LES LOIS DE ZIPF ET DE GIBRAT POUR LES VILLES : UNE INTRODUCTION

Alexandra SCHAFFAR*

Durant ces quinze dernières années, de nombreux travaux en sciences régionales se focalisent sur l'étude des lois de Gibrat et de Zipf pour les villes, deux représentations stylisées de la croissance et des hiérarchies urbaines. Ces lois ne sont pas récentes mais trouvent leurs origines au début du 20ème siècle, lorsque plusieurs chercheurs en sciences sociales investissent les propriétés des distributions scalaires et leurs applications, en s'inspirant de la prolifération exceptionnelle des travaux des mathématiciens sur les lois de distribution (Pearson, 1905 ; Yule, 1911 ; Lotka, 1926 ; Weibull, 1939).

Dans ce contexte transdisciplinaire, un bref article de quatre pages de Auerbach (1913) fait remarquer, en premier, la constance singulière de la valeur du produit de la population d'une ville par son rang, au sein d'un système urbain donné. 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. Auerbach établit, ainsi, un premier modèle rang-taille où le nombre des villes est en relation inverse avec leur poids démographique respectif.

Durant la première moitié du 20ème siècle, hormis une étude comparative de Singer (1936) sur la distribution des villes dans sept pays, le modèle rang-taille reste à la marge des constructions en économie spatiale, comme l'illustre le mépris affiché par Christaller (1933), selon lequel cette relation ne représente qu'*un jeu de chiffres,* dénué de tout fondement théorique. Il faut attendre les travaux de Zipf et surtout leur reformulation par Simon (1955), Beckmann (1958) et Berry (1961) pour que la loi rang-taille apparaisse comme une « *contrainte de comportement des masses* » (Isard, 1960) dont la régularité dans le temps exige une interprétation théorique de l'organisation des systèmes urbains. Leurs contributions confirment, a posteriori, l'intuition de Schumpeter, selon lequel « *il y a peu d'économistes 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 puisse être à l'origine d'une théorie économique complètement novatrice* » (Schumpeter, 1949).

Un nombre conséquent de chercheurs, initialement en géographie, a entrepris une série de travaux empiriques pour tester la validité de la loi de Zipf pour les villes, tout en s'interrogeant sur les interprétations théoriques qui en

* Université de Toulon, LEAD, EA 3163, 83957 La Garde, France ;
schaffar@univ-tln.fr

découlent quant à la structuration des systèmes urbains et la localisation des ménages et des firmes. Ces travaux ont permis d'affiner et d'enrichir les connaissances sur les hiérarchies urbaines, en multipliant les terrains d'études, en proposant des comparaisons internationales ou diachroniques, en utilisant multiples définitions statistiques de la ville, en testant différentes méthodes d'estimation pour limiter les biais d'échantillonnage. Néanmoins, ils ne permettent pas de dégager une vision unanime sur la loi rang-taille, qualifiée de façon controversée de « *mystère urbain* » par Krugman (1996) et « *d'illusion statistique* » par Henderson (1996) ; elle reste, ainsi « *un des faits les plus frappants en sciences sociales contemporaines en général* » (Gabaix et Ioannides, 2004).

Sur le plan théorique, Simon (1955), Robson (1973) puis Pumain (1982) introduisent l'hypothèse que la forme de la distribution des tailles urbaines dépend des processus de changement démographique des villes. Ils amorcent, par ce biais, les approches sur la croissance urbaine qui s'articulent autour de la relation entre la croissance des villes et leur taille démographique, synonyme d'un certain nombre de caractéristiques économiques, telles que le poids des externalités (positives ou négatives) d'agglomération, la concentration du capital humain ou le périmètre de débordement des connaissances codifiées. L'absence d'une telle relation valide la loi de Gibrat pour les villes et met en cause toute une série de constructions théoriques en sciences régionales qui reposent sur les effets d'agglomération. Son affirmation conduit, à l'inverse, à l'hypothèse d'une convergence des villes vers une taille urbaine « optimale », objectif convoité depuis la cité idéale de Platon jusqu'aux modèles récents de Henderson (1974) qui posent les fondements de l'économie urbaine contemporaine.

Ce numéro spécial de *Région et Développement* comprend un ensemble de contributions qui permettent à la fois de présenter l'état de l'art sur les lois de Zipf et de Gibrat pour les villes, apporter des éclairages sur les dernières avancées théoriques et méthodologiques et ouvrir le débat entre des chercheurs de disciplines différentes qui se partagent l'ambition d'analyser les hiérarchies urbaines et les processus de changement démographique dans les villes.

Dans la première contribution, **Gilles Duranton** propose une relecture des différents modèles de croissance urbaine en mettant en évidence le clivage fondamental entre les modèles classiques et les modèles de croissance aléatoire. Les premiers permettent d'étudier le changement démographique urbain, tout en considérant une série de sujets connexes tels que la spécialisation économique des villes, tandis que les seconds cherchent à générer une distribution rang-taille des villes conforme à la loi de Zipf sur la base d'hypothèses plus originales. Gilles Duranton montre que, sous certaines conditions restrictives, ces modèles peuvent coexister dans le sens où les premiers traitent les tendances, tandis que les seconds les chocs. Un ensemble de travaux récents s'inscrivent dans cette dynamique.

