



biblio.ugent.be

The UGent Institutional Repository is the electronic archiving and dissemination platform for all UGent research publications. Ghent University has implemented a mandate stipulating that all academic publications of UGent researchers should be deposited and archived in this repository. Except for items where current copyright restrictions apply, these papers are available in Open Access.

This item is the archived peer-reviewed author-version of:

Measuring hierarchical differentiation: connectivity and dominance in the European urban network

Van Nuffel, N., Saey, P., Derudder, B., Devriendt, L., Witlox, F.

In: *Transportation Planning and Technology*, 33 (4), 343-366, 2010.

doi: 10.1080/03081060.2010.494028

To refer to or to cite this work, please use the citation to the published version:

Van Nuffel, N., Saey, P., Derudder, B., Devriendt, L., Witlox, F. (2010). Measuring hierarchical differentiation: connectivity and dominance in the European urban network. *Transportation Planning and Technology* 33 (4), 343-366. doi 10.1080/03081060.2010.494028.

Measuring Hierarchical Differentiation: Connectivity and Dominance in the European Urban Network

Nathalie Van Nuffel, Piet Saey, Ben Derudder, Lomme Devriendt, Frank Witlox*

Geography Department, Ghent University, Belgium

Krijgslaan 281 (S8), 9000 Ghent, Belgium

t: (+32)92644553, f: (+32)92644985

e: nathalie.vannuffel@fulladsl.be; pieter.saeyl@telenet.be;

ben.derudder@ugent.be; lomme.devriendt@ugent.be; frank.witlox@ugent.be

w: <http://www.geoweb.ugent.be/seg>

** = corresponding author. Email: frank.witlox@ugent.be*

Abstract

This paper presents an examination of the empirical merits of a set of spatial interaction indices for measuring hierarchical differentiation (i.e. dominance and connectivity) in a spatial network. To allow for the comparison of the degree of hierarchical differentiation in networks with different number of nodes/links, we propose to normalize the ratio between the real measures and the corresponding values for a rank size distribution in order to obtain readily interpretable measures of hierarchical differentiation. When applied to data on

air passenger flows within Europe, the normalized indices, interpreted together, appear to give a good idea of the tendency towards hierarchical differentiation. The potential usefulness of this analytical framework is discussed in the context of studies on (transnational) inter-city relations and empirical assessments of changes in the spatial configuration of airline networks.

Keywords: hierarchical differentiation; European urban network; airline networks; spatial interaction indices

1. Introduction

The objective of this paper is to develop an analytical framework that allows for the examination of hierarchical differentiation in spatial networks. In this respect we seek to complement the work of Patuelli *et al.* (2007, 2010) and Schintler *et al.* (2007), who propose to ‘decode’ the complexity of spatial networks by developing appropriate measures of network connectivity (see Reggiani and Nijkamp 2007). The research presented here builds on the work of Limtanakool *et al.* (2007), who introduce a number of spatial interaction indices with the aim of examining the pattern of interaction between Functional Urban Areas (FURs) in France and Germany. Based on the values for these indices, the urban network configuration in both countries was located on the continuum between the archetypal fully monocentric and fully polycentric networks. In this paper, we propose to extend their framework by calculating two additional measures and by normalizing the ratio between the different measures and their corresponding values for a rank size distribution. These extensions are deemed necessary because of possible interpretation problems with the initial framework, which primarily stem from the fact that the clear-cut interpretation of these measures seems to depend on the number of nodes/links in the network, especially when the latter becomes large and complex. The relevance of this extended analytical framework will be assessed by applying it to data on air passenger flows within Europe.

In the context of this paper, we interpret the spatial interaction indices as measures of ‘connectivity’ and ‘dominance’. Connectivity and dominance are two aspects of hierarchical differentiation. This concept, borrowed from Pumain (2006), refers to the ranking of elements from large to small (e.g. the

rank size rule for cities). It differs from hierarchical organisation, which indicates the existence of different levels, with new properties emerging at each level (e.g. a Christaller pattern of central places). Hierarchical differentiation in a spatial network has three features, i.e. (i) dominance (at the nodal level), which relates to the degree to which flows are evenly distributed across the different nodes in the network; (ii) connectivity (at the flow level), which relates to the degree to which flows are evenly distributed across the different links in the network; and (iii) symmetry (at the flow level), which relates to the degree of reflexivity of the flows. In principle, the analysis of hierarchical differentiation involves the measurement of each of these characteristics, because a network with relatively little hierarchical differentiation in terms of dominance may well exhibit extensive hierarchical differentiation in terms of connectivity and/or symmetry. However, because of data constraints which will be detailed in the discussion of our dataset, we are unable to measure symmetry, and this paper therefore exclusively deals with the measurement of connectivity and dominance.

The paper is organized as follows. The next section details the context of our research, i.e. the numerical study of human infrastructure networks and its implications for a wider set of use of related societal and economic phenomena. We then introduce our dataset, and explain the transformations that were needed to make the data suitable for the purposes of our analysis. The subsequent section presents the four spatial interaction indices based on the work of Limtanakool *et al.* (2007), after which some preliminary results are used to call for the extension of this framework. The two next sections focus on the proposed changes and the main results respectively, after which the paper is concluded with an overview of our main findings.

2. Context: the numerical study of human infrastructure networks

The numerical study of human infrastructure networks has received growing attention in recent years (Reggiani and Nijkamp 2007). An obvious reason for this mounting success is the awareness of the profound implications of the structure and functioning of these networks for a wider set of related societal and economic phenomena. Commuting networks are an obvious example here, as it is clear that the structure of these networks is closely associated with questions of residential choice and job location.

Another field of enquiry that has received increasing attention in this research domain is the study of the spatial configuration of transportation networks in the context of a wider analysis of global inter-city relations (e.g., Smith and Timberlake 2001, Matsumoto 2004, Derudder *et al.* 2007, Taylor *et al.* 2007). The research presented in this paper is part of this literature, and makes use of information on air passenger data for measuring transnational urban networks¹. Within research on contemporary globalization, studies on transnational urban networks have taken on a prominent place. Since the 1990s, for instance, it has become commonplace to emphasize the importance of border-crossing relations between cities (e.g., Castells 1996, Sassen 2001, Taylor 2004). As a consequence, most urban scholars have come to acknowledge that major cities

¹ The research presented in this paper is part of a larger research project on the quantification of the world city network. This research is funded by the Fund for Scientific Research Flanders and carried out at the Department of Geography of Ghent University in co-operation with the Globalisation and World Cities (GaWC) Study Group, a research group founded by Peter Taylor that focuses on the external relations of world cities. Its principal aim is a thorough investigation of the specification, measurement and analysis of the relation between world cities and globalization.

are increasingly produced and reproduced by what flows through them, rather than by what remains fixed within them. In this context, numerous authors have pointed to severe data deficiencies as regards these inter-city relations. Taylor (1999) has argued that there are two main reasons for this empirical deficit: “First, statistics in general have their origin in servicing the information needs of states and this has resulted in our contemporary world being measured through state-centric data. [...] Second, statistics in general have developed a critical bias towards measuring attributes at the expense of connections.” In recent years, two different approaches have been adopted in averting this empirical deficit (Derudder 2006), i.e. (i) the corporate organization solution, analysing organization networks created by firms that pursue global strategies, and (ii) the infrastructure solution, describing the telecommunication and transport networks that have enabled organisations to go global. Air passenger data have become popular for analysing urban networks centred on the world’s major cities for three reasons: (i) this kind of information is publicly available, (ii) air transport is all about connections and flows, and (iii) such data allow for a relatively straightforward analysis of the spatial patterning of transnational inter-city relations (Derudder and Witlox 2005, 2009).

Besides being popular for research on inter-city relations, airline data are also suitable for examining the effects of developments in the airline industry on the spatial organization of airline networks. In the U.S., for instance, the deregulation of the passenger aviation market in 1978 has resulted in a radical reorganization of the airline network. More specifically, agreements between airports and airlines have tended to result in hub-and-spoke configurations in which a small number of key airports (hubs) serve as transfer points where passengers change planes. From these hubs, the spoke flights then take

passengers to their final destinations (Burghouwt *et al.* 2003). The process of deregulation also took place in Europe, albeit in a more gradual way. Three packages of deregulation measures (in 1987, 1989 and 1992) have led to a shifting of power from governments towards the European airlines (Button *et al.* 1998, Hakfoort 1999). Because European carriers already showed a very high traffic concentration rate before deregulation, the deregulation process did not result in a restructuring as radical as in the US (Burghouwt *et al.* 2003). However, Burghouwt *et al.* (2003) equally notice that “at a smaller scale, radial strategies can be observed among regional carriers”. The advantages of such a radial hub-and-spoke configuration, as compared to a point-to-point configuration (Figure 1), are obvious: for the same number of destinations, there are fewer routes to serve, which in turn yields the possibility of higher flight frequencies and bigger aircrafts (Burghouwt and Hakfoort 2001). Although the gradual deregulation of the airline market may seem to be coupled with a continuous evolution towards hub-and-spoke configurations (further facilitated by mergers and alliance building), a number of important countertendencies have emerged in the last few years. A first major countertrend is induced by the mounting success of low-cost carriers, which tend to prefer a point-to-point organization to avoid costs associated with the organization of an elaborate transfer system (Alderighi *et al.* 2005)². This resurgence of point-to-point forms of spatial organization is, however, also apparent in the airline sector more generally; this can for instance be read from the different visions developed by Boeing and Airbus – the world’s leading commercial jet producers – as to the future organization of airline networks.

