, 1949, p.155).

La loi de Zipf figure parmi ces régularités statistiques fortes. Elle caractérise la distribution des tailles des villes et permet de résumer en une seule valeur, celle du coefficient de hiérarchisation, les informations qui régissent la formation et l'évolution des hiérarchies urbaines d'un pays ou d'une région. Qualifiée jadis de *« mystère urbain »* par Krugman (1996, p.40), la loi de Zipf est, selon Gabaix et Ioannides (2004), *« un des faits les plus frappants en sciences sociales en général »* (Gabaix et Ioannides, 2004, p.739). Ce papier propose une relecture critique des différents travaux sur la loi de Zipf en science régionale, en mettant fortement l'accent sur plusieurs points de controverse que suscite cette loi.

La première section propose un rappel condensé de l'application des modèles rang-taille dans l'étude des hiérarchies urbaines des différents pays et régions, durant les trente dernières années. Ces travaux ont cherché à utiliser l'information procurée par le calcul du coefficient de hiérarchisation en deux sens : d'une part, proposer des comparaisons internationales de différents systèmes urbains et, d'autre part, mettre en relation, dans un même pays, l'évolution de sa structuration urbaine et de son développement économique.

La seconde section aborde le premier point de contestation des modèles rang-taille, selon lequel, le choix arbitraire d'une taille minimale des villes considérées, conduisant à une distribution *« tronquée »*, ne permet pas d'étudier la forme de la distribution réelle des villes. Or, cette dernière ne suit pas nécessairement une loi de Pareto, ce qui signifierait tout simplement que la loi de Zipf n'existe pas mais est une construction fictive, liée aux pratiques d'échantillonnage.

La troisième section porte sur le deuxième point de contestation des modèles rang-taille, qui concerne le calcul du coefficient de hiérarchisation. De nombreuses études montrent le fort biais que comportent les méthodes d'estimation les plus populaires, conduisant à des erreurs importantes d'interprétation de la structuration d'un système de villes.

Enfin, la quatrième section explore l'intégration de la loi de Zipf dans une approche dynamique, qui cherche à mettre en relation la croissance urbaine avec la distribution rang-taille des villes. Dans ce contexte, l'évolution du coefficient de hiérarchisation et la loi de Zipf, apparaissent comme des aboutissements de processus démographiques dont la nature est régie par des modèles de croissance urbaine spécifiques.

2 La loi de Zipf dans la science régionale

2.1 Les origines du modèle rang-taille

En 1913, Auerbach (1913) fait remarquer, en premier, la constance singulière de la valeur du produit de la population d'une ville multipliée par son rang, au sein d'un système urbain donné. Dans ce schéma, la population de la deuxième plus grande ville représente à peu près la moitié de la population de la première, celle de la troisième ville un tiers de la première et, de façon plus générale, la population de la ville de rang n , une proportion $1/n$ de la population de la plus grande ville. Il établit, ainsi, un premier modèle de hiérarchies urbaines où le nombre des villes est en relation inverse avec leur taille respective. Malgré sa primauté dans

l'étude des systèmes urbains, l'intuition d'Auerbach n'est pas fortuite. Elle s'intègre, au contraire, dans un ensemble de contributions de chercheurs en sciences sociales, tels que Pareto (1895) et Singer (1936) en économie, Goodrich (1928) en géographie, Bradford (1934) en scientométrie, qui s'inspirent de la prolifération exceptionnelle des travaux des mathématiciens de cette époque sur les lois de distribution statistique (Pearson, 1895, Gosset, 1909, Yule, 1911, Lotka, 1926, Weibull, 1939).

Au lendemain de la deuxième guerre mondiale, Zipf (1949) poursuit ces recherches, en appliquant initialement la loi de Pareto sur les sciences du langage, puis étend son travail sur d'autres domaines, tels que les systèmes de villes. En s'appuyant sur la base de données de Wickens (1921), il étudie, alors, la distribution rang taille de 256 villes australiennes de plus de 3000 habitants en 1921 (Zipf, 1949, p.137).

Selon Zipf, si X est la variable qui associe à chaque ville sa population, X suit une loi de Pareto, dont la fonction densité est de la forme :

$$f(x) = Cx^{-a} \quad (1)$$

où C est une constante qui dépend de la taille de la plus grande ville du pays et a est un coefficient de hiérarchisation. Si $\Pr(X > x)$ est la probabilité qu'une ville ait une population supérieure à x , la fonction de répartition complémentaire associée à X est :

$$F_{>}(x) = \Pr(X > x) = \int_x^{\infty} f(t)dt = \int_x^{\infty} Ct^{-a} dt$$

$$F_{>}(x) = \frac{Cx^{-a+1}}{a-1} \text{ avec } a > 1 \quad (2)$$

Lorsque les villes sont rangées, selon leur taille, de la façon suivante : $x_1 > x_2 > x_3 > \dots > x_i > \dots > x_n$, avec $r(x)$ le rang de la taille de la ville x , on obtient :

$$\Pr(X > x) = \frac{r(x)}{n} \quad (3)$$

En tenant compte de la relation 2, on peut alors écrire l'égalité suivante :

$$\frac{r(x)}{n} = \frac{C}{a-1} x^{-a+1} \text{ avec } a \neq 1 \quad (4)$$

En posant $k = \frac{Cn}{a-1}$ et $\beta = a-1$ (avec $\beta > 0$), on obtient :

$$r(x) = k \cdot x^{-\beta} \quad (5)$$

avec k un paramètre qui dépend de la taille de la plus grande ville et β un coefficient de hiérarchisation, appelé communément coefficient de Pareto. La relation linéaire entre les logarithmes de rang et de taille correspond à la version la plus connue de la loi rang taille qui stipule que le rang d'une ville donnée est inversement proportionnel à sa taille :

$$\ln r(x) = -\beta \ln x + \ln k \quad (6)$$

Lorsque β est égal à 1, on obtient la loi de Zipf. Ce dernier avance une vague justification de la régularité de cette valeur dans la nature humaine qui conditionne les choix de vie et de localisation des individus. Néanmoins, dans son étude initiale sur l'Australie, Zipf

trouve que la distribution des villes australiennes dévie d'une distribution de Pareto, notamment à cause de l'allure de sa queue, ce qui est, selon lui, dû au biais de l'échantillonnage. Ainsi, il considère qu'en retenant, dans un même échantillon, à la fois les communautés urbaines et rurales, on obtient une distribution composite, issue de la combinaison de deux distributions distinctes (Zipf, 1949, p.423). Il introduit, par là, un ensemble d'interrogations méthodologiques qui rejoignent l'approche fractale des distributions rang taille, évoquée par Bradford (1934). Cette question sera traitée par Mandelbrot (1959), qui proposera une généralisation de la loi rang taille (Feuerverger et Hall, 1999).

Quinze ans après les écrits de Zipf, la loi rang taille est introduite en science régionale par les articles fondateurs de Simon (1955) et de Beckmann (1958) et suscite, à partir des années quatre-vingts, un développement prolifique de travaux sur les hiérarchies urbaines (en Annexe 1, un tableau récapitulatif, mais non pas exhaustif des études empiriques récentes les plus représentatives sur la loi de Zipf).

2.2 Comparaisons internationales : le coefficient de Pareto, une mesure des différences des hiérarchies urbaines nationales

La loi rang taille a servi à la comparaison des hiérarchies urbaines nationales, par le biais de la comparaison des coefficients de Pareto dans différents pays et régions. Ainsi, dès 1936, Singer examine l'application d'une relation rang taille sur la distribution des villes de plus de 2000 habitants de sept pays, tandis que Allen (1954) entame une démarche identique sur un échantillon de 58 pays. Rosen et Resnick (1980) s'appuient, eux, sur des échantillons de villes de plus de 100000 habitants de 44 pays différents. Ils obtiennent des coefficients de hiérarchisation dans un intervalle [0,81 ; 1,96], avec 75% des pays affichant une valeur absolue de l'exposant supérieure à 1 (la moyenne des β est égale à 1,13, les deux valeurs extrêmes étant celles du Maroc, 0,809 et de l'Australie, 1,963). Ceci les conduit à conclure que la distribution rang taille des villes est plus égalitaire que ce que la loi Zipf laisserait supposer.

Dans sa revue de la littérature empirique sur la loi rang taille, de 1913 à 1980, et en s'appuyant sur l'étude de Rosen et Resnick, Carroll (1982) montre que les distributions de villes dont le coefficient de Pareto est égal à 1 ne représentent qu'un cas parmi d'autres hiérarchies urbaines existantes. Plutôt que de tester la validité de la loi de Zipf, il devient, ainsi, beaucoup plus intéressant d'étudier les pentes des différentes distributions rang taille afin de pouvoir comparer l'organisation des différents systèmes urbains dans un *continuum* entre deux configurations extrêmes, la primatie absolue d'une part, l'égalité parfaite des tailles des villes, d'autre part. Lorsque le coefficient de hiérarchisation β est inférieur à 1, l'effet agglomération est renforcé et les villes de grande taille ont un poids plus important que dans une distribution qui suit la loi de Zipf. A l'inverse, si β est supérieur à 1, on est en présence d'un espace polycentrique où le nombre des villes moyennes est plus important. Selon Gabaix et Ioannides (2004), le débat scientifique actuel devrait, donc, davantage se focaliser sur le degré d'adéquation des différentes distributions avec la loi de Zipf, plutôt que sur sa validité au sens strict du terme.

Dans un large travail de comparaison internationale, permettant de tester la relation entre niveau de développement économique et nature des hiérarchies urbaines Moriconi-Ebrard (1993) s'appuie sur le modèle de Lotka (qui met en relation la taille en fonction du rang) pour proposer une analyse de la distribution des villes de plus de 10000 habitants dans 78 pays, allant des pays industrialisés aux pays en développement, entre 1950 et 1980. L'écart des coefficients nationaux, qui se situent entre 0,73 et 1,38, montre une certaine différenciation des pays, mais le niveau de développement n'apparaît pas comme un critère significatif.

Moriconi-Ebrard trouve ainsi des coefficients très proches pour le groupe des pays développés et celui des pays en développement. Un véritable clivage apparaît, à l'inverse, entre l'Ancien (Europe, Asie, Afrique) et le Nouveau (Amérique, Océanie) Monde, ce qui conduit à émettre l'hypothèse que les différences internationales dans la distribution rang-taille des villes reposent sur des facteurs historiques. Cependant, dans une étude plus récente sur des échantillons de villes de 73 pays, Soo (2005) trouve, à l'inverse, un coefficient de hiérarchisation plus élevé pour les pays à revenu élevé (Amérique du Nord, Europe, Océanie) que pour les pays à revenu faible ou intermédiaire (Asie, Amérique du Sud, Afrique), ce qui signifie que ces derniers sont caractérisés par une structure plus hiérarchique que les premiers. Il conclut, alors, que le niveau du développement économique a une influence déterminante sur les hiérarchies urbaines.

Nitsch (2005) propose la comparaison la plus complète des différentes estimations du coefficient de hiérarchisation, en s'appuyant sur une méta-analyse qui porte sur 29 études et 515 estimations. Selon Nitsch (2005), l'ensemble des estimations du coefficient de Pareto varie dans un intervalle [0,49 ; 1,96], avec une moyenne de 1,09. La majorité des estimations (62%) dépasse l'unité, tandis que un tiers se trouve en dehors de l'intervalle [0,8 ; 1,2]. Il montre que la définition de la ville joue un rôle fondamental dans la valeur estimée du coefficient et met, en évidence, comme Rosen et Resnick (1980) et Alperovich (1993), le fait que les échantillons de villes administratives présentent des distributions plus égalitaires que les échantillons des agglomérations, sans que cela n'affecte, cependant, l'ordre des comparaisons établi.

