Departamento de Economía Aplicada

DOCUMENTOS
DE
TRABAJO

UNIVERSIDAD DE JAÉN

A definition of urban nucleus based on a rank-size test

WP 0104/ No. 22

Pablo Brañas Garza, Javier Rodero Cosano and
Francisco Alcalá Olid

Dirección para comentarios y críticas

Departamento de Economía Aplicada
Universidad de Jaén
Paraje las Lagunillas s/n
23071 Jaén
email: pbg@ujaen.es

A definition of urban nucleus based on a rank-size test

Pablo Brañas Garza Javier Rodero Cosano
Francisco Alcalá Olid

Department of Economics, University of Jaén, 23071 Jaén, Spain.
Tel: +34 953 01 22 29. Fax: +34 953 01 22 22. E-mail: pbg@ujaen.es

---

Abstract

This paper tries to be an application of Urzúa's (2000) statistical method to test the Zipf Law. Using this method we develop a definition of "urban" (non-rural) nucleus that allows us to approach the threshold between the urban and rural areas. This algorithm is empirically tested for the Spanish case along the XX century.

Key words: urban size, LMZ, LM tests, Zipf Law
JEL Class.: C12, R12

Usually, urban economists use to ask about the definition of *minimum size* (MS, hereafter) of any city to denominate it as "urban"; this size is the threshold value between rural and urban areas. We see that several authors (Zipf, 1949; Mills and Hamilton, 1984; Krugman, 1996; Brakman et al., 1999, among others) use different *ad hoc* sizes, without any clear theoretical or empirical foundation. We propose a MS definition based on the maximum level of acceptance (non-rejection) of the LM test supplied by Urzúa in this journal (Urzúa, 2000). That is, under the generally accepted Zipf Law, we consider as urban only the group of cities that fit better under this law. The paper is structured as follows: the first section deals with the Zipf Law (ZL, hereafter), the LMZ test and its maximization; the second analyzes an example with Spanish data; section third concludes.

1 Zipf Law, Lagrange Test and its maximization

ZL is a special case of a general rank-size distribution, on which the parameter "q" is equal to one. The size of each city is its population and the largest city

Preprint submitted to Elsevier Preprint *26 June 2001*

is given rank 1, the second-largest one rank 2, and so on. Under ZL, the largest city is r times as large as the r^{th} largest city (see Brakman et al., 1999; Fujita et al., 1999; Gabaix, 1999).

Under the hypothesis that ZL is a particular class of Pareto distribution, Urzúa (2000) proposes a test to contrast the fitness of this distribution to the real data. If the test is not rejected, the ZL is fulfilled. The Langrange multiplier associated to the test is:

$$LMZ = 4n[z_1^2 + 6z_1z_2 + 12z_2^2]; \tag{1}$$

$$\text{with,} \qquad z_1 \equiv 1 - \frac{1}{n}\sum_{i=1}^{n} \ln \frac{x_i}{x_{(n)}} \quad \text{and} \quad z_2 \equiv \frac{1}{2} - \frac{1}{n}\sum_{i=1}^{n} \frac{x_{(n)}}{x_i}.$$

Under the null hypothesis, LMZ is asymptotically distributed as chi-square with two degrees of freedom, χ_2^2.

Our objective is to derive a MS estimate based on the minimization of LMZ (or maximization of probability of non-rejection, so maximize the probability that ZL holds). Through recursive computer iterations (changing MS step by step) we got the minimum value of LMZ and therefore a definition of MS, that is the best fit for the ZL[1]. Under the Zipf postulates, this MS (ZMS hereafter) can be interpreted as the division point between the cities that follows a hierarchical structure (the urban network) and the rural world. So, this paper determines endogenously this point.

Finally, we have to point that, given that LMZ is asymptotically distributed as χ_2^2 only for $n > 50$, to ensure the accuracy of the algorithm results, the country must be large enough, i.e. with at least 50 cities.

2 An example with Spanish data

For an empirical example of this algorithm we use Spanish data of cities provided by the INE (National Statistic Institute). This database covers the whole XX century by decades, starting at 1900. Table 1 resumes the results.