² However, the results of a more recent empirical analysis by Alderighi *et al.* (2007) indicate that the *spatial* network configuration of full service carriers and low cost carriers is in fact largely similar. It is their *temporal* configuration that is very different in terms of the hub-and-spoke system: by adopting a wave-system structure in the airline flight schedule, full service carriers show high temporal concentration (allowing for the transfer of passengers), while low cost carriers have almost no temporal concentration.

Both firms' latest commercial airplane, i.e. Boeing's 787 Dreamliner and Airbus's A380, are based on diametrically opposite visions of the future. The Airbus A380 represents the hub-and-spoke model in that it is built around the assumption that airlines will continue to fly smaller planes on shorter routes (spokes) into a few large hubs, then onward to the next hub on giant airplanes. It also presumes that passengers will want to put up with the hassles of changing planes. Boeing's 787, in contrast, represents an alternative in that it does not take the hub-and-spoke model as a given. The company bets on increased point-to-point connectivity, and substantiates this based on the observation that since 1990 the number of city pairs more than 3000 nautical miles apart served by the world's airlines has doubled. This trend shows no sign of abating, while the average airplane size has actually declined slightly. All this suggests that customers have come to prefer more point-to-point flights on smaller airplanes, and this may well point to a change in the fortune of the point-to-point system at large (Bowen 2002, Graham and Goetz 2008).

Figure 1 about here

Both the emergence of transnational urban networks – facilitated by air transport flows – and the continuous developments in the air transport industry lead to a number of questions regarding the spatial configuration of airline networks at large. O'Connor (2003), for instance, has suggested that the geography of airline passenger movements through the major cities of the world has changed between 1990 and 2000 at the expense of the very large global cities and major hubs in favour of a group of next largest cities. This observation is, however, only substantiated by very general examinations of the changing spatial configuration of global airline networks, such as simple

comparisons of the shares of total passenger movement through cities and the exploration of longitudinal changes in the connectivity between cities. A number of other studies have tried to measure this shifting spatial configuration in more detail, whereby the predominant method has been to compare the real network configurations with ideal hub-and-spoke and point-to-point structures (for an overview, see Alderighi *et al.* 2007).

In this paper, we propose to take another approach by situating airline-based networks between both ideal-typical extremes. More specifically, we will do so by examining the degree of hierarchical differentiation in the network. In this context, the spatial configuration of transnational urban networks can be located on the continuum between an archetypal primate system (maximal hierarchical differentiation) and a system characterized by a perfectly equal distribution (no hierarchical differentiation). In terms of airline networks *per se*, in turn, the point-to-point organization (fuelled by Boeing's vision for the future and the success of low-cost carriers) results in limited hierarchical differentiation, while hub-and-spoke networks (fuelled by Airbus's vision for the future and the continued dominance of legacy carriers) result in heightened hierarchical differentiation. A longitudinal analysis of the degree of hierarchical differentiation in the network, then, may be used to estimate the net effect of both tendencies that have opposite effects on the degree of primacy/equality in the network at large.

It can be noted that the study of transnational urban networks and air transport networks imply different (but strongly related) units of analysis, i.e. cities versus airports. However, in view of the larger research project of which our analysis is part (see footnote 1 for more details), our principal aim is not to

measure the airline network as such. Rather, we will use air passenger flows to examine hierarchical differentiation in the European urban network, which in turn serves as the yardstick for elaborating on the usefulness of our framework in general. As a consequence, air transport flows are aggregated at the city level in this paper (e.g. the flows to/from Heathrow, Stansted, Gatwick, City Airport and Luton are aggregated in a single London measure).

3. Data

3.1. Description of the AEA-database

The analysis in this paper makes use of AEA-data. AEA stands for Association of European Airlines, a non-profit-making organisation that brings together 36 major European airlines³. The AEA represents its member airlines in dialogue with all the relevant European and international organisations in the aviation value chain, thus ensuring the sustainable growth of the European airline industry in a global context (<http://www.aea.be>). Through the cooperation of an airline, we were able to obtain a AEA-database that contains, for each connection, information on the carrier⁴, origin and destination (airport, city, country and region), number of passengers (subdivided into first class, business class and economy class), freight, mail, number of flights (subdivided into passenger flights and freight flights) and distance between origin and destination. The data is summarized on a monthly basis for the period January

³ Situation as of 12 May 2010.

⁴ In case of so-called code-share flights, a single connection has two or more carrier codes. In these instances, the connection was allocated to the carrier that actually operated the flight (and which is traditionally mentioned first on the flight schedule).

2001 – December 2005, which allows a detailed analysis on recent data. Because of some difficulties with the homogeneity of the data for different years⁵, in this paper we will only discuss the results for 2005.

Obviously, the AEA-data is a very rich source of information. There are, however, two drawbacks that should be taken into account when interpreting the results. The first problem is the fact that no low-cost carriers are member of the AEA. According to the European Low Fares Airline Association (www.elfaa.com), the low cost carrier sector accounted for approximately 30% of intra-European traffic in 2006. As mentioned above, low-cost carriers tend to prefer a point-to-point organisation of air traffic. Not including them will therefore have some influence on the results. A second and more important disadvantage of the AEA-data (at least in the context of the analysis of transnational inter-city relations), and a potential source of distortions and misinterpretations, is the lack of real origin-destination data: the database records the individual legs of a trip rather than the trip as a whole. For example, a flight from Oslo to Madrid via London will be recorded as two separate flights, one from Oslo to London and the other from London to Madrid. Any possible stopovers are not registered as such, which implies that the connectivity of cities with an important hub function, like London and Paris, will be overestimated.

3.2. Transformation of the data

⁵ Between 2003 and 2004, the number of passengers shows a major increase, mainly caused by a growth in domestic passengers (passengers flying within one country), which is due to a change in the registration procedure.

The AEA-database includes flights within Europe, as well as flights between Europe and other regions. For our research purposes, we only selected those flights where both the origin city and destination city are European cities. Since our analysis is part of a research on flows between cities, we converted the airport-to-airport database into a city-to-city database by summing the number of passengers over all the airports for a given city. Finally, given that we do not know the home-based location of the travelers, we summed the passengers travelling from city A to city B with those travelling in the opposite direction, and grouped the same connections, resulting in a database of non-directional flows. After these transformations, our database contains 130,663,329 passengers (of which 90% economy class and 10% business and first class), divided over 22 carriers, 35 countries and 183 cities.

As put forward in the introduction, this lack of directional information implies that we are unable to measure the third component of hierarchical differentiation, i.e. the degree of symmetry at the flow level. An earlier study by Lo and Yeung (1996), focusing on world city-formation in Pacific Asia, reveals that airline networks can indeed be quite asymmetrical. For instance, they note that Hong Kong, Kuala Lumpur, Manila, Taipei, and Tokyo are ‘source’ nodes with a large excess of departing passengers, while Seoul, Singapore and Bangkok, Osaka and Jakarta are ‘sink’ nodes with far more arriving than departing passengers, whereas Beijing and Shanghai are almost evenly balanced. At the same time, this overview shows that it is very hard to predict how hierarchical differentiation in terms of symmetry relates to the degree of dominance and connectivity in the network.

3.3 Possible analyses

The AEA data offers the possibility for various analyses and comparisons: the evolution in time (2001-2005), the difference between business class and economy class, and separate analyses per carrier or country. With regard to the latter, we choose the three carriers with the largest number of passengers flying within Europe, namely Air France, British Airways and Lufthansa, and their home base countries France, the United Kingdom and Germany. The evolution in time will not be dealt with in this paper because of the above mentioned homogeneity problem.