2.3 Comparaisons diachroniques : le coefficient de Pareto, une mesure du niveau de développement économique d'un pays

Une deuxième série de travaux met l'accent sur les comparaisons diachroniques afin d'examiner l'évolution du coefficient de hiérarchisation. A l'instar du travail pionnier de Williamson (1962), ces travaux cherchent à lier le niveau de développement d'un pays à son niveau d'urbanisation. En actualisant le travail de Berry (1964) qui, dans une comparaison de 37 pays, trouvait une corrélation entre ce coefficient et le niveau de développement, Parr (1985) engage une hypothèse extrêmement intéressante selon laquelle, dans de nombreux pays, ce coefficient suit, dans le temps, une courbe en U, ce qui implique la succession d'une phase de concentration puis de déconcentration urbaine.

A partir des résultats obtenus par la comparaison des hiérarchies urbaines de douze pays, Parr (1985) intègre les modifications des valeurs du coefficient de hiérarchisation dans un schéma d'évolution de la concentration urbaine qui peut être décliné en cinq étapes (Catin et Van Huffel, 2004) : une étape préurbaine, caractérisée par la faiblesse des infrastructures de transport et l'absence d'économies d'agglomération ; une étape de spécialisation urbaine où les infrastructures publiques se développent et certaines villes se spécialisent dans des productions particulières, d'où l'apparition de certaines économies d'échelle ; une étape de consolidation urbaine, durant laquelle les mécanismes précédents s'accroissent et la concentration urbaine augmente ; une étape de transformation urbaine qui correspond au seuil de retournement de la courbe en U, sous l'effet de l'apparition de déséconomies d'agglomération ; puis, enfin, une étape de dispersion urbaine où l'amélioration des infrastructures de transport interrégional permet une relocalisation des activités de production dans les régions et villes périphériques.

Plusieurs auteurs ont cherché à tester l'hypothèse de la courbe en U. Ainsi, Lepetit (1990) évalue l'évolution du coefficient de hiérarchisation en France entre le 18^{ème} et 19^{ème} siècle, mais trouve des changements très lents de la distribution rang taille, marquée surtout par la faiblesse du nombre des grandes villes. De son côté, en appliquant le modèle de Lotka, Guérin-Pace (1995) étudie le cas français sur les siècles suivants, entre 1831 et 1982, en

utilisant un échantillon qui contient toutes les unités urbaines de plus de 2000 habitants. Elle trouve une augmentation constante du coefficient de hiérarchisation (de Lotka) ce qui induit l'hypothèse d'une hiérarchisation croissante du paysage urbain français. Brakman et al. (2004) étudient l'évolution des hiérarchies urbaines aux Pays Bas et trouvent que le coefficient de hiérarchisation passe de 0,55 en 1600 à 1,03 en 1990, puis baisse à 0,72 en 1990, ce qui semble conforter l'hypothèse de Parr. Bosker et al. (2006) entreprennent ce même travail pour les villes italiennes entre 1300 et 1861, soit la veille de la réunification de la péninsule. Leurs résultats vont aussi dans ce sens, même si les auteurs trouvent une différenciation entre le Nord et le Sud.

D'autres travaux explorent la question de l'évolution de ce coefficient, en s'interrogeant sur la stabilité (ou l'instabilité) de la pente de la distribution dans le long terme. Eaton et Eckstein (1997) montrent une très forte stabilité des hiérarchies urbaines en France et au Japon. Dobkins et Ioannides (2000) et Black et Henderson (2003) étudient l'évolution de la pente de la distribution rang taille des villes américaines entre 1900 et 1990. Tandis que les premiers concluent à une baisse systématique de β , et donc à un renforcement démographique des grandes agglomérations durant le vingtième siècle, les seconds arrivent à des résultats plus contrastés, en utilisant un échantillon différent qui s'appuie sur une définition plus complexe et diachronique de l'agglomération. Black et Henderson (2003) dessinent, par ce biais, un espace américain fortement dichotomique où les divergences entre les grandes agglomérations et les villes moyennes s'accroissent durant la deuxième moitié du vingtième siècle.

Différents travaux traitent de l'évolution de la distribution des tailles des villes dans tel ou tel pays, sans pouvoir dégager une tendance claire. Ainsi, Kundak et Dökmeci (2000) montrent que la distribution rang taille des villes turques passe d'une distribution initialement convexe à une distribution linéaire, sous l'impulsion de la croissance des villes moyennes, durant le vingtième siècle. Cori (1984) montre une évolution inverse pour l'Italie entre 1921 et 1971 où le coefficient de hiérarchisation baisse systématiquement, traduisant, par là, un mouvement de concentration vers les grands centres urbains. Petrakos et Arvanitidis (2008) trouvent, quant à eux, une stabilité du coefficient, proche de 1, de la distribution rang-taille grecque, durant toute cette période, tandis que Le Gallo et Chasco (2008) estiment que le coefficient de hiérarchisation de la distribution espagnole baisse entre 1900 et 1980, puis augmente, à nouveau, à partir de cette date, sous les effets de la périurbanisation et de la décongestion urbaine, ce qui confirmerait l'hypothèse de Parr. Enfin, Rose (2006) trouve deux résultats contradictoires : lorsqu'il étudie l'évolution des hiérarchies urbaines aux Etats-Unis, entre 1950 et 2000, il obtient une remarquable stabilité de la pente de la distribution, quasiment égale à -1, peu importe la taille de l'échantillon ; par contre, lorsqu'il applique ce travail aux échantillons issus de 50 différents pays, il trouve une augmentation progressive de la moyenne des coefficients obtenus (0,78 en 1950 ; 0,95 en 2000).

De façon plus générale, l'ensemble de ces travaux montre que le modèle rang-taille permet une bonne approximation des hiérarchies urbaines d'un pays et/ou d'une région, mais le lien entre cet indicateur de structuration urbaine et des indicateurs de développement économique (tels que le revenu ou le PIB par habitant) n'est pas clairement établi. La synthèse comparative proposée par Nitsch (2005) ne semble pas rejeter l'hypothèse de la courbe en U, même si la partie ascendante de la courbe ne se confirme pas de façon claire. Ainsi la moyenne des estimations baisse de façon conséquente jusqu'au milieu du 20^{ème} siècle puis se stabilise, voire remonte légèrement pour les estimations des coefficients de hiérarchisation après 1950.

3 L'hypothèse des distributions alternatives

3.1 La déviation vis-à-vis du modèle rang-taille : l'hypothèse d'une distribution parabolique

A côté de l'engouement scientifique pour la loi de Zipf, un certain nombre de travaux apportent un regard plus sceptique sur la validité de la relation rang-taille, en émettant – dès les travaux fondateurs de Beckman et de Simon – l'hypothèse que la distribution des tailles des villes ne suit pas une loi de Pareto. Ainsi, dans leur étude, Rosen et Resnick (1980) considèrent la possibilité d'une déviation vis-à-vis de la stricte linéarité entre le logarithme du rang et le logarithme de la taille des villes qui caractérise la loi de Pareto. Cette relation serait, alors, davantage représentée par une courbe parabolique qui peut être étudiée en ajoutant un terme quadratique (ou même cubique) à l'équation de Zipf :

$$\ln R_i = \alpha + \beta \ln T_i + \delta (\ln T_i)^2 \quad (7)$$

Lorsque δ est significativement différent de 0, on s'éloigne de la loi de Pareto. Si $\delta > 0$, la courbe de la distribution rang taille est strictement convexe, ce qui signifie que le nombre de villes moyennes est inférieur à celui préconisé par la loi de Zipf. Si, au contraire, $\delta < 0$ la courbe de la distribution est strictement concave, ce qui implique un nombre important de villes moyennes dont le poids démographique contrebalance celui des grandes agglomérations et des petites villes. Rosen et Resnick trouvent une déviation vis-à-vis d'une distribution de Pareto pour 30 des 44 pays étudiés. Lorsque Soo (2005) applique ce modèle sur les distributions rang taille des villes de 73 pays, il trouve, lui aussi, un δ significativement différent de 0 dans 68% des cas.

Gabaix et Ioannides (2004) montrent cependant que la plupart des études comparatives qui appliquent le modèle quadratique de Rosen et Resnick ont tendance à rejeter trop systématiquement l'hypothèse d'un δ égal à 0 ; or, selon les deux auteurs, ce rejet est artificiel, car lié tout simplement aux pratiques d'échantillonnage. En appliquant la simulation Monte Carlo sur vingt mille échantillons extraits d'une population de villes virtuelle qui obéit parfaitement à la loi de Zipf, Gabaix et Ioannides (2004) trouvent un δ significativement différent de 0 dans 93%, 80% et 65% des cas, pour des échantillons de taille $n = 20, 100$ et 500 respectivement. Cette erreur serait liée au biais de l'estimation par la méthode des moindres carrés ordinaires.

3.2 L'hypothèse d'une distribution lognormale

Tandis qu'un certain nombre de chercheurs intègrent l'hypothèse de convexité de la relation rang-taille des villes dans une approche qui utilise le modèle de Mandelbrot (Laherrere et Sornette, 1998, Malakarne et al, 2001), d'autres chercheurs, à l'instar du travail pionnier de Simon (1955), émettent l'hypothèse que la distribution des tailles des villes suit une loi lognormale, plutôt qu'une loi de Pareto. Si tel est le cas, il faudrait, alors, expliquer pourquoi un si grand nombre d'études obtiennent des résultats significatifs vis-à-vis de la loi de Pareto, lorsqu'elles régressent les logarithmes du rang et de taille des villes.

Sur un plan empirique, ces auteurs suggèrent que la distribution linéaire des logarithmes des tailles en fonction des logarithmes des rangs des villes ne représente que la partie haute (une distribution « tronquée ») d'une distribution qui épouse davantage les traits d'une loi lognormale, lorsque l'on considère la totalité de la population urbaine, c'est-à-dire également les villes de taille plus réduite. Selon Parr (1970), dans la plupart des cas, les données relatives aux petites villes, celles dont la population se situe en dessous d'un certain seuil de grandeur, ne sont pas disponibles. Ceci signifie que la plus grande partie de la distribution lognormale

pourrait ne pas être visible si la taille moyenne ou médiane de la population des villes se situait largement en dessous de la taille minimale des villes considérée par l'échantillonnage.

De nombreux auteurs testent l'hypothèse d'une adéquation des distributions des tailles des villes avec une loi lognormale (Malecki, 1970, Robson, 1973, Pumain, 1982, Guerin-Pace, 1985), avec des résultats contradictoires. Néanmoins, en s'appuyant sur les données du recensement 2000 du gouvernement américain et en considérant comme ville tout lieu désigné comme une entité administrative territoriale, Eeckhout (2004) développe une analyse robuste qui examine la relation entre le rang et la taille de 23539 villes des Etats-Unis (qui représentent 75% de la population de ce pays). En utilisant le test d'adéquation de Kolmogorov–Smirnov, Eeckhout montre que la distribution des tailles des villes américaines en fonction de leur rang suit une loi lognormale dont la fonction densité est la suivante :

$$\rho(T) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{1}{T} \exp\left[-\frac{1}{2}\left(\frac{\ln T - \mu}{\sigma}\right)^2\right] \quad (8)$$

où μ et σ sont respectivement la moyenne et l'écart-type du logarithme de la distribution normale.

