Economics Letters 133 (2015) 10–13

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

The pan-European population distribution across consistently defined functional urban areas

Kurt Schmidheiny [a], Jens Suedekum [b,*]

[a] University of Basel, Faculty of Business and Economics, Peter Merian-Weg 6, CH - 4002 Basel, Switzerland

[b] Heinrich-Heine-University, Duesseldorf Institute for Competition Economics (DICE), Universitätsstrasse 1, D–40225 Duesseldorf, Germany

HIGHLIGHTS

- We analyze the first data set on consistently defined functional urban areas in Europe.
- We compare the European to the US urban system.
- City sizes in Europe do not follow a power law.
- The largest cities in Europe are "too small" to follow Zipf's law.

ARTICLE INFO

Article history:
Received 22 April 2015
Received in revised form
2 May 2015
Accepted 8 May 2015
Available online 16 May 2015

JEL classification:
R11
R12

Keywords:
City size distributions
Zipf's law
Functional urban areas
Urban systems

ABSTRACT

We analyze the first data set on consistently defined functional urban areas in Europe and compare the European to the US urban system. City sizes in Europe do not follow a power law: the largest cities are "too small" to follow Zipf's law.

1. Introduction

Beginning with the seminal contributions by Auerbach (1913) and Zipf (1949), there is a long literature on the distribution of population across cities. Virtually all of this research is concerned with cities of the same country. Gabaix (1999) focuses on the United States (US) and shows that population sizes across metropolitan statistical areas (MSAs) closely follow a Pareto distribution with shape parameter close to minus one, also known as Zipf's law. Further studies on the US city size distribution and the underlying urban growth process include Eeckhout (2004), Overman and Ioannides (2003), and Black and Henderson (1999). Focusing on other countries, Eaton and Eckstein (1999) obtain evidence for France and Japan, and Giesen and Suedekum (2011) for Germany respectively, which is consistent with the US experience.

Much less is known about city sizes in a wider context than the nation state, however, even though national borders are steadily losing significance in the ongoing process of economic globalization. The reason is that "cities" are usually not consistently classified; instead, each country adopts its own methods of defining urban areas and delineating their boundaries according to administrative or economic criteria. This is even true in Europe, where official approaches and city definitions differ widely across countries. For this research, we use novel and unexplored data, which allow for a harmonized approach to defining urban areas in 31 European countries and, with the same methodology, in the US. Our goal is to address the pan-European distribution of city sizes, and to compare the European to the American urban system.

* Corresponding author. Tel.: +49 0 211 8811622.
E-mail addresses: kurt.schmidheiny@unibas.ch (K. Schmidheiny), suedekum@dice.hhu.de (J. Suedekum).

http://dx.doi.org/10.1016/j.econlet.2015.05.013

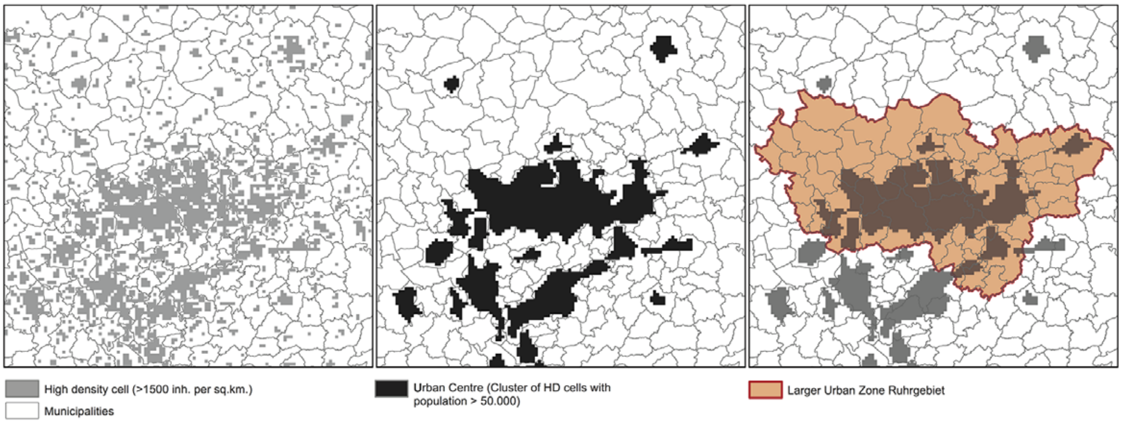

Fig. 1. Construction of the Ruhrgebiet (Germany) functional urban area. *Notes*: The left panel shows the high-density cells with more than 1500 inhabitants per square kilometer and administrative municipal boundaries. The middle panel illustrates the construction of urban centers with a total population of more than 50,000 inhabitants. The right panel shows the construction of the larger urban zone based on bilateral commuting flows. Source: European Commission, Directorate-General Regional and Urban Policy.