To test the relevance of our analytical framework, we will apply it to examine four hypotheses as regards the spatial structure of the European urban network. First, because not all cities are business centres, we expect business class flows to be more hierarchically differentiated than economy class flows. Put differently: the inclusion of a number of major tourist destinations in the dataset implies that we expect economy class flows to be more evenly distributed (and thus less hierarchically differentiated) than business class flows. Second, we expect a difference in hierarchical differentiation between France and the United Kingdom on the one hand and Germany on the other hand because of their different urban configuration: the primacy of London and Paris implies that we expect the national networks to be more hierarchically differentiated than the German urban network, which is notorious for its more polycentric structure (see e.g. Krätke 2001, Taylor *et al.* 2006, Van Nuffel *et al.*, 2010). Third, because of the hub-and-spoke strategy adopted by most full-service carriers, it can be expected that so-called ‘legacy carriers’ (i.e. full-service carriers that emerged from erstwhile ‘national airlines’) will exhibit more hierarchical differentiation than the urban networks in their respective home

countries. In other words, we expect that the hub-and-spoke strategies adopted by Air France (Paris), British Airways (London) and Lufthansa (Frankfurt and Munich) will result in a more hierarchically differentiated pattern than those of their respective home countries. And fourth, we expect the cities to be more hierarchically differentiated than the links. The adoption of a hub-and-spoke strategy implies that one node (or in the case of Lufthansa two nodes) will be frequented by a large number of passengers en route to their final destination, and will result in a higher degree of hierarchical differentiation at the nodal level. However, when hierarchical differentiation at the link level is computed for the actually existing connections (as we have done in this paper), then the degree of hierarchical differentiation will not necessarily rise after the adoption of a hub-and-spoke strategy (cf. Figure 1). As a consequence, the very existence of hub-and-spoke networks leads to the plausible hypothesis that in the network the nodes (cities) will be more hierarchically differentiated than the links (inter-city relations). It is on the basis of these four hypotheses that we will assess the empirical merits of our analytical framework elaborated in the next sections.

4. Spatial interaction indices

In this section, we discuss the measurement of hierarchical differentiation based on an adapted and extended version of the framework presented in Limtanakool *et al.* (2007). As mentioned in the introduction, we will focus on two aspects of hierarchical differentiation, i.e. dominance (at the nodal level) and connectivity (at the link level)⁶. Because the same degree of connectivity in a network can be

⁶ Because we do not know the home-based location of the travelers, the directional flow information of our AEA-database cannot be straightforwardly interpreted in terms of actual origins and destinations. For

associated with different levels of dominance (and the other way round)⁷, we need to combine indexes for both dimensions. We use a total of four indices, two for measuring dominance and two for measuring connectivity. One of the dominance indices (the overall distribution index based on cities ODI_c) is measured at the level of the overall network; the other (the non-directional dominance index DIT_i) is measured at the level of the individual cities. Similarly, for the connectivity measures, one index is calculated at the level of the overall network (the overall distribution index based on links ODI_l); the other index is calculated at the level of the individual connections or links (the relative strength index RSI_{ij}).

The first index, the overall distribution index based on cities ODI_c , is an entropy measure that measures the extent to which the total interaction is distributed evenly across all cities in the network:

$$ODI_c = - \sum_{i=1}^I \frac{(Z_i) \ln(Z_i)}{\ln(I)} \quad (1)$$

where Z_i is the share of passengers associated with city i in the total number of passengers, and I is the number of cities in the network. A value of 1 indicates an equal distribution over the I cities, while small values point to the presence of hierarchical differentiation.

the sake of convenience, we therefore assume that the general level of dominance also applies to each pair of cities/links.

⁷ For instance, a city boasting a large number of passengers in a network has a high degree of dominance. It can send/receive these passengers to/from a lot of other cities, but also to/from only a few other cities. In the first situation the network would have a low degree of hierarchical differentiation at the link level, in the second case a high degree.

The second index is the non-directional dominance index DIT_i , calculated as the ratio between the sum of the interactions associated with city i and the average size of the interactions associated with the other cities in the network:

$$DIT_i = \frac{T_i}{\sum_{j=1}^J \frac{T_j}{J}} \quad (2)$$

where T_i is the total number of passengers associated with city i and $i \neq j$. Cities with a DIT_i value above 1 are considered dominant cities because they are more important than the average of the other cities in the network. ‘Large’ differences between DIT_i values for different cities indicate a high degree of hierarchical differentiation.

The third index, the overall distribution index based on links ODI_l , is again an entropy index, measuring the extent to which the total interaction is distributed evenly across all links (city-pairs) in the network:

$$ODI_l = - \sum_{l=1}^{L_p} \frac{(Z_l) \ln(Z_l)}{\ln(L_p)} \quad (3)$$

where Z_l is the share of passengers travelling on link l in the total number of passengers, and L_p is the potential number of links in the network. The maximum ODI_l value of 1 indicates a fully connected structure. Small values point to the presence of hierarchical differentiation.

Finally, the fourth index is the relative strength index RSI_{ij} , which is simply the proportion of interaction on a single link between two cities relative to the total interaction in the network:

$$RSI_{ij} = \frac{T_{ij}}{\sum_{i=1}^I \sum_{j=1}^J T_{ij}} \quad (4)$$

where T_{ij} is the total number of passengers travelling between city i and city j , and $i \neq j$. The RSI_{ij} values for all links in the network sum to unity, while individual values range from 0 to 1. Similar to the DIT_i measure, ‘large’ differences between RSI_{ij} values point to the presence of hierarchical differentiation.

5. Extension of the analytical framework

Before applying the analytical framework adapted from Limtanakool *et al.* (2007) to airline data, we modified it in two ways. The first modification stems from the fact that intuitively clear notions such as ‘small differences’ or ‘large differences’ between the different DIT_i and RSI_{ij} values cannot be interpreted straightforwardly. Such interpretation poses little or no problems when only a small number of nodes is analysed, as is the case in the paper of Limtanakool *et al.* (8 FURs in Germany, and 6 in France). However, when the number of nodes is large – as is the case in our research –, then conclusions about the degree of hierarchical differentiation in terms of the differences between the individual values are not always straightforward to make. We therefore propose to

calculate the standard deviations of the values of both indices as a second overall measure of hierarchical differentiation that may be helpful in interpretation of the differences between the values of DIT_i and RSI_{ij} . High standard deviations reflect large differences in the values of the indices and thus point to more dominance and less connectivity. In other words: the higher the standard deviations, the less equally divided passengers are between cities and links.

The second modification stems from the fact that the measures are sensitive to the number of cities and links. This can be shown by calculating them for rank size distributions. A rank size distribution can be defined as:

$$\sum_{i=1}^I T_i = \sum_{r=1}^I \frac{1}{r} T'_1 \quad (\text{for cities}) \quad \text{and} \quad \sum_{l=1}^{L_p} T_l = \sum_{r=1}^{L_p} \frac{1}{r} T'_1 \quad (\text{for links}) \quad (5)$$

where T_i (resp. T_l) is the number of passengers associated with city i (resp. link l), I (resp. L_p) is the total number of cities (resp. the total number of potential links), T'_1 is the virtual number of passengers of the largest city/link, and r is the rank of the city/link. If the spatial interaction indexes were independent from the number of nodes/links, then one would expect the values of the entropy indices and the standard deviations to be the same for each rank size distribution, irrespective of the number of cities or links. The values of the indices and standard deviations for rank size distributions with the same number of cities/links as the real distributions of each of our empirical foci are given in Table I. If the number of cities/links would not matter at all, then all values in each column would be the same, which is clearly not the case. This implies that comparisons between countries, carriers or service classes cannot

be made without somehow normalizing the results first. Once again, this poses little or no problems when two networks with a largely similar number of nodes are compared as in Limtanakool *et al.* However, when comparing networks characterized by large differences in terms of the number of nodes/links, this may lead to interpretation problems, and our second proposed modification to the Limtanakool *et al.* (2007) framework therefore consists in the introduction of reference points that allow for a readily interpretable judgment of the measures of hierarchical differentiation, and this irrespective of the number of nodes/links in the network.

Table I about here

Indeed, the problem that leads to the necessity of a second modification stems from the sensitivity of the indices to the number of cities/links. This implies that statements about individual values cannot be substantiated⁸, because there is no reference point to judge them by (with the exception of the extreme values of 0 and 1). However, it also implies that the results of analyses that are based on different numbers of cities/links cannot be compared. Whereas for the judgment of individual values a simple comparison to the rank size value is sufficient⁹, matters are slightly more complicated when comparing results of different analyses. Imagine two countries X and Y, with x and y cities respectively, where $x > y$. If ODI_c for country X is smaller than ODI_c for country Y, this does not necessarily imply that country X is more hierarchically

⁸ E.g. the statement of Limtanakool *et al.* (2007) that an entropy value of 0.81 points to a fairly even distribution of interaction.