Selon Eeckhout (2004), le fait que la plupart des études sur les villes américaines confirment la loi de Zipf, provient du fait qu'elles s'appuient sur des échantillons de taille extrêmement réduite. Ainsi Krugman (1996), Gabaix (1999) ou Black et Henderson (1999) utilisent des échantillons de 135 villes, ce qui correspond à 0,5% de l'ensemble des villes américaines, totalisant 30,2% de la population urbaine. Cet échantillon correspond à la partie extrême haute de la distribution, pour laquelle l'allure de la distribution lognormale « tronquée » et l'allure de la distribution de Pareto se superposent ce qui explique le fait que tant d'études confirment la loi de Zipf.

Divers travaux (Anderson et Ge, 2005, pour la Chine, Zeyneloglu et al, 2005, pour la Turquie, Nota et Song, 2007, pour la Chine) suivent l'argumentation de Eeckhout et confirment l'adéquation de la distribution des tailles des villes avec une distribution lognormale, dès lors que l'on considère la totalité de la population des villes d'un pays. Ces auteurs contestent la définition arbitraire d'une borne inférieure pour l'échantillonnage qui conduit à une interprétation de la relation entre le rang et la taille des villes sur une distribution « tronquée ». Leur définition de la ville pose, néanmoins plusieurs soucis, car ils finissent par confondre localisation isolée et aire urbaine. Ainsi, dans le travail de Eeckhout, la taille médiane des 23539 villes se situe à 1337 personnes, ce qui signifie qu'un très grand nombre de villes représentent à peine un village. Or, 80% de la population américaine vit dans les 256 plus grandes agglomérations du pays, ce qui permet de considérer que la distribution tronquée est, au contraire, une bonne approximation des hiérarchies urbaines américaines. Cette réflexion peut être généralisée sur la plupart des pays développés et/ou à revenu élevé.

3.3 *Prolongements et interrogations méthodologiques*

A la lignée des travaux précédents, certains chercheurs ont voulu tester l'hypothèse de l'adéquation des distributions réelles des tailles des villes avec des distributions théoriques alternatives. Ainsi Reed (2001 et 2002) développe l'hypothèse que la distribution de la taille des villes suit une loi double Pareto – lognormale (dPIN), c'est-à-dire une distribution théorique qui, sur l'échelle logarithmique, a le corps d'une distribution lognormale et des queues inférieure et supérieure qui suivent une loi de Pareto. Chaque queue de la distribution épouse la forme suivante (Anderson et Ge, 2005) :

$$\ln\left(\frac{R_i}{n}\right) = \ln\left(\frac{\beta}{\alpha + \beta}\right) - \alpha\left(\frac{T_i}{T^*}\right) + e_i \quad \text{pour } T_i \geq T^* \quad (9)$$

$$\ln\left(\frac{R_i}{n}\right) = \ln\left(\frac{\beta}{\alpha + \beta}\right) - \beta\left(\frac{T_i}{T^*}\right) + e_i \quad \text{pour } T_i < T^* \quad (10)$$

où R_i et T_i représentent le rang et la taille de la ville i , T^* une valeur proche de la taille médiane des villes de l'échantillon et α et β deux estimateurs de hiérarchisation, déterminés par la méthode du Maximum de Vraisemblance. Reed (2001) trouve une bonne adéquation des distributions des tailles des villes de certaines régions espagnoles et américaines avec la distribution double Pareto.

De façon générale, il faut admettre que la littérature sur les hiérarchies urbaines et la loi de Zipf part d'un a priori théorique contestable, celui d'une distribution de la taille des villes qui suit une loi de Pareto, au lieu de s'interroger, de façon exploratoire, sur la meilleure adéquation possible entre la distribution réelle et les différentes distributions théoriques. En utilisant une transformation Box-Cox, Hsing (1990) rejette l'adéquation systématique entre la distribution des villes américaines et la loi de Pareto (voir aussi Cameron, 1990, Alperovich et Deutsch, 1995).

La plupart des études dans ce domaine utilisent des tests d'adéquation plus ou moins puissants, tels que le test de Pearson (Anderson et Ge, 2005), celui de Kolmogorov Smirnov (Eekhout, 2004) ou celui de Akaike (Chen et Fu, 2007). En s'appuyant partiellement sur le travail de Reed, Benguigui et Blumenfeld-Lieberthal (2007) proposent un modèle général, susceptible de décrire différentes formes de distributions réelles :

$$\ln T = \ln T_0 - H(\alpha)\mu[b + H(\alpha)\ln R]^\alpha \quad (11)$$

où R et T représentent le rang et la taille de la ville i , T_0 un paramètre qui dépend de la taille de la plus grande ville, μ et b deux constantes et α un coefficient de hiérarchisation positif. $H(\alpha)$ est une fonction égale à -1 si $\alpha < 1$ et à 1 si $\alpha \geq 1$. Lorsque $\alpha = 1$, elle décrit une distribution de Pareto et dans ce cas, μ est le coefficient de Pareto.

Par ce biais, Benguigui et Blumenfeld-Lieberthal (2007) arrivent à une classification des différentes distributions des tailles des villes : la première classe correspond aux distributions qui obéissent à une loi puissance (avec $\alpha = 1$), la seconde classe aux distributions exponentielles (avec $\alpha > 1$), enfin la troisième classe (avec $0 < \alpha < 1$) décrit une distribution lognormale lorsque $\alpha = 0,5$. Sur 41 distributions de différents pays testées par les deux auteurs, 17 suivent une distribution de Pareto ($\alpha = 1$) où le coefficient de Pareto μ varie entre 0,61 et 1,25, 13 affichent une valeur de α significativement supérieure à 1 (entre 1,3 et 2,59) et seule une distribution affiche une valeur de α inférieure à 1 (la Chine avec $\alpha = 0,41$), tandis que dix autres distributions ont une allure moins homogène qui ne peut être décrite par un seul modèle.

Il est intéressant de signaler que les résultats de Benguigui et Blumenfeld-Lieberthal ne semblent guère justifier une distribution de la taille des villes qui suit une distribution lognormale.

De façon générale, tous ces travaux soutiennent que, sur un plan méthodologique, il est préférable de chercher quelle est la meilleure loi théorique permettant de caractériser la distribution des tailles des villes, plutôt que de partir d'un a priori d'une loi de Pareto. Seulement, la plupart des modèles utilisés sont des modèles à deux variables explicatives, permettant de mieux prendre en compte la convexité de la courbe, tandis que le grand avantage de la distribution de Pareto est de concentrer toutes les informations sur un seul

indicateur, la valeur du coefficient de hiérarchisation. Dès lors, des choix d'arbitrage s'imposent, car en optant pour des lois théoriques alternatives à celle de Pareto, on gagne en qualité d'estimation, mais on perd en qualité d'interprétation. C'est probablement la raison pour laquelle ces études n'ont connu qu'un succès très limité au sein de la science régionale.

4 Le besoin d'une méthode d'estimation fiable

4.1 Estimation du coefficient de Pareto par les méthodes des Moindres Carrés

A partir du moment où la démonstration d'une linéarité entre les logarithmes de rang et de taille des villes d'un pays ou d'une région est établie, la méthode de l'estimation du coefficient de hiérarchisation acquiert un rôle fondamental quant aux résultats obtenus. La plupart des études sur la loi de Zipf estiment le coefficient de hiérarchisation, en utilisant le modèle log-log du rang en fonction de la taille ou son inverse (celui de Lotka), qui met en relation la taille en fonction du rang.

L'intérêt du second modèle réside sur le fait que la valeur du coefficient de hiérarchisation γ augmente lorsque le degré de hiérarchisation (le poids des grandes villes) augmente aussi et vice-versa. Lorsque la distribution de la taille des villes suit parfaitement une loi de Pareto, γ est l'inverse du coefficient de Pareto β , tandis que la constante est un indicateur de la taille de la plus grande ville du système. L'augmentation de cette constante, observée durant une grande partie du 20^{ème} siècle dans de nombreux systèmes urbains traduit un processus généralisé de concentration des populations urbaines dans les plus grandes villes. Cependant, depuis le début des années quatre-vingt-dix, cette constante a tendance à stagner, voire baisser, en partie à cause des effets de congestion observés dans les grandes capitales qui conduisent à une migration relative de la population vers des centres urbains secondaires dont le taux de croissance démographique s'accélère (Ades et Glaeser, 1995 ; Pumain et Moriconi-Ebrard, 1997, Dobkins et Ioannides, 2000).

Malgré la grande popularité de ces modèles, plusieurs auteurs accusent le biais de cette méthode d'estimation, notamment pour les échantillons de petite taille. Gabaix et Ioannides (2004) montrent ainsi que la méthode des Moindres Carrés Ordinaires (MCO) a tendance à sous-estimer la valeur du coefficient de hiérarchisation pour les petits échantillons, ce qui conduit à sur-dimensionner la taille de la plus grande ville du système. Selon Nishiyama, Osada et Morimune (2008), le biais dans l'estimation par les MCO est lié au processus de classement des villes par rang, c'est-à-dire au fait que l'on construise la variable expliquée R , en ordonnant la variable explicative T . Ceci est à l'origine de l'apparition de corrélations positives entre les résidus supposés indépendants, ce qui fait de la loi de Zipf un faux phénomène statistique (Gan, Li et Song, 2006).

De nombreux auteurs ont essayé d'apporter quelques corrections au biais de l'estimation du coefficient de Pareto par la méthode des MCO. En s'appuyant sur la généralisation du modèle de Zipf par Mandelbrot, Gabaix et Ibragimov (2006) considèrent la relation suivante :

$$\ln(R_i - \xi) = \ln C - \beta \ln T_i \quad (12)$$

avec $0 \leq \xi < 1$, où $\xi = 0$ correspond au modèle de Zipf. En suivant les démonstrations de Kratz et Resnick (1996) et de Csörgö et Viharos (1997), Gabaix et Ibragimov (2006) montrent que l'espérance d'estimation du coefficient de hiérarchisation $\hat{\beta}$ obéit à :

$$E(\hat{\beta}/\beta - 1) = \frac{(2\gamma - 1)\ln^2 n}{4n} + u\left(\frac{(\ln n)^2}{n}\right) \quad (13)$$

$$\Rightarrow \hat{\beta}/\beta = 1 + \sqrt{\frac{2}{n}}N(0,1) + \frac{(\ln n)^2(2\xi - 1)}{4n} + u\left(\frac{(\ln n)^2}{n}\right)$$

Il est aisé de montrer que la meilleure estimation de β est fournie lorsque $\xi = 1/2$. La relation rang taille des villes, pour laquelle le biais de l'estimation du coefficient pour des petits échantillons est minimale, prend alors la forme :

$$\ln(R_i - \frac{1}{2}) = \ln C - \hat{\beta} \ln T_i \quad (14)$$

tandis que son écart-type peut être estimé par :

$$\sigma(\hat{\beta}) = \sqrt{\frac{2}{n}}\beta \quad (15)$$

De leur côté, en travaillant sur le modèle de Lotka, Nishiyama et Osada (2004) proposent une correction de l'estimation du coefficient de hiérarchisation de Lotka, en le multipliant par une constante qui ne dépend que du nombre de villes n retenues dans l'échantillon :

$$\bar{\gamma} = \frac{n \sum (\ln i)^2 - (\sum \ln i)^2}{n \sum_{i=1}^n \ln i \left(\frac{1}{n} + \dots + \frac{1}{i} - 1 \right)} \gamma \quad (16)$$

Dans le prolongement de ce travail, Nishiyama, Osada et Morimune (2008) cherchent à améliorer l'estimation du coefficient d'hiérarchisation de Lotka, en utilisant la méthode des Moindres Carrés Généralisés (MCG). Si :

$$Y' = [\ln T_1 \quad \ln T_2 \quad \Lambda \quad \ln T_n] \quad (17)$$

$$X' = \begin{bmatrix} 1 & 1 & \Lambda & 1 \\ \ln 1 & \ln 2 & \Lambda & \ln n \end{bmatrix} \quad (18)$$

et $\Omega = \text{Var}(Y)$ (19)

alors, les estimations de γ et A , ainsi que leurs variances sont égales à :

$$\begin{bmatrix} \hat{A} \\ \hat{\gamma} \end{bmatrix} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}Y \quad (20)$$

$$\text{Var} \begin{bmatrix} \hat{A} \\ \hat{\gamma} \end{bmatrix} = (X'\Omega^{-1}X)^{-1} \quad (21)$$

Par le biais d'une simulation, les auteurs montrent que la méthode des MCG réduit fortement le biais de la méthode MCO pour les petits échantillons.