Where n, based on ZMS definition given above, is the resulting number of "urban" cities and $R \equiv N - n$, the number of "rural" ones (N is the total number of population centers; in Spain, $N \simeq 8000$).

---

[1] The working paper shows this routine that can be provided on request.

Table 1
Values for ZMS in Spain: 1900-2000

Year	ZMS	n	Year	ZMS	n	Year	ZMS	n	Year	ZMS	n
1900	31609	29	1930	45607	30	1960	25211	117	1991	20499	267
1910	35072	30	1940	61188	30	1970	22087	169	1998	26197	217
1920	37967	32	1950	81182	25	1980	24902	194			

Figure 1 shows the temporal values of ZMS_t, n_t and the resulting percentage of "urban" population, U_t, $(t = 1900, \ldots, 1998)$ in Spain along the XX century.

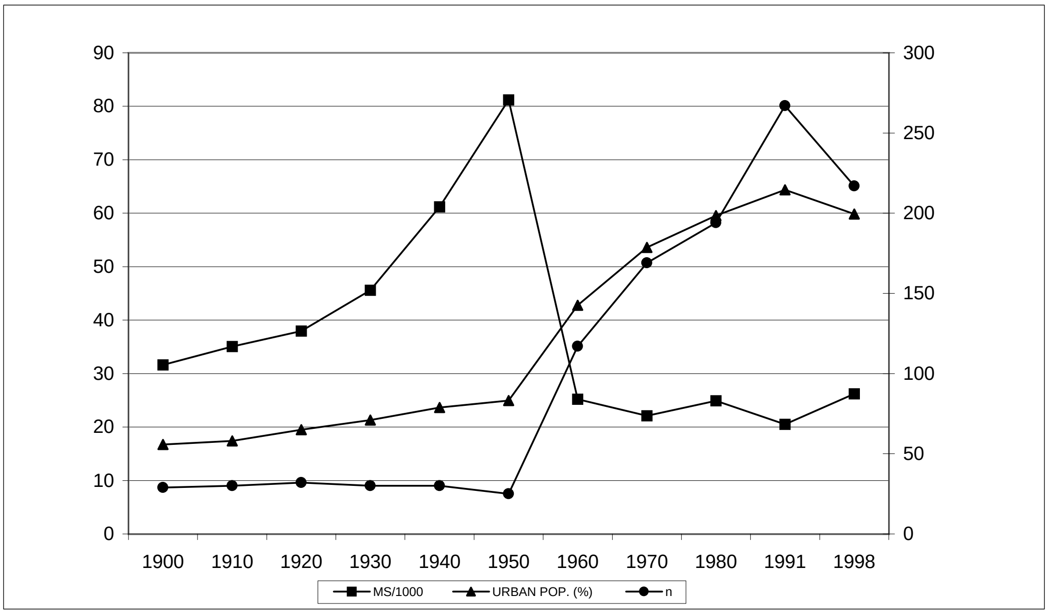

Fig. 1. Temporal path of ZMS, n and % of urban pop.

There is a clear decreasing trend in ZMS along time with a rise in the number of "urban" nucleus. Also, we see that the percentage of "urban" inhabitants is also growing. But we see a clear change of trend in the 60's. Until that date the increasing urban population tends to concentrate in a small number of cities (the n value even decreases slightly). After that date, the urban structure includes a rise number of cities parallel to the urbanization of the population.

A possible explanation for these facts is the increase in public infrastructure, specially that one devoted to roads, along the century. Due to the political situation of the country, the communication investment has follow a centralized pattern, creating, as in France, the proper conditions for the appearance of a unique system of cities. Also, the dramatic fall of commuting and transport costs (of raw goods) has allowed the generation of this system in a mid-size country like Spain. In such a structure, most cities do not need to be large to allow access to an ample "menu" of goods & services.