⁹ In the case of the entropy values, a value smaller than the value for the rank size distribution with the same number of cities/links is interpreted as a tendency towards hierarchical differentiation, a value larger than the corresponding rank size value as a tendency towards an equal distribution of the number of passengers. For the standard deviations, the interpretation is the other way round.

differentiated than country Y. We can however obtain comparable results if we use the rank size distribution (as defined above) as a reference point, a choice that is primarily guided by the fact that it provides us with a conceptually and mathematically transparent yardstick. However, simply dividing the real values of the entropy indices and standard deviations by their corresponding rank size distribution values does not solve the incomparability problem. What we need is a normalization that, indifferent of the number of cities/links in the network analysed, results in a value between 0 and 1, where 0 indicates a completely even distribution, 1 points to maximum hierarchical differentiation and 0.5 reflects the rank size distribution¹⁰. For the entropy values (ODI_c and ODI_l), this can be achieved by applying the following formula's (the subscript N refers to the normalized value of the index or standard deviation, while the subscript RS indicates the value of the index or standard deviation for a rank size distribution):

$$ODI_N = \frac{1 - ODI}{1 - ODI_{RS}} \quad \text{when } ODI \geq ODI_{RS} \quad (6)$$

$$ODI_N = \frac{2 - \frac{ODI}{ODI_{RS}}}{2} \quad \text{when } ODI \leq ODI_{RS} \quad (7)$$

The normalized value of the standard deviation (SD) of RSI_{ij} is computed as follows:

¹⁰ To allow for a straightforward understanding of the formulas, we interpret maximal hierarchical differentiation as 'all passengers concentrated in one city/link', even though this maximum is an asymptotical case that cannot be reached in reality.

$$SD_N(RSI_{ij}) = \frac{\frac{SD(RSI_{ij})}{SD_{RS}(RSI_{ij})}}{2} \quad \text{when } SD \leq SD_{RS} \quad (8)$$

$$SD_N(RSI_{ij}) = \frac{2 - \frac{\frac{1000}{L_p} \sqrt{L_p - 1} - SD(RSI_{ij})}{\frac{1000}{L_p} \sqrt{L_p - 1} - SD_{RS}(RSI_{ij})}}{2} \quad \text{when } SD \geq SD_{RS} \quad (9)$$

In a situation of complete absence of hierarchical differentiation, all links have the same RSI_{ij} value (i.e., $1/L_p$) and the standard deviation is 0, resulting in a normalized standard deviation of 0 (formula (8)). When there is maximal hierarchical differentiation (all passengers concentrated on one link), the standard deviation reaches its maximum value of $\frac{1000}{L_p} \sqrt{L_p - 1}$ (proof in Appendix), resulting in a normalized standard deviation of 1 (formula (9)).

For the standard deviation of DIT_i , the normalization formula's are:

$$SD_N(DIT_i) = \frac{\frac{SD(DIT_i)}{SD_{RS}(DIT_i)}}{2} \quad \text{when } SD \leq SD_{RS} \quad (10)$$

$$SD_N(DIT_i) = \frac{2 - \frac{SD_{RS}(DIT_i)}{SD(DIT_i)}}{2} \quad \text{when } SD \geq SD_{RS}$$

(11)

When there is no hierarchical differentiation at all, all cities have a DIT_i value of 1 and the standard deviation is 0, resulting in a normalized standard deviation of 0 (formula (10)). When all passengers are concentrated in one city (maximal hierarchical differentiation), the standard deviation becomes infinite, resulting in a normalized standard deviation of 1 (formula (11)).

Normalizing the indices by comparing them to the corresponding indices for a rank size distribution with the same number of cities/links allows for comparisons between the various analyses. In the example given above, if ODI_c for country X is larger than ODI_c for a rank size distribution with x cities, while ODI_c for country Y is smaller than ODI_c for a rank size distribution with y cities¹¹, then the conclusion (after applying the above formula's) would be that country Y is more hierarchically differentiated than country X, which will be reflected in a larger normalized value for country Y.

Before turning to the discussion of the results, two final comments should be made. The first comment relates to the way in which the different indices deal with differences from the mean. The entropy measures ODI_c and ODI_l are not very sensitive to changes in the values of the largest cities/links, because the proportions of passengers are multiplied by their logarithm. On the other hand, although the standard deviations treat positive and negative deviations from the mean in the same way, they are more sensitive to higher deviations because of

¹¹ Cf. Table I. Notice that ODI_c for a rank size distribution with x cities will be smaller than ODI_c for a rank size distribution with y cities, since $x > y$ (in the hypothetical case that the number of cities in country Y is larger than the number of cities in country X, the example given in the text would be mathematically impossible). The same holds for ODI_l and for RSI_{ij} : the fewer the number of links, the larger the index value for the rank size distribution. For DIT_i on the other hand, fewer cities will result in a smaller standard deviation when a rank size distribution is applied.

the squaring. Therefore, in interpretations, it is best to combine entropy values and standard deviations.

The second comment relates to the use of the potential links L_p in our actual analysis. In practice, a lot of links feature no passengers at all (e.g., there are at present no direct flights between Brussels and Glasgow), so that Z_i equals 0 for quite a lot of connections. However, because 0 does not have a logarithm, these values cannot be used in the numerator of ODI_i , and in our calculations we have therefore replaced L_p by the total number of ‘real links’ L_r . As a consequence, in (3) we assume that all potential links are actually existing links, while in (1) there are no cities where Z_i equals 0. In other words: in our calculations, we only employed those links that actually feature passengers, and accordingly make use of a rank size distribution that starts from the number of ‘real links’.

6. Results

6.1. City and link level

Table II lists the top-5 of the most important cities and links in 2005, subdivided into business class and economy class. From the table, it can be read that while Frankfurt holds the fourth place for economy class flows, it comes second for business class flows. This is no surprise, given the importance of this city for (international) business. However, at the same time, this conclusion regarding the dominance of cities such as Frankfurt is of course also somewhat data-driven in that a city such as Berlin – acting as the major base of low-cost carrier Air Berlin – is underestimated because of its relatively higher proportion

of low-cost passengers, which are not included in our AEA-database. Similar observations can be made with respect to cities such as Dublin (Ryanair and AerLingus, which recently left the AEA after its transformation to a low-cost carrier) and Barcelona (Vueling).

The analyses with regard to the links show that the fourth and fifth most important links for economy class are the connections of Madrid to Paris and to London. For business class flows on the other hand, places 4 and 5 are taken by Scandinavian capital cities: Oslo-Stockholm and Copenhagen-Stockholm. The latter is, however, not a reflection of the economic importance of these Scandinavian cities. Business class was first introduced in Europe by Scandinavian Airlines (SAS) and has since then been very popular in the Scandinavian countries. The cumulated RSI_{ij} values (multiplied by 1000) of the five most important connections amount to 61.01 for economy class and to 110.39 for business class, reflecting the more hierarchically differentiated nature of the business class network.

Table II about here

Table III shows the DIT_i values of all dominant cities in 2005. In other words, the DIT_i values larger than 1. Dominant cities within the same country are listed together. As can be seen from the table, only seven countries have more than one dominant city: the UK, France, Germany, Italy, Spain, Sweden and Switzerland. The three countries for which we did a separate intra-country analysis are marked in bold. Germany has six dominant cities, while the UK and France only have three. Moreover, the range of DIT_i values of the six German cities is smaller than that of the British and French cities. As mentioned above, we therefore expect domestic flights in Germany to show

more connectivity, or less hierarchical differentiation, than domestic flights in France and the United Kingdom. Of the other two countries, the UK and France, the former will probably be more hierarchically differentiated, given the large DIT_1 value for London.

Table III about here

6.2. Network level

Table IV shows the normalized values of the entropy values and the standard deviations, calculated by formula's (6)-(11). Remember that a figure larger than 0.5 indicates a distribution that is more hierarchically differentiated than the rank size distribution, while a value smaller than 0.5 indicates a less hierarchically differentiated distribution.

Table IV about here

To visualise the results, some graphs were made of the distributions of the number of passengers for the different analyses (Figures 2 to 9). The bold line indicates the real values, the dashed thin line the values for the corresponding rank size distribution. The x-axis is made logarithmic to ease interpretation. The corresponding normalized index values are indicated beside the graphs.

Figures 2 and 3 show the distributions for the three carriers (Air France, British Airways and Lufthansa). At the city level (Figure 2), all three have a more hierarchically differentiated distribution of passengers than the rank size distribution, which is reflected in values for the indices above 0.50. Of the three

carriers, Lufthansa has the least hierarchically differentiated network. At the link level however, the bold lines are less steep than the dashed lines, which means that the passengers are more evenly distributed than would be the case for a rank size distribution. Here again, Lufthansa's network is the least hierarchically distributed, which can be related to its gradual adoption of a 'double hub'-system centred on Frankfurt and Munich.

When we look at the network in the three selected countries as a whole (France, United Kingdom and Germany) rather than those of their legacy carriers, the distribution of the passengers at the link level (Figure 5) is again less hierarchically differentiated than the distribution at the city level (Figure 4), albeit that the difference in hierarchical differentiation between cities and links is less pronounced than was the case for the carriers. At the link level, France and the United Kingdom approach the rank size distribution, Germany is clearly less hierarchically differentiated. At the same time, however, it needs to be stressed that it remains unclear how the absence of low-cost carrier flows may impact these results.