4.2 Estimation du coefficient de Pareto par les méthodes semi-paramétriques

Comme toutes les distributions de Pareto, la distribution rang-taille des villes se caractérise par une probabilité d'apparition de valeurs extrêmes pouvant conduire à la

formation de « queues épaisses » qui affichent un comportement spécifique par rapport au reste de la distribution (Stoev, Michailidis et Taqqu, 2008). Ainsi, selon la valeur de la borne inférieure considérée, on peut se retrouver avec un échantillon qui affiche des écarts extrêmes (par exemple, le ratio entre New-York, la plus grande ville des Etats-Unis avec 8 millions d'habitants et Duffield, Virginia, avec 52 habitants, est de 1/150 000).

Le problème des queues épaisses à celui du « bornage » de la distribution rang-taille des villes (Malakarne et al, 2001) : en premier lieu, lorsque l'on procède à des comparaisons internationales, les bornes inférieures ne peuvent pas être les mêmes, pour chaque pays. A titre d'exemple, les villes avec une population supérieure à 100 000 habitants représentent 15% des villes américaines, mais seulement 4% des villes brésiliennes. En second lieu, lorsque l'on augmente considérablement la taille de l'échantillon, en incluant les plus petites villes, une déviation peut apparaître vis-à-vis de la loi de Pareto. La combinaison de ces deux problèmes fait que la loi de Zipf est vérifiée, le plus souvent, pour des échantillons d'une taille spécifique et lorsque l'on s'éloigne de cette taille « idéale » (en augmentant ou en réduisant le nombre de villes), on constate un écart vis-à-vis de la loi rang-taille également. C'est aussi la raison pour laquelle la loi de Zipf a tendance à se confirmer plus facilement pour des pays de grande taille qui possèdent de nombreuses grandes villes, par exemple les Etats-Unis, que les pays à taille plus réduite, par exemple de nombreuses nations européennes.

Afin d'améliorer le calcul de la pente des distributions avec une queue épaisse, de nombreux chercheurs font appel à l'utilisation de méthodes semi-paramétriques et plus particulièrement à l'estimateur de Hill qui correspond à la méthode du maximum de vraisemblance, lorsque la distribution étudiée épouse parfaitement la loi de Pareto.

Pour un échantillon de n villes avec des tailles $T_1 > \dots T_j > \dots T_n$, l'estimateur de Hill est donné par :

$$\hat{\beta} = \frac{n-1}{\sum_{j=1}^{n-1} (\ln T_j - \ln T_n)} \quad (22)$$

tandis que son écart-type est égal à (Gabaix et Ioannides, 2004) :

$$\sigma_n(\hat{\beta}) = \hat{\beta}^2 \left(\frac{\sum_{j=1}^{n-1} j (\ln T_j - \ln T_{j+1})^2}{n-1} - \frac{1}{\hat{\beta}^2} \right)^{\frac{1}{2}} (n-1)^{-\frac{1}{2}} \quad (23)$$

Dans sa comparaison internationale des distributions rang-taille des villes, Soo (2005) montre que l'utilisation de l'estimateur de Hill permet de baisser sensiblement le rejet de la loi de Zipf, par rapport à la méthode des MCO. Dimou et Schaffar (2007) concluent dans le même sens dans leur étude des distributions rang-taille de dix pays de la péninsule balkanique. Néanmoins, McCullogh (1997), Embrechts et al.(1997), Weron (2001) rappellent le biais que comporte cet estimateur sensible aussi aux choix de la borne inférieure et à l'apparition de valeurs extrêmes. Ainsi, si k est le nombre des plus grandes villes (qui forment une queue épaisse) d'un échantillon donné de taille n , le biais de l'estimateur de Hill augmente avec k , tandis que la variance de l'estimateur est proportionnelle à $1/k$. Ce biais peut être approximé de la façon suivante :

$$\hat{\beta} = \beta + \lambda k + u \quad \text{avec } k = 1, 2, \dots, m \quad (24)$$

Lorsque k tend vers 0, le biais de l'estimateur λk diminue, mais sa variance augmente, ce qui conduit à un choix « délicat » entre la réduction du biais et la pertinence de l'estimation (Huisman et al, 2001, Durfour et Kurz-Kim, 2005).

4.3 Une comparaison des différentes méthodes d'estimation par le biais d'une simulation Monte-Carlo

Afin de comparer les performances des différentes méthodes d'estimation du coefficient de hiérarchisation en fonction de la taille de l'échantillon, ce papier propose une simulation Monte Carlo. Cinq tailles différentes sont considérées : 20, 50, 100, 200 et 500 villes et, pour chacune de ces tailles, 20000 échantillons qui suivent la loi de Zipf (avec un coefficient de Pareto β égal à 1) sont construits. Pour chacune des tailles considérées, il y a, donc, 20000 estimations du coefficient de hiérarchisation selon chaque méthode : le calcul du coefficient de Pareto par les MCO, les MCO corrigés par Gabaix et Ibragimov et la méthode de Hill et le calcul du coefficient de hiérarchisation de Lotka par les MCO, en lui appliquant la correction de Nishiyama et Osada et, enfin, en utilisant la méthode des MCG.

| Taille échantillon | 20 | 50 | 100 | 200 | 500 |
|--|---------------|---------------|---------------|---------------|---------------|
| β (MCO) | | | | | |
| Moyenne | 0,9018 | 0,9225 | 0,9437 | 0,9613 | 0,9782 |
| Ecart-type réel | 0,2804 | 0,1813 | 0,1324 | 0,0962 | 0,0624 |
| Moyenne des e- types calculés | 0,0485 | 0,0235 | 0,0137 | 0,0078 | 0,0036 |
| IC à 5% | [0,52 ; 1,42] | [0,65 ; 1,24] | [0,74 ; 1,17] | [0,81 ; 1,13] | [0,88 ; 1,08] |
| β_{cor}(MCO corr) | | | | | |
| Moyenne | 1,0504 | 1,0101 | 1,0016 | 0,9986 | 0,9984 |
| Ecart-type réel | 0,3203 | 0,1944 | 0,1378 | 0,0983 | 0,0630 |
| Moyenne des e- types calculés | 0,0588 | 0,0263 | 0,0147 | 0,0081 | 0,0038 |
| Ecart type estimé | 0,3162 | 0,2000 | 0,1414 | 0,1001 | 0,0632 |
| IC à 5% | [0,61 ; 1,65] | [0,72 ; 1,35] | [0,79 ; 1,24] | [0,84 ; 1,16] | [0,90 ; 1,10] |
| β (Hill) | | | | | |
| Moyenne | 1,0586 | 1,0219 | 1,0112 | 1,0054 | 1,0021 |
| Ecart-type réel | 0,2570 | 0,1482 | 0,1019 | 0,0708 | 0,0449 |
| Moyenne des e- types calculés | 0,2324 | 0,1433 | 0,1008 | 0,0710 | 0,0448 |
| IC à 5% | [0,71 ; 1,53] | [0,80 ; 1,29] | [0,85 ; 1,19] | [0,89 ; 1,13] | [0,93 ; 1,08] |
| γ (MCO) | | | | | |
| Moyenne | 1,1395 | 1,0823 | 1,0529 | 1,0339 | 1,0180 |
| Ecart-type réel | 0,3321 | 0,2013 | 0,1403 | 0,0991 | 0,0629 |
| Moyenne des e- types calculés | 0,0646 | 0,0291 | 0,0159 | 0,0086 | 0,0038 |
| IC à 5% | [0,67 ; 1,75] | [0,78 ; 1,44] | [0,84 ; 1,30] | [0,88 ; 1,21] | [0,92 ; 1,13] |
| γ_{cor} (γ corrigé) | | | | | |
| Moyenne | 0,9987 | 1,0001 | 0,9995 | 0,9998 | 0,9997 |
| Ecart-type réel | 0,2910 | 0,1860 | 0,1332 | 0,0958 | 0,0618 |
| Moyenne des e- types calculés | 0,0566 | 0,0269 | 0,0151 | 0,0083 | 0,0037 |
| IC à 5% | [0,58 ; 1,53] | [0,72 ; 1,33] | [0,80 ; 1,23] | [0,85 ; 1,17] | [0,90 ; 1,11] |
| γ (MCG) | | | | | |
| Moyenne | 1,1025 | 1,0523 | 1,0298 | 1,0169 | 1,0078 |
| Ecart-type réel | 0,2657 | 0,1543 | 0,1049 | 0,0723 | 0,0453 |
| Moyenne des e- types calculés | 0,3655 | 0,2708 | 0,2104 | 0,1609 | 0,1109 |
| IC à 5% | [0,71 ; 1,58] | [0,81 ; 1,32] | [0,86 ; 1,21] | [0,90 ; 1,14] | [0,93 ; 1,08] |

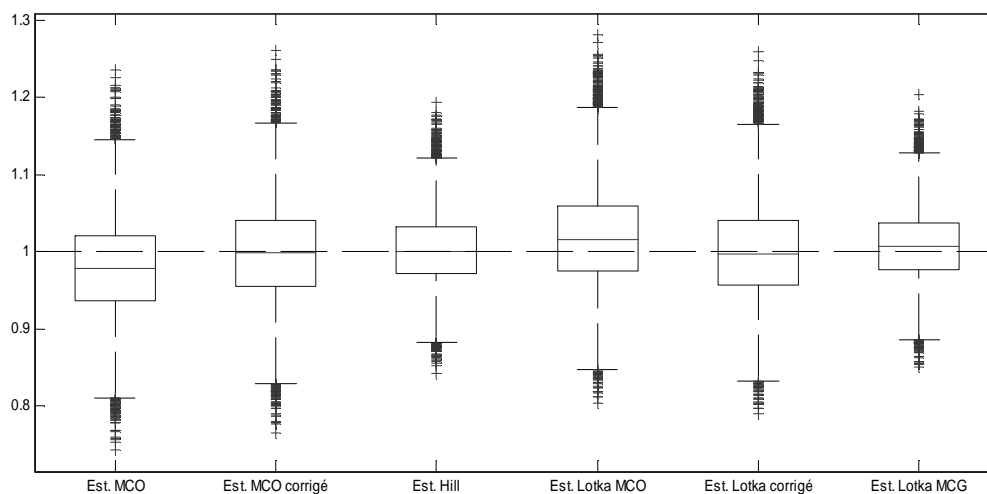
Tableau 1 : Estimations du coefficient de Pareto –Résultats de la simulation Monte Carlo (20000 estimations)