But, the congestion on these infrastructures could be an explanation for 90–98's break of trend. As a consequence of the entrance of Spain in the EU, the movement of commodities has risen exponentially causing congestion in the roads. When congestion increase, commuting and transportation costs also rise, so domestic production begin to be profitable and local markets (and cities) become larger.

3 Conclusions

This paper has intended to be an extension of Urzúa (2000). Through his LMZ, we have developed an algorithm that permits a definition of urban cities. This algorithm is consistent only for mid-size or big countries. The results from an application to the Spanish case seem to indicate that, the optimal minimum size has fallen and the number of urban cities and percentage of urban population have increased dramatically along the XX century.

References

Brakman, S., Garretsen, H., Marrewijk, C. V., Den Berg, M. V., 1999. The return of Zipf: Towards a further understanding of rank–size distributions. Journal of Regional Science 39, 183–213.

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 104, 739–767.

Krugman, P., 1996. The Self–Organizing Economy. Blackwell, Cambridge.

Mills, E. S., Hamilton, B. W., 1984. Urban Economics. Scott Foresman, Glenview, IL., Ch. Studies in the Structure of the Urban Economy.

Urzúa, C. M., 2000. A simple and efficient test for zipf's law. Economic Letters 66, 257–260.

Zipf, G. K., 1949. Human Behaviour and the Principle of Least Effort. Harvard University Press, Harvard, reimpresión en Hafner Publishing Co.: N.Y., 1972.

A Mathematica Code used to Estimate the ZMS value

```
PARTE II: ESTIMACION DE LOS TAMANOS DE CORTE OPTIMOS PARA CADA ANO
 Needs["Statistics‘NormalDistribution‘"]
 Needs["Graphics‘MultipleListPlot‘"]
 Pob = ReadList["C:\Pablo\meta.txt",{Number,
 Number, Number, Number, Number, Number, Number, Number,
 Number, Number, Number}];
 pobmin = 25;
 anos = {1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980, 1991, 1998};
 pvalue = List[{10, 15, 20, 25, 30, 50, 100, 200, 10000}];
 P1 = {6.19, 6.14, 6.09, 6.08, 6.03, 5.98, 5.98, 5.99,
       Quantile[ChiSquareDistribution[2], 0.95]};
 For[j = 1, j <= 11, \(j++\), Print[];
   Print["EN EL ANO DEL SENOR DE ",
     anos[[j]], " OBTENEMOS LOS SIGUIENTES RESULTADOS:"];
   Print[
   "================================================================"];
   Print[];
   Pop = Select[Pob[[All, j]], #1 > 0 &];
   ytam = Length[Pop];
   l = 2;
   TZ = {Pop[[1]]};
   LZ = {};
   LR = {};
   NZ = {};
   (*Print[TZ[1]];\)*)
   LiR_m = [Infinity];
   For[i = pobmin, i <= ytam, i++\,
     If[Pop[\([i]\)] < TZ[[l - 1]],
       TZ = Append[TZ, Pop[\([i]\)]];
       Pob_crit = Take[Pop, i];
       n = Length[Pob_crit];
       NZ = Append[NZ, n];
       x_n = Min[{Pob\_crit}];
       z_1 = 1 - (1/n) Log[Apply[Times, Pob_crit]/x_n^n];
       div = 1/Pob_crit;
       z_2 = 1/2 - (1/n)(x_n) Apply[Plus, div];
       LMZ = 4  n ((z\1^2 + 6 ( z_1) z_2 + 12 z_2^2));
       If[n <= 10, pos = 1,
         If[n <= 15, pos = 2,
           If[n <= \ 20, pos = 3,
             If[n <= \ 25, pos = 4,
               If[n\  <= \ 30, pos = 5,
                 If[n <= \ 50, pos = 6,
                   If[n <= \ 100, \ pos = 7,
```

```
         If[n <= 200, pos = 8,
          pos = 9]]]]]]]];
  pt1 = Extract[P1, {pos}];
  LMR = LMZ/pt1;
  (*\(Print[N[LMZ, 12], "n= ", n, " l= ", l, " i= ", i, " popi= ",
     Pop[[i]], " tz= ", TZ[l - 1, j], " Indice Rel: ", LMR];*)
  LZ = Append[LZ, N[LMZ, 6]];
  LR = Append[LR, N[LMR, 6]];
  If[LMR < LiR_m,
  LiR_m = LMR;
  NiZ_m = n;
  TiZ_m = x_n;]
  l++;
 ]
];
For[i = 1; potot = 0, i <= \ NiZ_m, i++, potot += Pop[[i]]];
(* Ahora empiezo a presentar los resultados*)
Print[N[LiR_m, 5], \ ", para n= ", NiZ_m, " un tamano de corte de ",
TiZ_m, " y una poblacion total urbana de ", potot, " personas"];
tabla = {};
tablalim = {{0, 6.2}};
For[i = 1, i <= Length[LZ], i++,
tabla = Append[tabla, {NZ[[i]], LZ[[i]]}];]
For[i = 1, i <= Length[P1], i++,
tablalim = Append[tablalim, {pvalue[[1, i]], P1[[i]]}];]
(*Print[tablalim];*)
ListPlot[tabla, ImageSize -> 72*8, PlotStyle -> {PointSize[0.003]}];
MultipleListPlot[tabla, tablalim,
 {ImageSize -> 72*8, PlotRange -> {{0, 1300}, {0, 30}},
 PlotStyle -> {PointSize[0.003]}, PlotJoined -> {False, True}}];
(*Calculo los intervalos de aceptacion*)
intervalos = {};
i = 1;
(*Print[LR, NZ];*)
While[i <= Length[NZ],
 If[LR[[i]] <= 1,
  liminf = NZ[[i]];
  k = i;
  If[k < Length[NZ],
   While[k < Length[LZ] && LR[[k]] <= 1,
    k++]];
  If[k == Length[LZ] && LR[[k]] <= 1, k++;];
  limsup = NZ[[k - 1]];
  i = k;
  intervalos = Append[intervalos, {liminf, limsup}];
 ];
 i++;
```

```
  ];
  Print["Los intervalos de aceptacin son: ", intervalos];
  For[i = 1; pobtam = 0, i <= ytam, i++, pobtam += Pop[[i]]];
  For[i = 1; poburb = 0, i <= limsup, i++, poburb += Pop[[i]]];
  Print["El porcentaje de poblacion urbana es ", N[100*potot/pobtam,2],
  " % "];
  Print["La poblacion urbanizable son ", poburb,
  "habitantes, es decir un ", N[100*poburb/pobtam, 2],"% del total"];
]
```