Dans une deuxième contribution, **Denise Pumain** revisite les travaux sur la loi de Zipf, afin de dégager les connaissances cumulables dans des disciplines différentes mais connexes telles que l'économie urbaine et la géographie. En apportant un regard critique sur certaines utilisations abusives de la loi rang-taille et en rappelant un ensemble de paradoxes qui en découlent, Denise Pumain propose une exploration des processus de croissance des villes et d'évolution des hiérarchies urbaines qui s'appuie sur une théorie géographique évolutive déduite de la comparaison d'un grand nombre de systèmes urbains régionaux et nationaux. Dans son article, Denise Pumain introduit l'hypothèse de la croissance urbaine en tant que processus complexe, composé d'une part d'une forme d'évolution générale propre à tout processus d'urbanisation et d'autre part des histoires locales.

En adoptant une approche assez proche de la précédente, dans laquelle les systèmes urbains apparaissent comme des systèmes complexes déterminés par des effets d'échelle et de variété, **Tarsha Eason** et **Ahjond Garmestani** examinent la relation entre la taille des villes, le taux de croissance urbaine et les facteurs qui ont un impact significatif sur la résilience et la dynamique démographique des systèmes urbains. En s'appuyant sur des données empiriques concernant le système urbain du sud-est des Etats-Unis, ils montrent que la loi de Gibrat n'est pas valide pour les villes les plus dynamiques sur le plan démographique. A l'inverse, dans ces villes la croissance urbaine semble fortement corrélée avec un certain nombre de caractéristiques économiques telles que le revenu par habitant et le niveau du capital humain localisé.

Dans sa contribution, **Boris Portnov** propose d'étudier la validité de la loi de Gibrat pour les villes, en s'appuyant sur une étude empirique des évolutions démographiques de 2189 aires urbaines européennes. Il émet l'hypothèse que la loi de Gibrat est validée lorsqu'on utilise le niveau d'agrégation le plus élevé, c'est-à-dire les aires urbaines. Cependant, lorsqu'on utilise des niveaux d'agrégation plus fins, tels que les villes administratives ou les quartiers, ceci n'est plus le cas, car les effets de taille et les mécanismes économiques conditionnent davantage les mouvements démographiques et les migrations. Au niveau d'agrégation le plus élevé, les mouvements démographiques intra-urbains enregistrés se compensent, ce qui peut induire une interprétation différente et/ou erronée de la relation entre la croissance et la taille urbaine.

Kristian Giesen et Jen Sudekum examinent de leur côté la forme distributionnelle des villes françaises, en considérant non pas une distribution tronquée à une taille urbaine minimale mais l'ensemble des localités habitées. Ils montrent que cette distribution n'épouse pas la loi de Zipf, mais plutôt une distribution double Paréto lognormale (DPLN). Ils arrivent, par ce biais, à concilier la littérature sur la loi de Zipf, animée par des travaux qui s'appuient sur des distributions de grandes villes et certaines approches plus ou moins récentes selon lesquelles la distribution réelle adopte davantage les allures d'une lognormale, lorsqu'on considère l'ensemble de la population des villes. Kristian Giesen et Jens Sudekum construisent un modèle de croissance urbaine microéconomique adéquat, qui débouche sur une distribution DPLN. Un trait intéressant dans leur article est l'introduction de « l'âge des villes » comme élément explicatif de la forme de la distribution des tailles urbaines.

L'âge des villes se trouve au cœur de la contribution de **Lena Sanders** qui propose une exploration des travaux des économistes, des géographes et des archéologues sur la loi de Zipf. En puisant dans cette diversité disciplinaire des analyses qui se focalisent sur l'organisation des systèmes urbains, elle montre que la relation rang-taille ne doit être utilisée que comme un critère parmi d'autres pour étudier les subtilités des hiérarchies urbaines. Les pratiques des différentes disciplines sont source d'inspiration non seulement pour diversifier les méthodes d'investigations, mais également pour expliquer la récurrence ou pas d'une relation rang-taille et décrire empiriquement les caractéristiques d'un système de peuplement.

Rafael Gonzalez-Val entame, de son côté, une analyse des questions méthodologiques qui pèsent sur l'étude de la loi de Zipf. Dans sa contribution, il parcourt trois séries de choix méthodologiques qui doivent être tranchés par toute étude sur le sujet : ils concernent la définition de la ville, la construction de l'échantillon et la méthode d'estimation du coefficient de hiérarchisation. A partir de données sur les villes américaines, il montre que la méthode du maximum de vraisemblance (Hill) est la plus adéquate pour calculer le coefficient de hiérarchisation sous la condition stricte que la distribution des tailles urbaines soit une distribution de Pareto. En cas de doute, il est préférable d'utiliser la méthode corrigée des moindres carrés ordinaires qui souffre moins des biais liés à une certaine convexité de la relation rang-taille.