Le tableau 1 fournit les résultats obtenus : la moyenne, l'écart-type réel, la moyenne des écart-types calculés pour chaque coefficient de la régression, ainsi que l'intervalle de confiance à 5% du coefficient pour chaque série d'estimations. Pour les estimations avec la méthode des MCO corrigée, l'estimation de l'écart-type avec la formule proposée par Gabaix et Ibragimov (2006) est également calculée. Comme on peut le constater, toutes les estimations s'améliorent au fur et à mesure que la taille de l'échantillon augmente : les moyennes s'approchent de 1, les écart-types diminuent et les intervalles de confiance se

réduisent. Le biais est positif pour le calcul du coefficient de hiérarchisation avec le modèle de Zipf et négatif pour le calcul du coefficient de hiérarchisation avec le modèle de Lotka (en Annexe 2, les histogrammes des 20000 estimations pour les échantillons de taille de 20, 100 et 500 villes qui permettent de se rendre compte des effets d'asymétrie qui caractérisent les petits échantillons).

Le tableau 1 permet de constater l'efficacité des améliorations proposées par Gabaix et Ibragimov (2006) et par Nishiyama et Osada (2004) pour le calcul des coefficients de Pareto et de Lotka respectivement. Avec la correction de Gabaix et Ibragimov (2006), le biais de l'estimateur diminue de 0,10 à 0,05 pour les échantillons de taille 20 et de 0,022 à 0,002 pour les échantillons de taille 500. La correction de Nishiyama et Osada (2004) est encore plus performante : le biais diminue de 0,14 à 0,01 pour les échantillons de taille 20 et de 0,018 à 0,0003 pour les échantillons de taille 500. Cependant, la relation entre la réduction du biais et l'augmentation de la taille de l'échantillon est plus asymptotique, lorsque l'on applique ces deux corrections, sauf pour les échantillons de plus petite taille ($T = 20$). Il semblerait ainsi qu'il y ait une taille optimale de l'échantillon (entre 100 et 200), pour laquelle la correction de Gabaix et Ibragimov procure une estimation quasi-parfaite (β égal à 1), tandis que le coefficient de Lotka, corrigé par Nishiyama et Osada est moins sensible à la modification de la taille de l'échantillon, au-delà d'une taille 50.

L'ensemble des résultats obtenus montre que le calcul du coefficient de hiérarchisation par la méthode des MCO, pour le modèle de Zipf ou de Lotka, sans les corrections de Gabaix et Ibragimov (2006) ou de Nishiyama et Osada (2004), est nettement moins performant que celui proposé par les autres méthodes d'estimation. Or, la très grande majorité d'études sur les distributions rang taille des villes emploient cette méthode.



Graphique 1 : Comportement des différents estimateurs pour les échantillons de taille 500 (20000 simulations Monte Carlo)

Ceci est également confirmé par le graphique 1 qui permet de constater que la méthode Hill procure des meilleurs résultats. Notons, enfin, que contrairement à l'idée communément admise, selon laquelle au dessus d'une certaine taille de l'échantillon, les différentes méthodes pourraient permuer, l'étude observée arrive à la conclusion inverse. La corrélation (R^2) entre l'estimateur de Hill et les autres estimateurs baisse systématiquement au fur et à mesure que la taille de l'échantillon augmente. Ce résultat étonnant, lié à la baisse conjointe des écarts-types des différents estimateurs, montre que la spécificité de chaque estimateur se confirme au fur et à mesure que le nombre d'informations augmente. Sur un plan

méthodologique, ceci implique le besoin de procéder à un choix de la méthode d'estimation du coefficient de hiérarchisation, indépendamment de la taille de l'échantillon considérée.

5 La loi de Zipf dans un contexte dynamique

5.1 *Les approches de la croissance aléatoire*

Les sections précédentes traitent des modèles rang taille de façon statique, en cherchant essentiellement à décrire la distribution des tailles des villes à un moment donné par le biais de l'estimation d'un coefficient de hiérarchisation qui peut ainsi être comparé, dans l'espace et dans le temps, avec les coefficients d'autres distributions. Or, ce qui semble indispensable est de comprendre le processus de croissance qui conduit à une évolution du coefficient de hiérarchisation ou même qui conduit à une distribution rang-taille conforme à la loi de Zipf. C'est dans ce contexte qu'un ensemble de travaux cherche à lier les théories de la croissance urbaine avec celles sur les distributions rang-taille des villes.

Ces travaux peuvent schématiquement être divisés en deux sous-groupes : d'une part les approches en termes de croissance aléatoire et, d'autre part, les approches en termes de croissance endogène. Leur opposition fondamentale concerne l'existence d'une relation cause-à-effet entre la taille et la croissance des villes : dans les premières, la croissance ne dépend pas de la taille, ce qui vérifie une autre loi statistique, celle de Gibrat, en invalidant, par-là, tout un ensemble d'analyses dans les sciences régionales, qui se focalisent sur le rôle des effets d'échelle. Dans les secondes, cette relation existe, mais il convient alors de déterminer quelle est sa nature. Selon le modèle de croissance urbaine choisi, la loi de Zipf apparaît comme l'aboutissement ou pas d'un processus de croissance urbaine spécifique.

Les prémices des théories de la croissance aléatoire se trouvent dans l'article de Simon (1955) qui s'appuie sur la loi de l'effet proportionnel de Gibrat, mais c'est Gabaix (1999) qui élabore le modèle de base actuellement utilisé, construit sous des hypothèses très restrictives telles qu'une population croissante, la libre mobilité du travail - mais uniquement pour les jeunes générations - et des technologies de production à rendement constant. Dans le cadre de ces hypothèses, la croissance des villes dépend de l'apparition de chocs ponctuels exogènes, distribués de façon aléatoire entre les villes, qui sont générateurs d'aménités urbaines et modifient donc les choix de localisation des ménages mobiles. Ces chocs sont liés à des causes exogènes, naturelles (catastrophes naturelles, tremblements de terres, famines) ou historiques (guerres), mais aussi à des choix de politique économique locale (mesures modificatives de la taxation locale, des services municipaux, mise en place d'une politique de préservation de l'environnement, etc.) qui modifient la fonction d'utilité des ménages résidant dans une ville.

Dans ce contexte, et à partir d'un certain seuil de taille, la dynamique d'un système urbain suit un processus stochastique, au sein duquel chaque ville a la même espérance de croissance démographique (Dimou et al. 2008). La dynamique urbaine suit une loi de Gibrat, ce qui signifie que le taux de croissance d'une ville est indépendant de sa taille, tandis que sa variance est proportionnelle à $(\beta - 1)$ (où β est le coefficient de hiérarchisation de la distribution rang-taille). Cordoba (2008), prolonge l'argumentation de Gabaix, en montrant que la loi de Gibrat est la condition nécessaire et suffisante pour produire, à l'état stationnaire, une distribution rang-taille qui obéit à la loi de Zipf.

De nombreux travaux empiriques récents ont cherché à tester la validité de la croissance aléatoire. Ces travaux concernent essentiellement les systèmes urbains affectés par des guerres. Ainsi, Davis et Weinstein (2002) supposent que chaque localité subît, de façon permanente et aléatoire, des chocs ponctuels qui conduisent à une réallocation de la

population et, par-là, à une instabilité temporelle de la pente de la distribution de la taille des villes. Ceci signifie que lorsqu'une ville perd son rang, suite à un choc exogène positif ou négatif, il n'y a aucune raison pour qu'elle le retrouve par la suite. La distribution rang taille des villes se trouve, ainsi, définitivement modifiée.

Davis et Weinstein (2002) étudient le cas d'un pays dont les villes ont subi des chocs importants durant le vingtième siècle : le Japon d'avant et d'après les bombardements alliés de la deuxième guerre mondiale. Ils trouvent que ces bombardements ont eu des effets importants mais temporaires et que les villes japonaises ont retrouvé leur place initiale d'avant-guerre, dans la distribution rang-taille, au bout d'une quinzaine d'années, ce qui va à l'encontre des prédictions des approches de la croissance aléatoire. Ceci signifie qu'il y a un effet mémoire de la taille et de la primauté urbaine, auprès de la population.

Brakman et al (2004) qui travaillent sur le cas analogue des villes allemandes, arrivent, de leur côté, à une conclusion quelque peu différente, car ils observent que les effets structurels de ces bombardements sur la dynamique démographique des villes sont beaucoup plus longs que ne le prétendent Davis et Weinstein. Ils montrent que les petites villes, relativement épargnées des bombardements alliés, ont sensiblement et définitivement « gagné des places » dans la distribution rang-taille des villes allemandes, entre l'avant et l'après-guerre, même si leur progression s'estompe. Sharma (2003) arrive à une conclusion quasi-identique, lorsqu'elle étudie les effets du choc de la partition entre l'Inde et le Pakistan, tandis que d'autres études empiriques récentes sur le Portugal, les Etats-Unis ou la Chine réfutent également la loi de Gibrat en matière de croissance urbaine (Delgado and Godinho, 2007, Garmestani *et al*, 2007, Dimou *et al*, 2008).

5.2 Les approches de la croissance endogène

Les approches de la croissance urbaine endogène ont été essentiellement développées par Henderson (Black et Henderson, 1999 et 2003, Henderson, 2004, Henderson et Wang, 2004). L'hypothèse de base est que les firmes se concentrent géographiquement afin d'exploiter des effets d'agglomération qui peuvent être liés soit à une dynamique de spécialisation (externalités de localisation) soit à une dynamique de diversification (externalités d'urbanisation). A l'opposé, elles subissent des déséconomies d'agglomération, liées à l'augmentation des coûts migratoires pendulaires et aux effets de congestion. Les villes ont des tailles variables car elles dépendent des choix de localisation des firmes, effectués en fonction des avantages et des inconvénients de chacune d'entre elles.

Dans ce contexte, la croissance urbaine est un processus déterministe, sensible aux effets de taille. Plusieurs configurations possibles apparaissent : une convergence des villes vers une taille unique, avec les petites villes qui croient plus vite que les grandes, une divergence, impliquant un processus inverse ou même une croissance parallèle qui confirmerait une stabilité de la distribution rang-taille. Il est évident que ces modèles ne produisent pas nécessairement une distribution rang-taille qui, à l'état stationnaire, suit la loi de Zipf.