SERIES DE DOCUMENTOS DE TRABAJO PUBLICADOS

WP 9801/Nº 1
PROPUESTA DE UN ANÁLISIS ECONOMÉTRICO PARA EL ESTUDIO DEL PRECIO DE LA VIVIENDA URBANA
Pablo Brañas Garza, Pablo Fernández-Álvarez y José Mª Caridad y Ocerin

WP 9802/Nº 2
UN ANÁLISIS DEL CRECIMIENTO Y LA CONVERGENCIA DE LA ECONOMÍA ANDALUZA ENTRE 1985 Y 1995
José García Roa

WP 9803/Nº 3
PHYSICAL AND NOT SO PHYSICAL DISTANCES IN A SIMPLE URBAN MODEL: AN ANALYSIS
Pablo Brañas Garza, Javier Rodero Cosano y Joan Carles Martori

WP 9804/Nº 4
UNA EVALUACIÓN DEL CAMBIO DE ESTRATEGIA DE LA POLÍTICA MONETARIA EN ESPAÑA: PERSPECTIVAS DE FUTURO
Francisco Alcalá Olid y Antonio Martín Mesa

WP 9805/Nº 5
URBAN MICROECONOMICS WITHOUT MUTH-MILLS: A NEW THEORETICAL FRAME
Javier Rodero Cosano, Pablo Brañas Garza e Inmaculada Fernández Piñar

WP 9806/Nº 6
LAS EXTERNALIDADES URBANAS: ENTRE ALPEROVICH Y FUJITA
Pablo Brañas Garza, Alejandro Lorca Corrons, Javier Rodero Cosano y Mª Angustias Dávila Vargas-Machuca