At the city level, a clear distinction can be made between the graphs for France and the United Kingdom on the one hand, and the graph for Germany on the other. The first two are steeper than the rank size distribution, while the graph for Germany approaches the rank size distribution, although it seems somewhat less steep. The index values for France and the United Kingdom are above 0.50, while the values of the indices for Germany seem to contradict each other: the entropy value is above 0.50, the standard deviation is below 0.50. In other words: the entropy value points to more hierarchical differentiation than the rank size distribution, the standard deviation to less hierarchical differentiation. The same is true for both economy class and business class passengers at the city level (Figure 6), and the explanation for this seeming contradiction is very

simple. When the y-axis is also made logarithmic (Figure 7), the graph clearly shows how the real distribution turns downwards at a relatively early stage. This means that there are a lot of small values and these small values are overrepresented in the entropy value (because of the multiplication by the logarithm). At the link level, the real distribution also goes downwards (Figure 8), but at a much later stage, resulting in an entropy value below 0.50. The indices for the links (Figure 9) are substantially smaller than those for the cities, which indicates that the links are indeed much less hierarchically differentiated than the cities.

Figures 2, 3, 4, 5, 6, 7, 8 and 9 about here

7. Conclusions

The formative aim of this paper was to examine the empirical merits of an analytical framework that allows measuring the degree of hierarchical differentiation in a spatial network. Our analysis of hierarchical differentiation in the European urban network as exemplified through airline flows was used as a heuristic device to assess the merits of this empirical framework. We have argued that the basic set of four spatial interaction indices suggested by Limtanakool *et al.* (2007) needs to be extended in two ways. First, the standard deviations of two of the indices (DIT_i and RSI_{ij}) may be calculated as a second overall measure of hierarchical differentiation. Second, because of their sensitivity to the number of cities or links, these standard deviations and the entropy indices ODI_c and ODI_l should be normalized, especially when dealing with large and/or complex networks. This can be done by comparing them to

their corresponding values for a rank size distribution with the same number of cities or links. The four normalized measures, when interpreted together, appear to give a good idea of the tendency to hierarchical differentiation. This appreciation is based on the fact that our four hypotheses are confirmed by our results: business class is more hierarchically differentiated than economy class; Germany is less hierarchically differentiated than France and the United Kingdom; the carriers are more hierarchically differentiated than their home countries; and the links are less hierarchically differentiated than the cities.

The analytical framework elaborated in this paper can be used as a new tool in the scientific literature seeking to develop adequate measurement tools to analyse spatial networks. The normalization of the indices allows for the comparison of the organizational structure of networks irrespective of the number of nodes/links, which may for instance be useful in longitudinal analyses that involve new and/or vanishing nodes and links in the development of the network. Although the potential usefulness of this framework has been discussed in the context of studies on (transnational) inter-city relations and empirical assessments of changes in the spatial configuration of airline networks, it is important to stress that it can be extended to other kinds of flow data and/or networks at other scales.

References

- Alderighi, M., Cento, A., Nijkamp, P. and Rietveld, P., 2005. Network competition – the coexistence of hub-and-spoke and point-to-point systems. *Journal of Air Transport Management*, 11(5), 328-334.
- Alderighi, M., Cento, A., Nijkamp, P. and Rietveld, P., 2007. Assessment of new hub-and-spoke and point-to-point airline network configurations. *Transport Reviews*, 275, 529-549.
- Bowen, J., 2002. Network change, deregulation, and access in the global airline industry. *Economic Geography*, 78, 425–439.
- Burghouwt, G. and Hakfoort, J., 2001. The evolution of the European aviation network, 1990-1998. *Journal of Air Transport Management*, 7(5), 311-318.
- Burghouwt, G., Hakfoort, J. and Ritsema van Eck, J., 2003. The spatial configuration of airline networks in Europe. *Journal of Air Transport Management*, 9(5), 309-323.
- Button, K.J., Haynes, K. and Stough, R., 1998. *Flying into the Future. Air Transport Policy in the European Union* (Cheltenham: Edward Elgar).
- Castells, M., 1996. *The Rise of the Network Society* (Oxford: Blackwell).
- Derudder, B., 2006. On conceptual confusion in empirical analyses of a transnational urban network. *Urban Studies*, 43(11), 2027-2046.

- Derudder, B. and Witlox, F., 2005. An appraisal of the use of airline data in assessing the world city network: a research note on data. *Urban Studies*, 42(13), 2371-2388.
- Graham, B. and Goetz, A.R., 2008. Global air transport. In: R. Knowles, J. Shaw and J. Docherty, eds. *Transport Geographies - Mobilities, Flows and Spaces*. Oxford: Blackwell, 137-155.
- Hakfoort, J.R., 1999. The deregulation of European air transport: a dream come true? *Tijdschrift voor Economische en Sociale Geografie*, 90(2), 226-233.
- Krätke, S., 2001. Strengthening the polycentric urban system in Europe: conclusions from the ESDP. *European Planning Studies*, 9(1), 105-116.
- Limtanakool, N., Dijst, M. and Schwanen, T., 2007. A theoretical framework and methodology for characterising national urban systems on the basis of flows of people: empirical evidence for France and Germany. *Urban Studies*, 44(11), 2123-2145.
- Lo, F.-C. and Yeung, Y.-M., eds, 1998. *Globalization and the World of Large Cities*. Tokyo: United Nations University Press.
- Matsumoto, H., 2004. International urban systems and air passenger and cargo flows: some calculations. *Journal of Air Transport Management*, 10, 241-249.

O'Connor, K., 2003. Global air travel: toward concentration or dispersal? *Journal of Transport Geography*. 11, 83–92.

Patuelli, R., Reggiani, A., Gorman, S.P., Nijkamp, P. and Bade, F.J., 2007. Network analysis of commuting flows: a comparative static approach to German data. *Networks and Spatial Economics*, 7, 315–331.

Patuelli, R., Reggiani, A., Nijkamp, P. and Bade, F.-J., 2010. The evolution of the commuting network in Germany. Spatial and connectivity patterns. *Journal of Transport and Land Use*, 2 (3/4), 5-37.

Pumain, D., 2006. *Hierarchy in Natural and Social Sciences*. Dordrecht: Springer.

Reggiani, A. and Nijkamp, P., 2007. Transport networks and metropolitan development: new analytical departures. *Networks and Spatial Economics*, 7, 297–300.

Sassen, S., 2001. *The global city: New York, London, Tokyo*. Princeton: Princeton University Press.

Schintler, L.A., Kulkarni, R., Gorman, S. and Stough, R., 2007. Using raster-based GIS and graph theory to analyze complex networks. *Networks and Spatial Economics*, 7, 301–313.

Smith, D.A. and Timberlake, M., 2001. World city networks and hierarchies, 1977-1997: an empirical analysis of global air travel links. *American Behavioral Scientist*, 44, 1656-1678.

Taylor, P.J., 1999. "So-called world cities": the evidential structure within a literature. *Environment and Planning A*, 31(11), 1901-1904.

Taylor, P.J., 2004. *World city network: a global urban analysis*. London: Routledge.

Taylor, P.J., Derudder, B. and Witlox, F., 2007. Comparing airline passenger destinations with global service connectivities: a worldwide empirical study of 214 cities. *Urban Geography*, 28(3), 232-248.

Taylor, P.J., Evans, D. and Pain, K., 2006. Organization of the polycentric metropolis: corporate structures and networks. In: P. Hall and K. Pain, eds. *The Polycentric Metropolis. Learning from mega-city regions in Europe*. London: Earthscan, 53-69.

Van Nuffel, N., Derudder, B. and Witlox, F., 2010. Even important connections are not always meaningful: on the use of a polarization measure in a typology of European cities in air transport networks. *Tijdschrift voor Economische en Sociale Geografie*, 101(3), 333-348.

Appendix

The standard deviation (SD) for a population is given by:

$$SD = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}}$$

with x_i the value for element i , \bar{x} the average value and N the number of elements.

The maximum SD for RSI_{ij} is reached when all passengers are concentrated on one link. The RSI_{ij} value for this link is 1, while the RSI_{ij} values for all other links are 0. If L_p is the total number of potential links in the network, then the maximum SD for RSI_{ij} is given by:

$$SD_{\max} = \sqrt{\frac{(1000 - \frac{1000}{L_p})^2 + (L_p - 1) \left(\frac{1000}{L_p} \right)^2}{L_p}}$$

The use of 1000 in the formula reflects the fact that we multiplied the RSI_{ij} values by 1000 to ease their interpretation. If the original RSI_{ij} values are used, the 1000 should self-evidently be changed to 1.