Différentes tentatives de tester les modèles de croissance endogène ont été développés. Dans leur étude des villes françaises et japonaises, Eaton et Eckstein (1997) montrent que les plus grandes villes, dans chaque pays, ont conservé leur rang durant toute le vingtième siècle, ce qui implique que leur croissance est parallèle, plutôt que convergente ou divergente. Black et Henderson (1999), eux, introduisent l'hétérogénéité des villes, en proposant un modèle avec deux types de villes où la taille urbaine est fonction des externalités d'échelle et de la concentration du capital humain. En admettant un système fermé, sans apparition de nouvelles villes et une libre mobilité des ménages dont les choix migratoires dépendent du stock de capital humain de chaque ville, ils montrent que la croissance urbaine est exclusivement liée à la croissance du capital humain.

Dans un premier travail sur l'évolution de la distribution des villes américaines entre 1900 et 1990, Black et Henderson (1999) trouvent que la taille moyenne des villes augmente sous l'impulsion du changement technologique et l'accumulation des connaissances, mais aussi que les petites villes affichent une croissance supérieure aux grandes, ce qui conduit à une convergence vers une taille de ville optimale. Cependant, en relâchant quelques hypothèses, et notamment la forte relation entre la croissance urbaine et le rôle de l'éducation, à laquelle est associée de façon quasi-exclusive la croissance du capital humain, les deux auteurs trouvent, dans un travail plus récent, une croissance parallèle des villes américaines durant cette même période, ce qui semble plus en conformité avec les résultats des autres chercheurs (Black et Henderson, 2003).

Plusieurs études testent l'existence d'une croissance urbaine parallèle. Sharma (2003), Bosker et al (2006), Chen et Fu (2007), Dimou et al (2008), Delgado et Godinho (2007), rejettent de façon assez conséquente l'hypothèse d'une croissance parallèle des villes, respectivement pour les systèmes urbains indien, allemand, chinois et portugais, tandis que Schaffar et Dimou (2010) arrivent à une conclusion plus mitigée dans leur étude du système urbain de la péninsule balkanique.

5.3 Les approches hybrides

A côté des approches précédentes, un certain nombre d'autres travaux se sont développés.

Dans un premier temps, les modèles de la nouvelle économie géographique apportent des contributions spécifiques à l'étude de la croissance urbaine, sans néanmoins, s'intéresser aux questions liées l'évolution de la distribution rang-taille, sauf exception (Krugman, 1996, Fujita, Krugman et Venables, 1999). Ces modèles admettent de façon générale que l'émergence des villes, liée aux choix de localisation des firmes, est initialement un processus aléatoire où les facteurs géographiques (proximité de l'eau, climat, sol montagneux, etc.) peuvent jouer un rôle limité. L'apparition de rendements croissants localisés renforce, par la suite, les choix initiaux et génère une croissance urbaine autoentretenu.

Plus récemment, un certain nombre de contributions ont proposé des modèles qui cherchent à rompre la dichotomie entre croissance aléatoire et croissance endogène. Ainsi, en prolongeant la démarche de Ioannides et Overman, Rossi-Hansberg et Wright (2007) introduisent un modèle de croissance urbaine où des chocs ponctuels spécifiques à un secteur ou un groupe de firmes (notamment des innovations autonomes qui correspondent à un changement technologique exogène), distribués de façon aléatoire, modifient la productivité totale des facteurs dans chaque ville. Dans ce contexte, l'espérance du taux de croissance urbaine et sa variance restent indépendantes vis-à-vis de la taille des villes, conformément à la loi de Gibrat. Cependant, ce processus ne génère pas, nécessairement, une distribution rang-taille des villes qui obéit à la loi de Zipf, à cause des effets démographiques et migratoires de court terme des chocs de productivité.

Duranton (2002) modélise les « micro-fondements » du processus stochastique qui détermine l'évolution de la distribution rang-taille des villes. Dans un environnement de concurrence monopolistique, avec libre circulation des travailleurs et un nombre fixe de villes, il montre que des innovations sectorielles (chocs technologiques) conduisent à un processus de localisation/délocalisation des activités productives qui peut expliquer la croissance ou le déclin des villes, une partie des travailleurs se déplaçant ainsi, au gré des innovations, de ville en ville. Duranton (2002) émet, dans un premier temps, l'hypothèse que la localisation de l'innovation est un processus aléatoire (se déplaçant en « *saut de grenouille* ») qui dépend de chocs exogènes liés, entre autres, aux politiques de R&D, engagées par chaque ville. Par la suite, il introduit les effets des économies d'agglomération et de la congestion, ce qui conduit à une croissance urbaine convergente et à des distributions rang-taille légèrement concaves, qui semblent reproduire assez fidèlement les distributions

rang-taille des villes françaises et américaines (Duranton, 2008). Le modèle de Duranton n'engendre, donc, pas la loi de Zipf.

Enfin Pumain et al (2006) proposent une théorie évolutionniste des systèmes urbains, permettant d'interpréter leurs caractéristiques de hiérarchisation. En se rapprochant des travaux de Catin et Van Huffel, ils proposent une évolution des systèmes urbains où les grandes villes captent, par leur diversité et leur complexité, l'innovation et la nouveauté, en relayant aux villes moyennes les activités plus banalisées. Dans cette approche, la diffusion de l'innovation des grandes vers les petites villes ne se fait pas à vitesse constante ni égale, mais dépend des capacités d'adaptation de chaque système urbain, d'où une instabilité de la pente de la distribution rang-taille (Pumain, 2006). Bretagnolle et al (2006) mettent ainsi, en évidence, les trajectoires relatives des tailles des plus grandes villes françaises.

Ces travaux théoriques récents restent encore largement exploratoires de la relation entre la nature de la croissance urbaine et la forme de la distribution de la taille des villes. Par le biais d'un certain nombre de simulations, les approches hybrides parviennent à calibrer des modèles de croissance urbaine qui paraissent pertinents, permettant de reproduire la loi de Zipf ou, tout au moins, expliquer les éventuelles déviations de la distribution des tailles des villes vis-à-vis de cette loi. En s'appuyant essentiellement sur l'étude des systèmes urbains américain et français, leur justification empirique reste, néanmoins, encore limitée et des études appliquées supplémentaires sont nécessaires afin de renforcer leur validité.

6 Conclusion

Ce papier propose une relecture critique des principaux travaux des trente dernières années sur la loi de Zipf. Tout en rappelant les enjeux des modèles rang-taille en science régionale ainsi que leur faculté de proposer une mesure de comparaison simple des différentes hiérarchies urbaines dans l'espace et dans le temps, une grande importance est accordée, ici, aux controverses qui les accompagnent et qui sont liées au bornage de l'échantillon des villes et à la méthode d'estimation du coefficient de hiérarchisation.

Ces controverses mettent en lumière les risques d'erreur d'interprétation, liés essentiellement à des aspects méthodologiques. Se dessine, néanmoins, par derrière, une question plus conséquente qui concerne la définition même de la ville et de l'unité d'analyse urbaine, de façon plus générale (Guerois et Paulus, 2002). D'une part, la définition administrative, économique ou géographique de la ville, d'autre part la détermination du seuil de population à partir duquel une localité peut être désignée en tant qu'espace urbain sont des éléments qui modifient de façon conséquente les résultats obtenus par les modèles rang-taille (Nitsch, 2005, Soo, 2005). A cet égard, Black et Henderson (1999) et Bretagnolle et al (2002) proposent des analyses où le périmètre des villes considérées change dans le temps, afin de tenir compte d'une définition diachronique de l'espace urbain.

L'introduction des modèles rang-taille dans une approche dynamique permet d'ouvrir des nouvelles perspectives de recherche, en faisant de la loi de Zipf, en particulier, l'aboutissement d'un processus de croissance urbaine. Cette perspective exige, à présent, une validation empirique plus robuste, portant sur l'étude de différents systèmes urbains.

BIBLIOGRAPHIE

- Ades A, Glaeser E, 1995, Trade and circuses: explaining urban giants, *Quarterly Journal of Economics*, 110, 195—228.
- Allen G., 1954, The Courbe des populations, a further analysis, *Bulletin of the Oxford University Institute of Statistics*, Vol.16, 179-189.
- Alperovich, G., 1993, An explanatory Model of City-size distribution: Evidence from Cross country Data, *Urban Studies*. 9, 1591-1601.
- Alperovich, G. Deutsch J., 1995, The Size Distribution of Urban Areas: Testing For the Appropriateness of the Pareto Distribution Using a Generalized Box-Cox Transformation Function, *Journal of Regional Science*, 35, 267-276.
- Anderson G., Ge Y., 2005, The size distribution of Chinese cities, *Regional Science and Urban Economics* 35, 756-776.
- Auerbach, F., 1913, Das Gesetz der Bevölkerungskonzentration, reproduit dans *Regional Science and Urban Economics*, 31, 601-615.
- Beirlant P. *et al*, 1999, Tail index estimation and an exponential regression model, *Extremes*, 2177, 177-200.
- Beckmann M.J. 1958, City hierarchies and the distribution of city sizes. *Economic Development and Cultural Change*, 6, 243-248.
- Benguigui L., Blumenfeld-Lieberthal E., 2007, A dynamic model for city size distribution beyond Zipf 's law, *Physica A*, 384 , 613–627.
- Berry B.J.L., 1964, Cities as systems within systems of cities, *Papers of the Regional Science Association*, 13, 147-163.
- Black D., Henderson J.V., 1999, A Theory of Urban Growth, *Journal of Political Economy*, 107, 252-284.
- Black D., Henderson J.V., 2003, Urban evolution in the USA, *Journal of Economic Geography*, 3, 343-372.
- Bosker E., Brakman D., Garretsen H., Schramm M., 2006, The strategic bombing of German cities during World War II and its impact on city growth, *CESIFO working Paper*, n°1726.
- Bradford, S. C., 1934, Sources of informations on specific subjects. *Engineering*, 137, 85-86.
- Brakman, S., *et al*, 1999, The return of Zipf : A further understanding of the rank-size distribution, *Journal of Regional Science* 39, 183-213.
- Brakman D., Garretsen H., Schramm M., 2004, The strategic bombing of German cities during World War II and its impact on city growth, *Journal of Economic Geography*, 4, 201-218.
- Bretagnolle A., Paulus F., Pumain D., 2002, Time and space scales for measuring urban growth, *CyberGeo*, 219, pp.1-20.
- Bretagnolle A., Daudé E., Pumain D., 2006, From theory to modelling: urban systems as complex systems , *CyberGeo*, 335, 1-17.
- Cameron, T. A., 1990, One-Stage Structural Models to Explain City Size, *Journal of Urban Economics*, 27, 294-307.
- Carroll B, 1982, National City Size Distributions: What Do We Know After 67 Years of Research? *Progress in Human Geography*, Vol.9, 1-43.
- Catin M., Van Huffel C., 2004, L'impact de l'ouverture économique sur la concentration spatiale dans les pays en développement, *Région et Développement*, 20, 123-157.
- Chen Z., Fu S., 2007, Searching for the parallel growth of cities, *Acts of the European Regional Science Association Conference*, Paris, 28-31 August.