Daniel Arribas-Bel, **Fernando Sanz Gracia** et **Domingo Ximenez de Embun** étudient la validité des lois de Zipf et de Gibrat pour le système urbain australien. Ils ne confirment ni l'une ni l'autre, mais émettent l'hypothèse que la distribution rang-taille des villes australiennes reste fortement soumise à des effets de distance et aux caractéristiques d'un continent fortement sous-peuplé. Ils introduisent une analyse exploratoire des données spatiales, en s'appuyant sur certains outils d'économétrie spatiale, afin de détecter des effets de voisinage dans la taille des villes et leur croissance démographique. Ils montrent l'existence d'effets d'auto-corrélation spatiale dans les processus de croissance urbaine mais non pas dans la distribution rang-taille des villes.

Enfin, dans la dernière contribution de ce numéro spécial de *Région et Développement*, **Alexandra Schaffar** utilise un modèle rang-taille des villes sur des données de panel pour étudier la validité de la loi de Zipf. En introduisant la dimension géo-localisée des données, Alexandra Schaffar cherche à détecter les effets d'auto-corrélation spatiale dans la distribution rang-taille des villes chinoises et indiennes. Elle montre qu'en Chine, la prise en compte des effets d'auto-corrélation spatiale ne modifie pas fondamentalement les conclusions du modèle a-spatial tout en permettant de l'affiner : la distribution rang-taille des villes suit la loi de Zipf dans le long terme. En Inde, le modèle a-spatial conduit à un rejet de la loi de Zipf. L'introduction de la dimension spatiale conduit à valider la loi de Zipf sur le long terme.

L'ensemble de ces contributions permet de mettre en lumière les nouveaux défis des analyses sur la croissance et les hiérarchies urbaines. Le premier défi est d'ordre épistémologique : il concerne le croisement de disciplines autour d'un même objet d'étude, celui de l'organisation et la structuration des systèmes urbains. Le second défi est d'ordre méthodologique : il concerne la définition même de la ville et du système urbain et par la suite les méthodes utilisées pour étudier les lois de Zipf et de Gibrat pour les villes, qui restent deux faits stylisés majeurs dans la formation des hiérarchies urbaines. Le troisième défi est théorique : il faut s'interroger sur les modèles de localisation des firmes et des ménages et des mécanismes économiques sous-jacents qui expliquent l'apparition de telles régularités dans l'espace et dans le temps.

REFERENCES

Auerbach, F., 1913, Das Gesetz der Bevölkerungskonzentration, reproduit dans *Regional Science and Urban Economics*, 31, pp.601-615.

Beckmann M.J. 1958, City hierarchies and the distribution of city sizes, *Economic Development and Cultural Change*, 6, pp.243-248.

Berry B.J., 1964, Cities as systems within systems of cities, *Papers of the Regional Science Association*, 13, pp.147-163.

Christaller W., 1933, *Central Places in Southern Germany*, Englewood Cliffs, New Jersey.

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, pp.2341-2376.

Gibrat, R., 1931, *Les inégalités économiques*. Librairie du Recueil Sirey, Paris.

Goodrich P., 1926, The statistical relationship between population and the city plan, in Burgess N. (eds) *The Urban Community*, University of Chicago, pp.144-156.

Henderson V., 1974, The size and types of cities, *American Economic Review*, 64, pp.640-656.

Henderson V., 1988, *Urban Development: theory fact and illusion*, Oxford University Press.

Isard W., 1956, *Location and space economy*, MIT Press, Cambridge.

Krugman P., 1996a, Confronting the Mistery of Urban Hierarchy, *Journal of the Japanese and the International Economies*, 10, pp.399-418.

Lotka A., 1926, The frequency distribution of scientific productivity, *Journal of the Washington Academy of Science*, 16, pp.317-323.

Parr J., 1976, A class of deviations from rank-size regularity: Three interpretations, *Regional Studies* 10(3), pp.285–292.

Pumain D., 1982, *La dynamique des Villes*, Economica, Paris.

Robson, B. T. 1973. *Urban Growth, an approach*. London: Methuen.

Schumpeter J., 1949, Economic Theory and Entrepreneurial History, reprinted in Schumpeter, J., 1989, *Essays on Entrepreneurs*, *Innovations*, *Business*, Transaction Publishers, N.Brunswick.

Simon, H., 1955, On a class of skew distribution functions, *Biometrika*, 44, pp.425-440.

Singer H.,1936, The “courbe des populations”: a parallel to Pareto’s law, *The Economic Journal*, vol.XLVI, n°182, pp.254-263.

Yule G. (1911) [1944], *The statistical study of library vocabulary*, Cambridge University Press.

Zipf G.K., 1941, *National Unity and Disunity: the Nation as a Bio-Social Organism*, Principia Press, Bloomington.

Zipf G.K., 1949, *Human Behavior and the Principle of Least Effort*, Addison-Welsey, Cambridge, MA.