The formula can be further worked out as follows:

$$SD_{\max} = \sqrt{\frac{(1000)^2 + (\frac{1000}{L_p})^2 - (2)(1000)\frac{1000}{L_p} + L_p(\frac{1000}{L_p})^2 - (\frac{1000}{L_p})^2}{L_p}}$$

$$SD_{\max} = \sqrt{\frac{(1000)^2 - (2)\frac{(1000)^2}{L_p} + \frac{L_p(1000)^2}{(L_p)^2}}{L_p}}$$

$$SD_{\max} = \sqrt{\frac{(1000)^2(1 - \frac{2}{L_p} + \frac{1}{L_p})}{L_p}}$$

$$SD_{\max} = \sqrt{\frac{(1000)^2(1 - \frac{1}{L_p})}{L_p}}$$

$$SD_{\max} = \frac{1000}{L_p} \sqrt{(1 - \frac{1}{L_p})L_p}$$

$$SD_{\max} = \frac{1000}{L_p} \sqrt{L_p - 1}$$

Table I: Results for rank size distributions

	Cities			Links		
	SD DIT _i	ODI _c	# cities	SD RSI _{ij}	ODI _l	# potential links
Total	3.27	0,79	183	5.05	0.70	33306
Economy	3.27	0,79	183	5.05	0.70	33306
Business	3.12	0,79	160	5.57	0.71	25440
France	1.85	0,83	33	22.82	0.75	1056
Germany	1.52	0,85	18	36.08	0.77	306
United Kingdom	1.52	0,85	18	42.51	0.77	306
Air France	2.42	0,81	75	20.31	0.73	5550
British Airways	2.33	0,81	67	21.34	0.73	4422
Lufthansa	2.57	0,80	87	15.10	0.72	7482

Table II: Top-5 DIT_i and RSI_{ij} values in 2005

	DIT_i (DIT_i value between brackets)		RSI_{ij}	
	Economy class	Business class	Economy class	Business class
Rank	City	City	Link	Link
1	London (21.36)	London (25.26)	Amsterdam-London	Amsterdam-London
2	Paris (17.12)	Frankfurt (15.27)	London-Paris	London-Paris
3	Amsterdam (13.74)	Paris (14.11)	Frankfurt-London	Frankfurt-London
4	Frankfurt (12.06)	Amsterdam (12.54)	Madrid-Paris	Oslo-Stockholm
5	Munich (7.34)	Munich (7.82)	London-Madrid	Copenhagen-Stockholm
Cumulated RSI_{ij} value (x 1000)			61.01	110.39

Table III: Dominant cities in 2005

Country	City	DIT _i value	Country	City	DIT _i value
United Kingdom	London	22.08	Austria	Vienna	5.57
	Manchester	2.16	Belgium	Brussels	5.32
	Birmingham	1.37	Sweden	Stockholm	4.20
France	Paris	17.03		Göteborg	1.05
	Lyons	1.25	Turkey	Istanbul	4.11
	Nice	1.18	Czech Republic	Prague	3.92
The Netherlands	Amsterdam	13.80	Greece	Athens	3.75
Germany	Frankfurt	12.58	Portugal	Lisbon	3.35
	Munich	7.50	Norway	Oslo	3.34
	Düsseldorf	2.34	Finland	Helsinki	3.09
	Berlin	1.61	Hungary	Budapest	2.77
	Hamburg	1.42	Switzerland	Geneva	2.56
	Stuttgart	1.12		Zurich	2.35
Denmark	Copenhagen	7.23	Romania	Bucharest	1.55
Italy	Milan	6.99	Ireland	Dublin	1.55
	Rome	5.78	Cyprus	Larnaca	1.32
	Venice	1.48	Malta	Gudja/Luqa	1.05
Spain	Madrid	6.89	Poland	Warsaw	1.03
	Barcelona	4.63			

Table IV: Normalized results for 2005

	# pass.	Cities			Links		
		SD DIT _i	ODI _c	# cities	SD RSI _{ij}	ODI _l	# links
Total	130,663,329	0.42	0.55	183	0.15	0.40	1088
Economy	117,853,550	0.42	0.55	183	0.15	0.40	1088
Business	12,809,779	0.48	0.56	160	0.21	0.43	929
France	17,929,179	0.75	0.60	33	0.50	0.53	98
Germany	13,206,507	0.48	0.54	18	0.32	0.44	47
United Kingdom	9,923,747	0.73	0.60	18	0.53	0.55	36
Air France	14,999,274	0.83	0.60	75	0.28	0.47	117
British Airways	15,910,731	0.82	0.59	67	0.27	0.46	108
Lufthansa	22,748,302	0.67	0.56	87	0.19	0.42	188

Figure 1: Point-to-point network and hub-and-spoke network

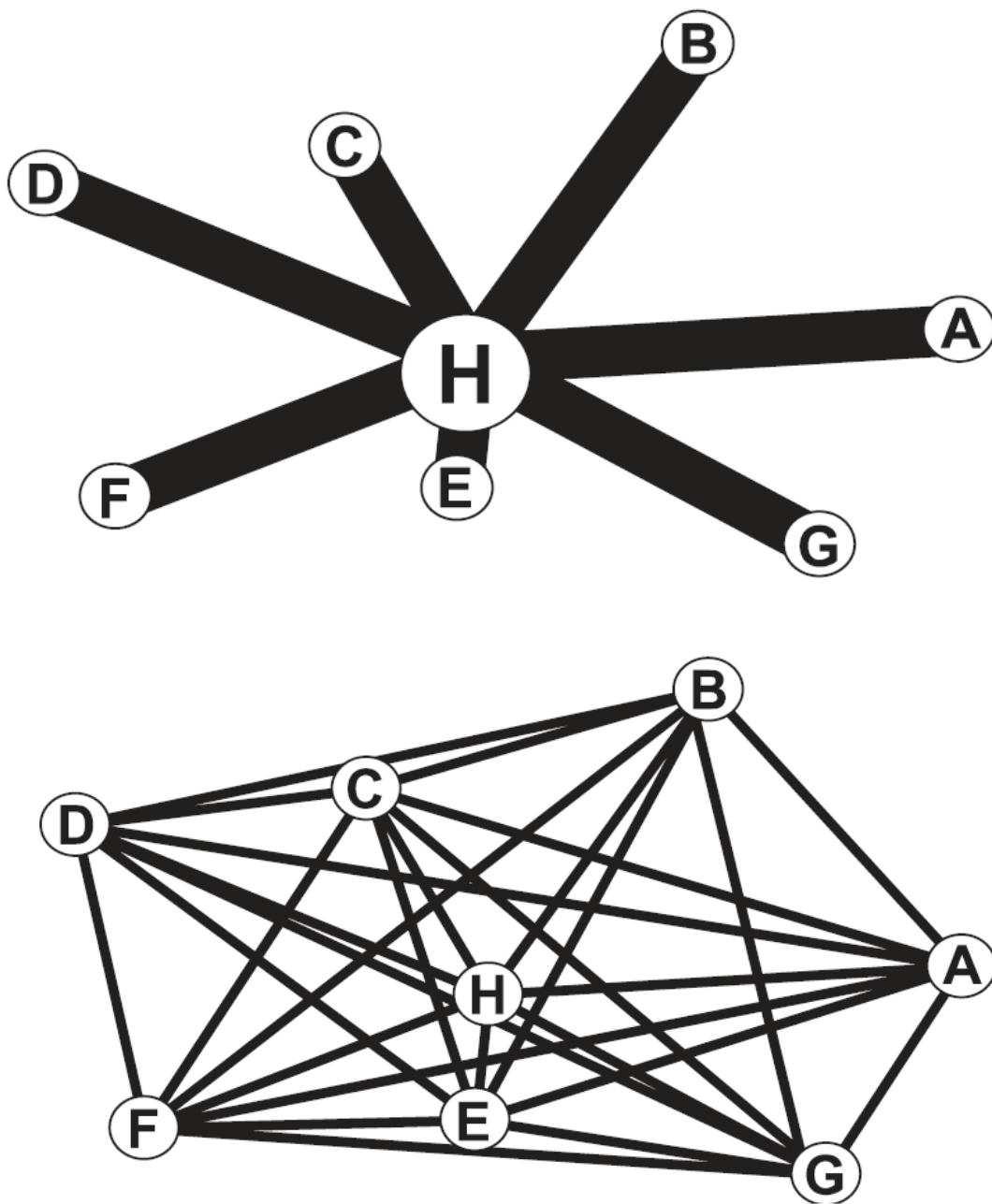
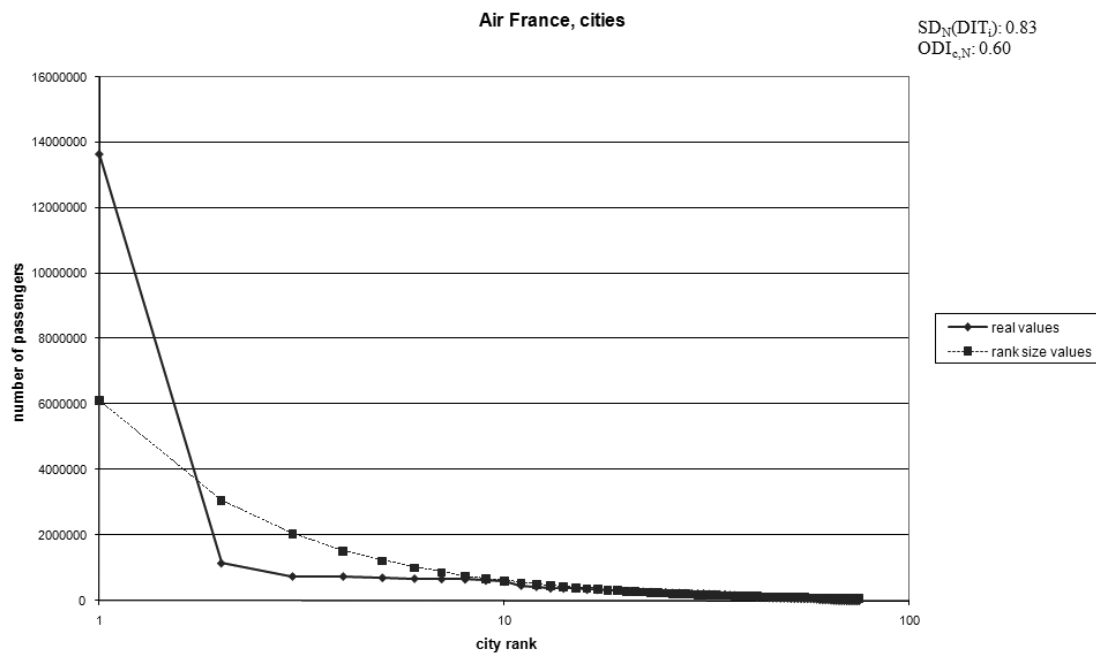
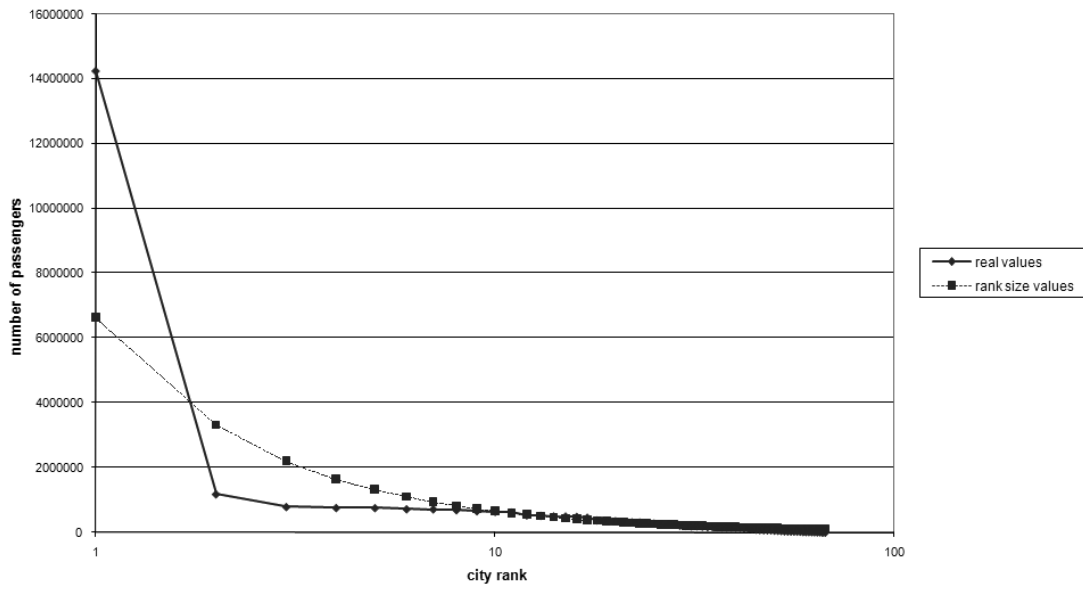