- Cordoba J.C., 2008, On the Distribution of City Sizes, *Journal of Urban Economics*, 63, forthcoming.
- Csörgo, S., Viharos, L., 1997, Asymptotic normality of least squares estimators of tails indices, *Bernoulli*, 3, 351-370.
- Davis D.R., Weinstein D.E., 2002, Bones, Bombs and Breakpoint: the Geography of Economic Activity, *American Economic Review*, 92, 1269-1289.
- Delgado A.P., Godinho I, 2007, Growth and Change in the Portuguese urban system: 1890-1901, *Acts of the International Conference on Regional and Urban modelling Conference*, Brussels, 1-2 June.
- Dimou M., Schaffar A., 2007, Evolution des hiérarchies urbaines et loi de Zipf. Le cas des Balkans, *Région et Développement*, 25, 65-86.
- Dimou M., Schaffar A., Chen Z., Fu S., 2008, La croissance urbaine chinoise reconsidérée, *Région et Développement*, 26, 109-131.
- Dobkins, L.H., Ioannides, Y.M., 2000, Dynamic evolution of U.S. cities, in Huriot, J., Thisse, J (Eds.), *The Economics of Cities, Theoretical Perspectives*. Cambridge University Press, Cambridge, 217-260.
- Duranton G., 2002, City size distributions as consequence of the growth process, *Mimeo, London, School of Economics*.
- Duranton G., 2008, Urban evolutions: the fast, the slow and the still, *American Economic Review*, March, 197-221.
- Dufour, J.-M., Kurz-Kim, J.-R., 2005, Exact inference and optimal invariant estimation for the tail coefficient of symmetric unstable distributions, *Working paper, Université de Montréal*
- Eaton J., Eckstein Z., 1997, Cities and Growth: Theory and Evidence from France and Japan, *Regional Science and Urban Economics*, XXVII, 443-474.
- Eeckhout J., 2003, Gibrat's Law for (all) Cities, *American Economic Review*, 94, 1429-1451.
- Embrechts, P., Kluppelberg, C, Mikosch, T., 1997, *Modelling Extremal Events for Insurance and Finance*, Springer, New York.
- Feurverger, A., Hall, P., 1999, Estimating a tail exponent by modelling departure from a pareto distribution, *Annals of Statistics* 27, 760-781.
- Fujita M., Krugman P., Venables A., 1999, *The Spatial Economy*, MIT Press, Cambridge, MA.
- Gabaix X., 1999, Zipf's Law for Cities: an Explanation, *Quarterly Journal of Economics*, 114, 739-767.
- Gabaix, X., Ibragimov, R., 2006, *Log(Rank - 1/2) : a simple way to improve the OLS estimation of tail exponents*, Discussion paper 2106, Harvard Institute of Economic Research, Harvard University.
- Gabaix, X., Ioannides, Y., 2004, The evolution of city sizes' distribution in Henderson J.V et Thisse J-F. (eds) *Handbook of regional and urban economics*, vol.4, Elsevier Science B.B, Amsterdam, 2341-2376.
- Gan L., Li D., Song S., 2006, Is the Zipf law spurious in explaining city-size distributions? *Economic Letters*, 92, 256-262.
- Garmestani *et al*, 2007, Departures from Gibrat's Law, discontinuities and city size distributions, *Urban Studies*, 44, 1997-2007.
- Gibrat, R., 1931, *Les Inégalités Economiques*, Librairie du Recueil Sirey, Paris.
- Glaeser *et al*, 1995, Economic growth in a cross-section of cities, *Journal of Monetary Economics*, 36, 117-143.
- Goodrich P., 1926, The statistical relationship between population and the city plan, in Burgess N. (eds) *The urban Community*, University of Chicago, 144-156.

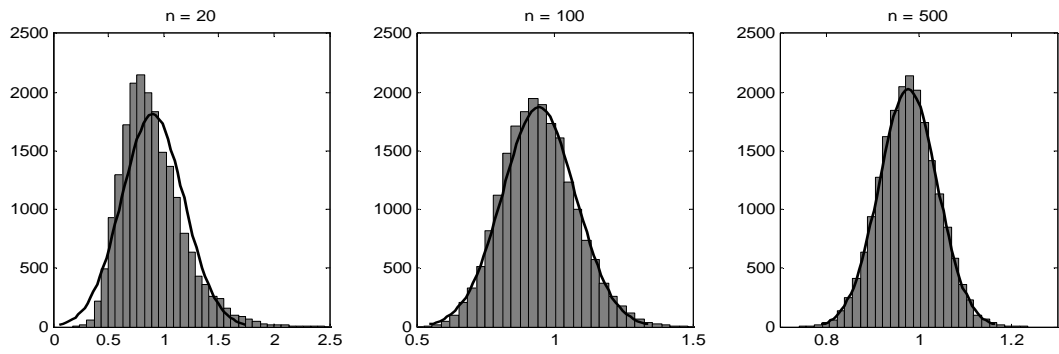
- Guérin-Pace F., 1995, Rank size distributions and the process of urban growth, *Urban Studies*, vol.32, n°3, 551-562.
- Guérois M., Paulus, F., 2002, Commune centre, agglomération, aire urbaine : quelle pertinence pour l'étude des villes ?, *Cybergeo*, 212, pp.1-18.
- Henderson V., Wang H., 2007, Aspects of the rural-urban transformation of countries, *Journal of Economic Geography*, 5(1), 23-42.
- Henderson V., 2004, Urbanization and Growth, *Working Paper*, Brown University.
- Hsing, Y., 1990, A Note on Functional Forms and the Urban Size Distribution, *Journal of Urban Economics*, 27, 73-79.
- Huisman, R., Koedijk, K. G., Kool, C. J. M. and Palm, F., 2001, Tail-index estimates in small samples, *Journal of Business and Economic Statistics* 19, 208-216.
- Kratz M., Resnick S., 1996, The QQ-estimator and heavy tails, *Communications in Statistics, Stochastic models*, 12, 699-724
- Krugman P., 1996, Confronting the mystery of Urban Hierarchy, *Journal of the Japanese and the International Economies*, 10, 399-418.
- Kundak S, Dökmeci V., 2000, *Growth, Distribution and Rank Stability of Urban Settlements in Turkey*, 6th World Congress of the RSAI, Lugano, May 16-20
- Laherrere J, Sornette D., 1998, Power laws, *European Physics*, 2, 525 -528.
- Le Gallo J., Chasco C., 2008, Spatial analysis of urban growth in Spain: 1900–2001, *Empirical Economics*, vol. 34(1), 59-80.
- Lepetit B, 1990, Patterns of settlement and political changes: The French Revolution and the National Urban Hierarchy, in Van der Woude, Hayami, De Vries (Eds.), *Urbanization in History*, Clarendon Press, Oxford
- Lotka A., 1941, The law of urban concentration, *Science*, n°94, 164.
- Malakarne et al, 2001, q -exponential distribution in urban agglomeration, *Physical review*, vol.65, 017106.
- Malecki E., 1980, Growth and change in the analysis of rank–size distributions: empirical findings, *Environment and Planning A*, n°12, 41–52.
- Mandelbrot B., 1959, A note on a class of skew distribution function: analysis and critique of a paper by H.A. Simon. *Information and Control*, 2, 90-99.
- McCulloch, J. H., 1997, Measuring tail thickness to estimate the stable index alpha: A critique, *Journal of Business and Economic Statistics* 15, 74-81.
- Moriconi-Ebrard F., 1993, *L'urbanisation du monde depuis 1950*, Paris, Anthropos.
- Nishiyama Y., et al, 2008, OLS-t-test revisited in rank-size rule regression, *DEE Discussion paper 06-3*, Kyoto Institute of economic research.
- Nota S., Song F., 2006, Further Analysis of the Zipf's Law: Does the Rank-Size Rule Really Exist? UNR Joint Economics Working Paper Series, *Working Paper*, 07-2006, University of Nevada.
- Nishiyama Y., Osada S., 2004, *Statistical theory of rank-size rule regression under Pareto distribution*, working paper, Kyoto Institute of economic research.
- Nitsch V., 2005, Zipf zipped, *Journal of Urban Economics*, 57, 86–100
- Nota S., Song F., 2006, Further Analysis of the Zipf's Law: Does the Rank-Size Rule Really Exist? *Working Paper*, 07-2006, University of Nevada.
- Pareto, V., 1897. *Cours d'Economie Politique*, Rouge et Cie, Paris.
- Parr, J.B., 1970 Models of city size in an urban system, *Papers and Proceedings of the Regional Science Association*, 25, 221-253.
- Parr, J., 1985, A Note on the Size Distribution of Cities Over Time, *Journal of Urban Economics*, 18, 199-212.
- Petrakos G, Arvanitidis P, 2008, Metropolitan Development and cooperation in Southeastern Europe, *Région et Développement*, 27, 33-59

- Pumain D. 1982, *La dynamique des villes*. Paris, Economica.
- Pumain D, Moriconi-Ebrard F., 1997, City Size distributions and metropolisation, *Geojournal*, 43(4), 307-314.
- Pumain D., 2006, Alternative explanations of hierarchical differentiation in urban systems, in Pumain D. (ed). Hierarchy in natural and social sciences, *Methodos*, series 3, 169-222.
- Pumain D, Paulus, F, Vacchiani-Marcuzzo C., Lobo J., 2006, Une théorie évolutive pour expliquer les lois d'échelle dans les systèmes de villes, *Cybergeo*, 243, juillet.
- Reed, W., 2001, The Pareto, Zipf and other power laws, *Economics Letters*, 74, 15-19.
- Rose A, 2006, *Cities and Countries*, Draft paper, Haas School of Business, University of Berkeley.
- Robson B. T. 1973, *Urban growth, an approach*. London, Methuen.
- Rosen, K., Resnick, M., 1980, The size distribution of cities: an examination of the Pareto law primacy, *Journal of Urban Economics*, 8, 165-186.
- Rossi-Hansberg E., Wright M., 2007, Urban Structure and Growth, *Review of Economic Studies*, 74, 597-624.
- Schaffar A, Dimou M, 2010, Urban Hierarchies and city-growth in the Balkans, *Urban Studies*, 47:2, forthcoming.
- Schumpeter J., 1949, *The theory of economic development*. Cambridge, Harvard University Press, 2nd edition.
- Sharma S., 2003, Persistence and stability in city growth, *Journal of Urban Economics*, 53, 300-320.
- Simon H, 1955, On a class of skew distribution functions *Biometrika*, 44, 425-440
- Singer H.W. 1936, The "courbe des populations": a parallel to Pareto's law. *Economic Journal*, 46, 254-263.
- Song, S., Zhang, K., 2002, Urbanization and City Size Distribution in China, *Urban Studies*, 39, 2317-2327.
- Soo K.T., 2005, Zipf's Law for cities: a cross-country investigation, *Regional Science and Urban Economics*, 35, 239-263.
- Stoev S, Michailidis G, Taqqu M, 2008, Estimating heavy-tail exponents through max self-similarity, draft paper, university of Michigan.
- Weron R, 2001, Levy-stable distributions revisited: tail index > 2 does not exclude the Levy-stable regime, *Journal of Modern Physics C*, vol. 12 no. 2
- Wickens, 1921, *Census of the Commonwealth of Australia*, in H.J.Green, Government Printer, Melbourne.
- Williamson J., 1965, Regional inequality and the process of national development, *Economic Development and Cultural Change*, 4, 3-47.
- Zeyneloglu S. et al, 2005, Methods and Data consideration related to the rank size distributions of settlements, *Acts of the ERSA*, Amsterdam, August.
- Zipf, G.K., 1949, Human Behavior and the Principle of Least Effort, *Addison-Welsey*, Cambridge, MA.