WP 9807/Nº 7
LA ECONOMIA ISLÁMICA Y SUS CONTRATOS: UNA PANORÁMICA
Pablo Brañas Garza, Alejandro Lorca Corrons y Javier Rodero Cosano

WP 9808/Nº8
SIZE, PROFITABILITY AND AGENCY PROBLEMS IN PROFIT LOSS SHARING IN ISLAMIC FINANCE
Humayon A. Dar, David I. Harvey y John R. Presley

WP 9901/Nº9
CAPITAL HUMANO Y CRECIMIENTO EN EL MEDITERRÁNEO: ¿SPILLOVERS O DETERMINISMO GEOGRÁFICO
Javier Rodero Cosano, Pablo Brañas Garza, Mª Lucia Cabañes Argudo y Alejandro V. Lorca Corrons

WP 9902/Nº10
SOBRE EL RUIDO Y SU PERCEPCIÓN: UNA APROXIMACIÓN EXPERIMENTAL
Pablo Brañas Garza; M. D. Alcántara Moral y Javier Rodero Cosano

WP 9903/Nº11
CRECIMIENTO ECONÓMICO ENDÓGENO Y CAPITAL PÚBLICO DESDE UNA PERSPECTIVA REGIONAL: UNA APROXIMACIÓN
Diego Martínez López

WP 0001/Nº12
DIFFERENT PATHS OF URBAN AGGLOMERATION IN SPANISH REGIONS: EVIDENCE FROM 1960-1998
Pablo Brañas Garza y Francisco Alcalá Olid

WP 0002/Nº13
IS THERE ANY RELATIONSHIP BETWEEN PUBLIC INVESTMENT AND ECONOMIC GROWTH IN THE SPANISH REGIONS?
Diego Martínez López

WP 0003/Nº14
CONTRACTS IN THE AGRICULTULTURAL SECTOR WITH MORAL HAZARD AND HIDDEN INFORMATION: SPECULATIONS, TRUTHS AND RISK-SHARING
Francisca Jiménez Jiménez

WP 0004/Nº15
HOTELLING AND THE OLYMPUS: MODELLING DIFFERENCES IN RELIGIOUS PRICES
Javier Rodero Cosano y Pablo Brañas Garza

WP 0005/Nº16
AN EMPIRICAL MEASUREMENT OF THE EFFECTS OF EXTERNALITIES ON LOCATION CHOICE
Pablo Brañas Garza y Javier Rodero Cosano

WP 0006/Nº17
EL ENDEUDAMIENTO A LARGO PLAZO DE LA HACIENDA PÚBLICA ANDALUZA: UNA VISIÓN PANORÁMICA
Diego Martínez López

WP 0007/Nº18
CRECIMIENTO Y SISTEMAS DE CIUDADES: UN MODELO DE DIFERENCIACIÓN DE PRODUCTO
Jose Luis Sáez Lozano y Pablo Brañas Garza

WP 0101/Nº19
AGLOMERACIÓN URBANA EN ESPAÑA (1900-2000): "ESTIMACIÓN RANK-SIZE VS. TEST LM"
Pablo Brañas Garza y Francisco Alcalá Olid

WP 0102/Nº 20
INTRODUCCIÓN A LOS EFECTOS DEL CAPITAL PÚBLICO DESDE UNA PERSPECTIVA DUAL
Diego Martínez López

WP 0103/Nº 21
MEASURING COMPETITION AMONG STUDENTS THROUGH EXPERIMENTAL BEAUTY CONTEST GAMES: AN OVERVIEW OF RESULTS
Pablo Brañas Garza, Francisca Jiménez Jiménez, Virtudes Alba Fernández y Javier Rodero Cosano

WP 0104 /Nº 22
A DEFINITION OF URBAN NUCLEUS BASED ON A RANK-SIZE TEST
Pablo Brañas Garza, Javier Rodero Cosano y Francisco Alcalá Olid