Figure 2: Carriers, city level



British Airways, cities

$SD_N(DIT_i): 0.82$
 $ODI_{c,N}: 0.59$



Lufthansa, cities

$SD_N(DIT_i): 0.67$
 $ODI_{c,N}: 0.56$

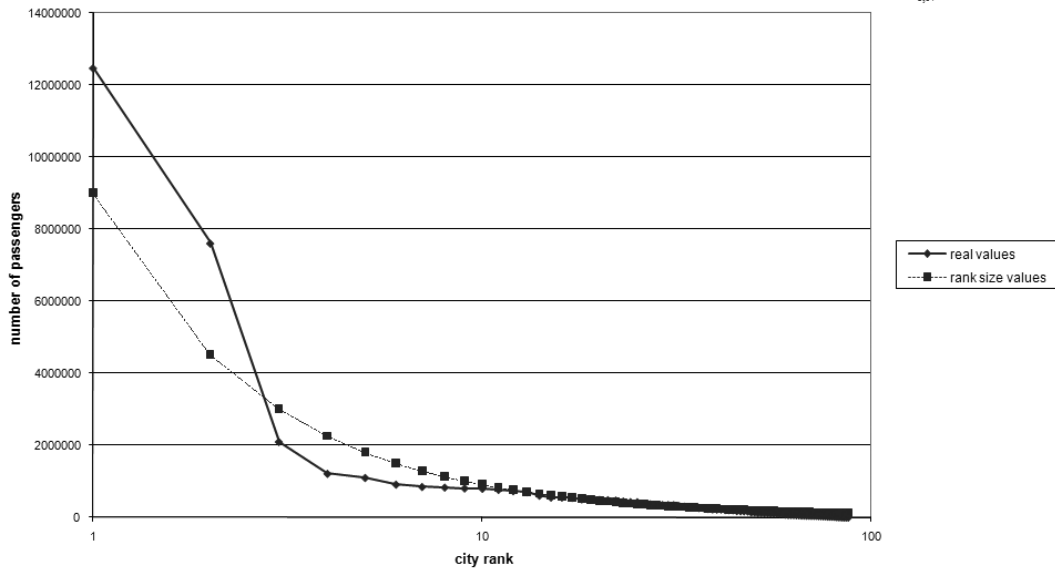
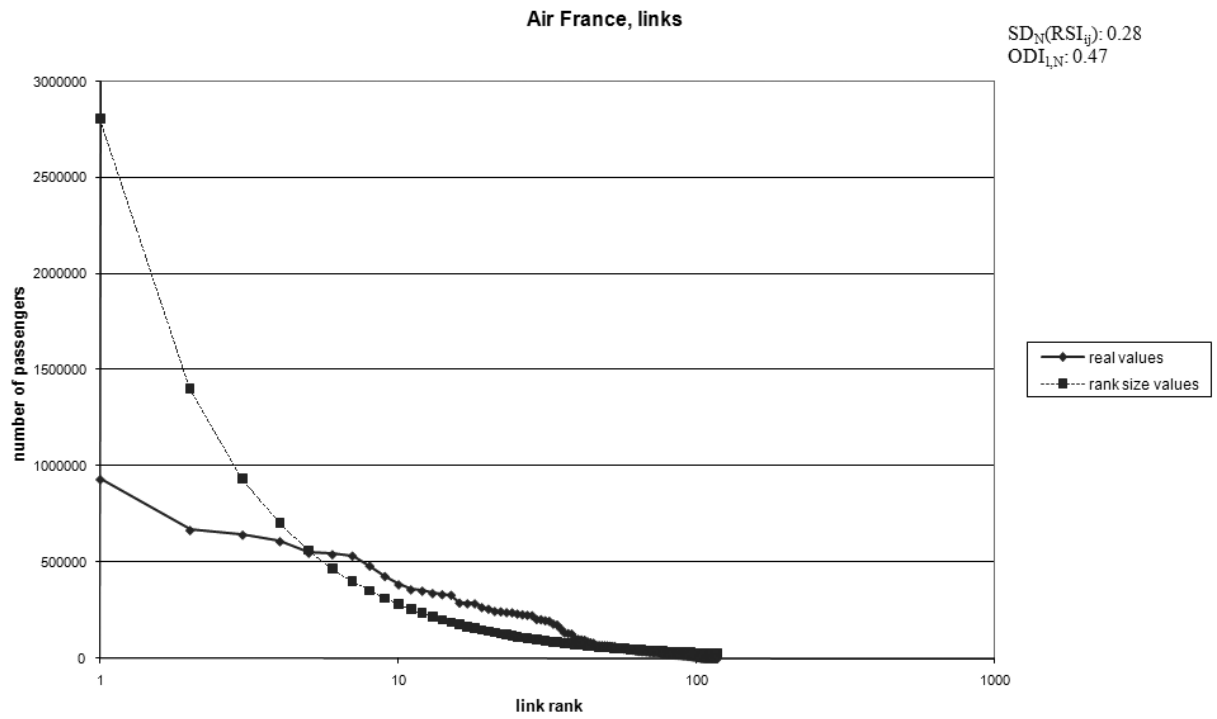
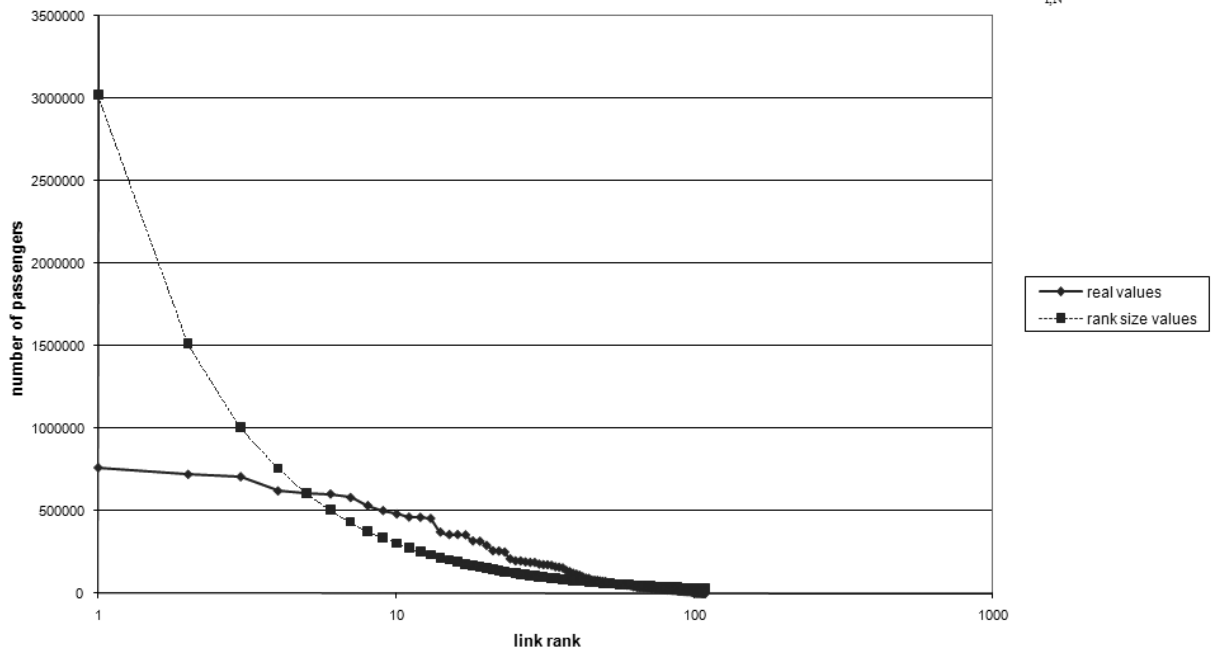