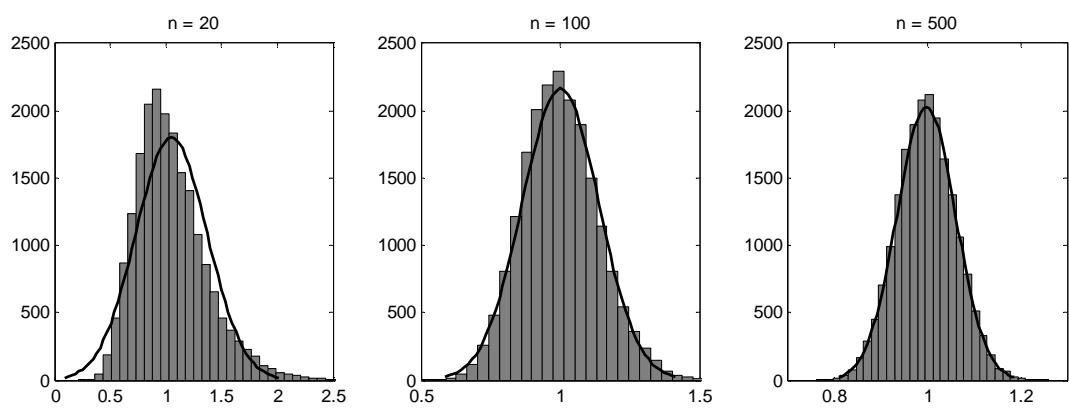
ANNEXES

| Auteurs | Pays | Dates | Modèle | Valeurs du coefficient |
|----------------------------------|---------------------------------------|-------------------------------------|---------------|-------------------------------------|
| Singer (1936) | 7 pays (villes de plus de 2000 h) | Début 20ème siècle | Pareto | 1,15 moyenne (min 0,93 - max 1,64) |
| Rosen & Resnick (1980) | 44 pays (les 50 plus grandes villes) | 1970 | Zipf | 1,13 moyenne (min 0,81 – max 1,96) |
| Malecki (1980) | Etats-Unis (Etats du Midwest) | 1940-1970 | Zipf | 1,06 moyenne (min 1,00 – max 1,54) |
| Parr (1985) | 12 pays (villes de plus de 100 000 h) | 1920-1980 | Zipf | 1,15 moyenne (min 0,75 – max 1,64) |
| Lepetit (1990) | France (villes de plus de 10000 h) | 1750-1780-1836 | Lotka | 0,74 moyenne (min 0,72 – max 0,76) |
| Alperovich (1993) | 15 pays (villes de plus de 100 000 h) | 1984 et 1989 | Zipf | 0,95 moyenne (min 0,75 – max 1,61) |
| Moriconi- Ebrard (1993) | 78 pays (villes de plus de 10000 h) | 1950 à 1980 | Lotka | 1,06 moyenne (min 0,87 – max 1,52) |
| Guérin-Pace (1995) | France (villes de plus de 2000 h) | 1831 et 1982 | Lotka | 0,934 moyenne (min 0,70 – max 1,13) |
| Eaton et Eckstein (1997) | France/Japon (40 plus grandes villes) | France 1876/1980 Japon 1925/1985 | Zipf | 0,984 moyenne (min 0,88 – max 1,03) |
| Black et Henderson (1998) | Etats-Unis (194/247/282 villes) | 1900-1950-1990 | Zipf | 0,89 moyenne (min 0,84 – max 0,97) |
| Brakman et al. (2004) | Pays Bas (42 villes) | 1990 | Lotka | 1,13 moyenne (min 0,70 – max 1,13) |
| Fujita, Krugman, Venables (1999) | Etats-Unis (130 villes) | 1990 | Zipf | 1,00 moyenne (min 1,00 – max 1,00) |
| Gabaix (1999) | Etats-Unis (135 villes) | 1990 | Zipf | 1,00 moyenne (min 1,00 – max 1,00) |
| Dobkins et Ioannides (2000) | Etats-Unis (116/162/392 villes) | 1900-1950-1990 | Zipf | 0,99 moyenne (min 0,95 – max 1,04) |
| Petrakos et Arvanitids (2008) | Grèce (villes de plus de 5000 h) | 1951-1991 | Zipf | 0,96 moyenne (min 0,95 – max 0,98) |
| Soo (2005) | 73 pays (villes de plus de 15000 h) | Dernière année disponible | Zipf | 1,11 moyenne (min 0,73 – max 1,72) |
| Crampton (2005) | 14 pays (taille des villes variable) | Dernière année disponible | Zipf | 0,90 moyenne (min 0,71 – max 1,24) |
| Dimou et Schaffar (2006) | Balkans (villes de plus de 10000 h) | 1981-1991-2001 | Lotka | 0,96 moyenne (min 0,81 – max 1,12) |
| Bosker et al. (2006) | Italie (villes de plus de 10000 h) | 1300-1861 | Zipf | 1,44 moyenne (min 1,24 – max 1,61) |
| Le Gallo et Chasco (2008) | Espagne (villes de plus de 10000 h) | 1900-2001 | Zipf | 0,60 moyenne (min 0,54 – max 0,66) |

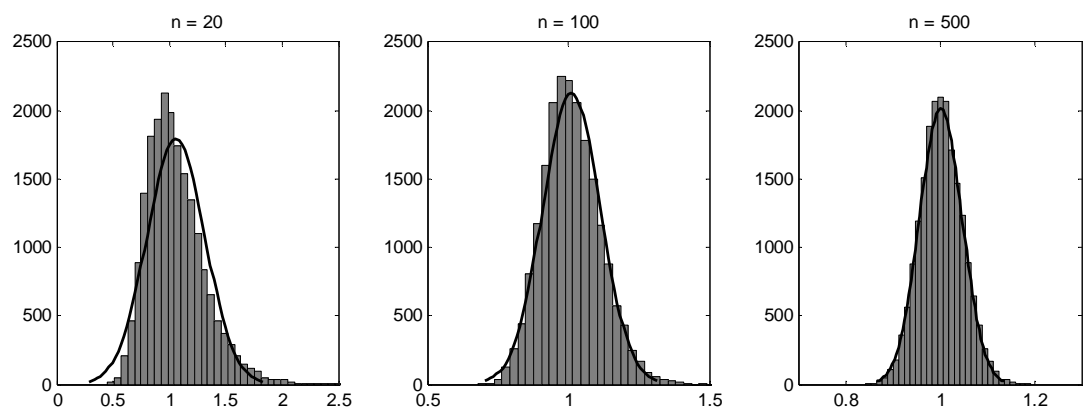
Annexe 1 : Quelques études sur la validité de la loi de Zipf



β par la méthode des MCO

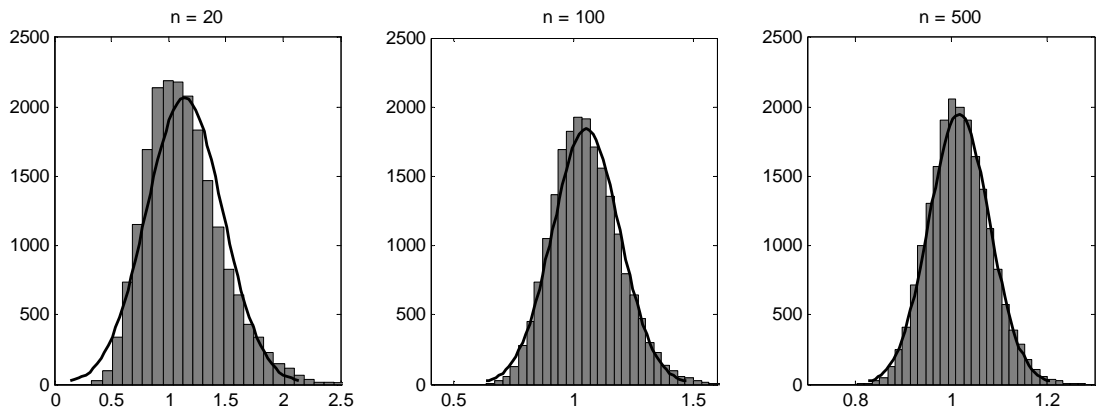


β par la méthode des MCO corrigés

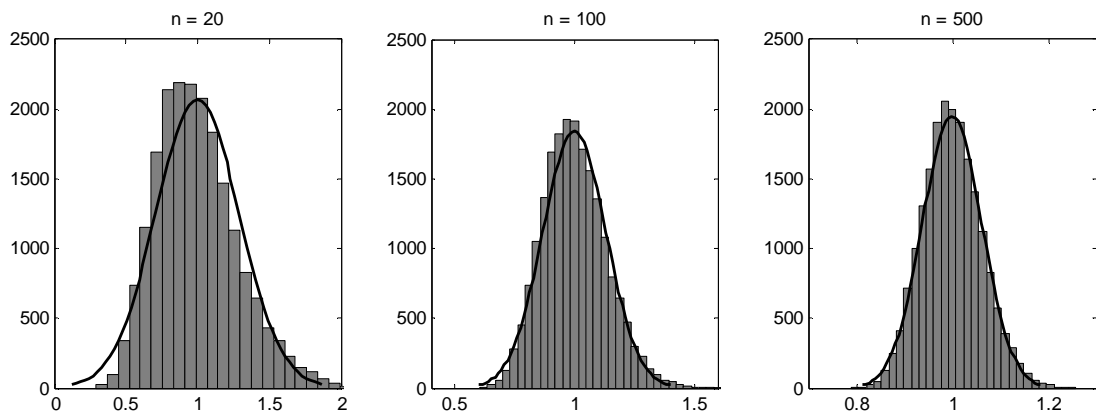


β par la méthode de Hill

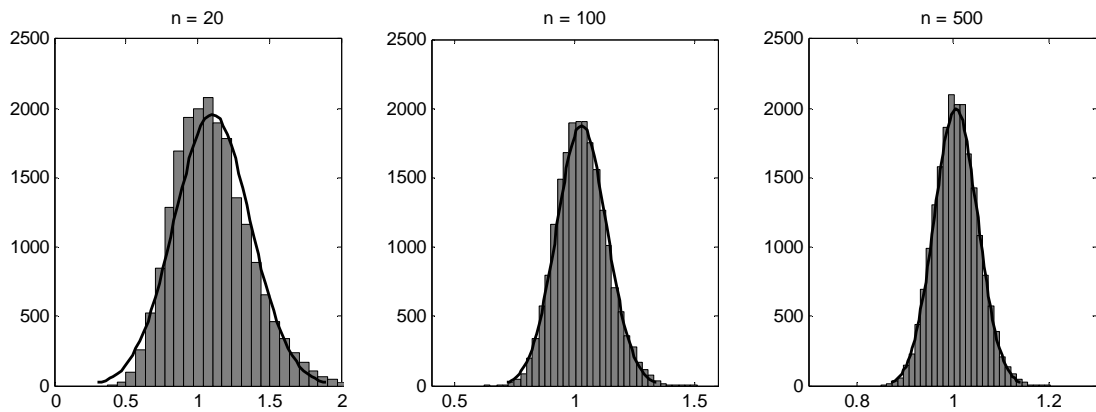
**Annexe 2 : Histogrammes des estimateurs du coefficient de hiérarchisation
obtenus par 20 000 simulations Monte Carlo**



γ par la méthode des MCO



γ corrigé



γ par la méthode des MCG

Annexe 2 (suite) : Histogrammes des estimateurs du coefficient de hiérarchisation obtenus par 20 000 simulations Monte Carlo