2. Data

The novel data stem from a collaboration of the European Commission (EC), see Dijkstra and Poelman (2012) and the Organization for Economic Co-operation and Development (OECD), see Brezzi et al. (2012). The EC–OECD definition of functional urban areas proceeds in three steps: Step 1 partitions the European surface into 1 km^2 grid cells and identifies *high-density cells* with a population density greater than 1500 inhabitants per km^2 based on categorized satellite images. Step 2 generates clusters of contiguous (sharing at least one border) high-density cells. Low-density cells encircled by high-density cells are added. Clusters with a total population of at least 50,000 inhabitants are identified as *urban centers*. Step 3 uses administrative data to calculate commuting flows from local administrative units (municipalities) into urban centers. Local administrative units with 15% of employed persons working in an urban center are assigned to the urban center. A contiguous set of assigned local administrative units form a *larger urban zone*. Non-contiguous local urban centers with bilateral commuting flows of more than 15% of employed persons are combined into a polycentric larger urban zone.

Fig. 1 provides an example, where the single panels illustrate the three steps for the case of the Ruhr area (*Ruhrgebiet*) in Germany. Table 1 gives an overview of the European urban hierarchy across the resulting 692 functional urban areas in Europe in the year 2006.

This EC–OECD definition of urban areas has important advantages over using population data for administratively defined cities. The algorithm, for example, identifies the *Ruhrgebiet* as the largest German city. This larger urban zone comprises the four administrative cities Duisburg, Essen, Bochum, and Dortmund, which form a contiguously populated cluster but are reported as individual cities in traditional data. The algorithm also assigns larger urban zones across national borders, for example Geneva and Basel, which consist of urban centers not only in Switzerland but extend into France and Germany/France, respectively. Finally, the EC–OECD data is complimentary to other approaches that draw on fine-grained satellite images, such as Rozenfeld et al. (2008, 2011) who build on population clusters, or Henderson et al. (2012) who exploit data on night-lights. While those data also ignore artificial administrative boundaries when defining "cities", they neglect economic linkages across cities such as commuting flows. The EC–OECD data considers such linkages, and thus combines the key advantages of the satellite-based approach and the traditional delineations of functional urban areas. To the best of our knowledge, this paper is the first to analyze this novel data source for city sizes.

Table 1
Population size (number of inhabitants) across 692 urban areas in Europe in 2006.

Rank	Urban area name	Population size
1	Paris (FR)	11,370,846
2	London (UK)	11,256,669
3	Madrid (ES)	5,993,683
4	Ruhrgebiet (DE)	5,280,039
5	Berlin (DE)	4,980,394
6	Barcelona (ES)	4,374,747
7	Milano (IT)	4,052,933
8	Athens (GR)	4,045,748
9	Roma (IT)	3,850,688
10	Napoli (IT)	3,545,095
…	…	…
22	Amsterdam (NL)	2,381,265
…	…	…
135	Bydgoszcz (PL)	489,204
…	…	…
282	Algeciras (ES)	263,244
283	Bayreuth (DE)	259,547
…	…	…
570	Targoviste (ROM)	120,141
571	Cáceres (ES)	119,493
…	…	…
690	Acireale (IT)	54,978
691	Santa Lucía de Tirajana (ES)	53,630
692	Mollet del Vallès (ES)	51,648

3. Analysis for Europe

In Fig. 2, we depict the pan-European population distribution across all 692 functional urban areas. The data are arranged as a logarithmic rank–size plot, which is typically used to illustrate city size distributions. We observe that the relationship is straight in the intermediate range of city sizes with populations roughly between $\exp(11.7) \approx 120{,}000$ and $\exp(14.7) \approx 24{,}000{,}000$ inhabitants. In that range, which spans the ranks 22–570 in the urban hierarchy, city sizes can be approximated by a power law distribution. Outside this range, there are three notable deviations.

First, the plot turns concave for smaller cities. This, however, is a typical feature that is observed in many individual countries. See Eeckhout (2004), Rozenfeld et al. (2011), or Giesen and Suedekum (2014), who emphasize that the power law behavior pertains to the upper tail of the distribution only. Second, the plot also turns concave for large cities. On a pan-European scale, the largest cities are thus "too small" relative to a power law distribution. Within single European countries, this often tends to be the opposite. Here we observe that the largest city within a country is often "too large" for a power law; examples include Vienna in Austria or Budapest in