Figure 3: Carriers, link level



British Airways, links

$SD_N(RSI_{ij}): 0.19$
 $ODI_{i,N}: 0.42$



Lufthansa, links

$SD_N(RSI_{ij}): 0.27$
 $ODI_{i,N}: 0.46$

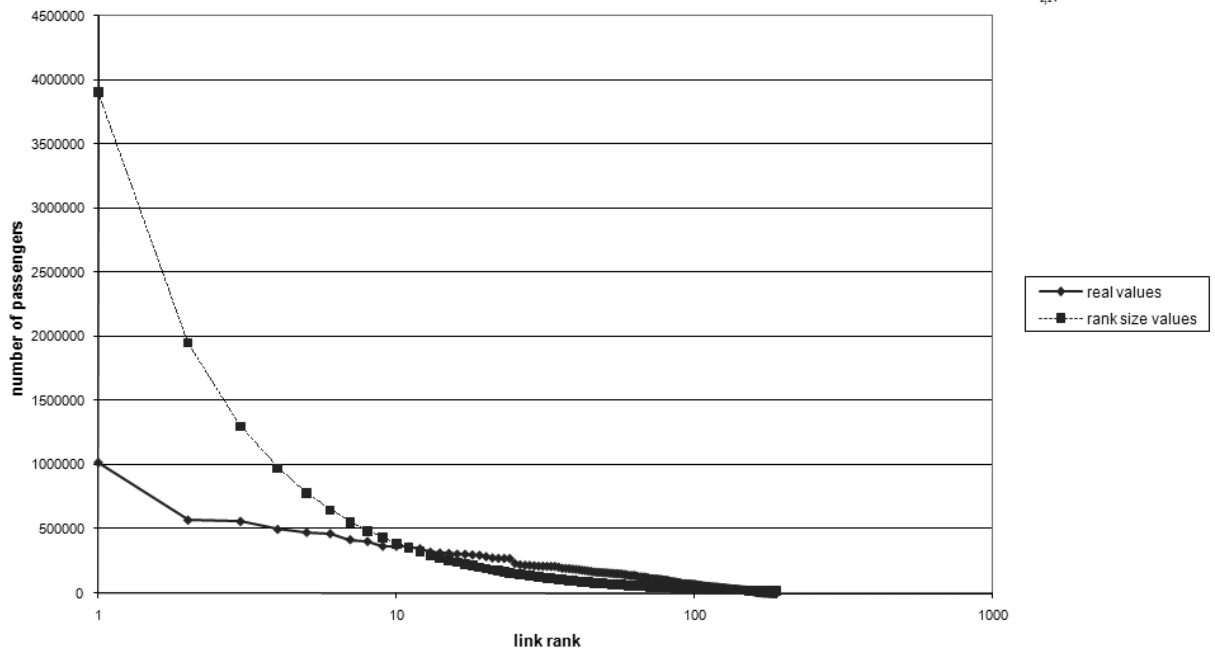
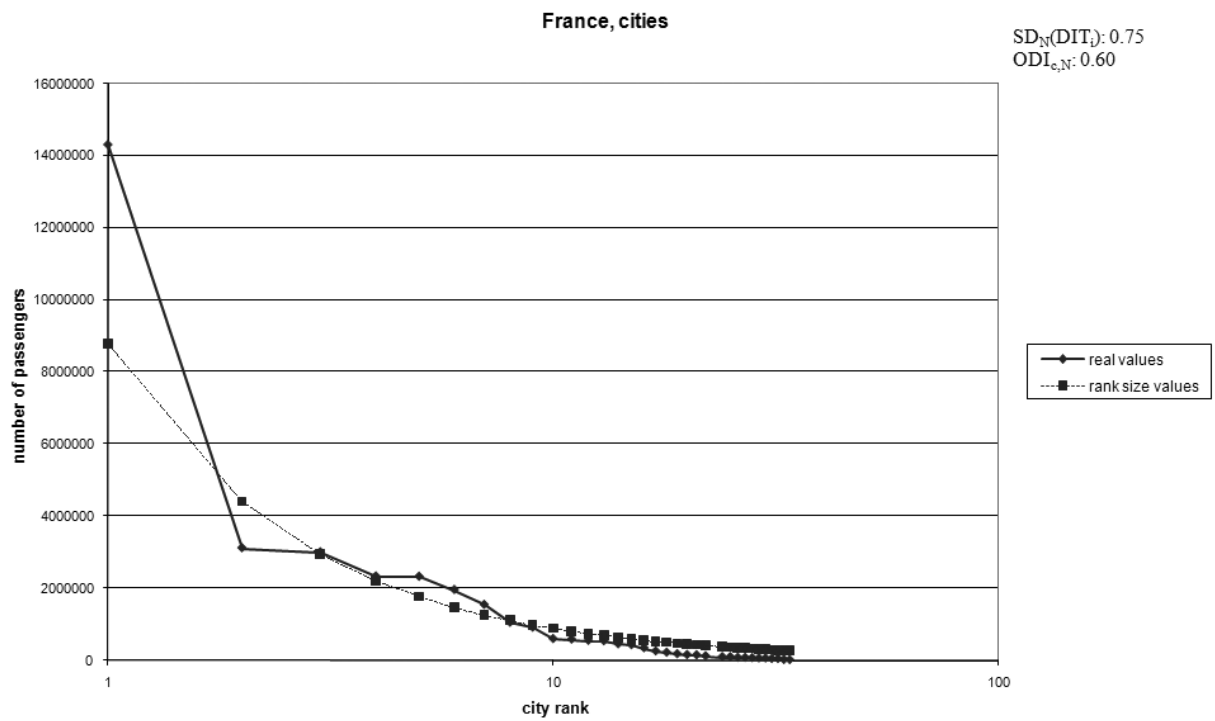
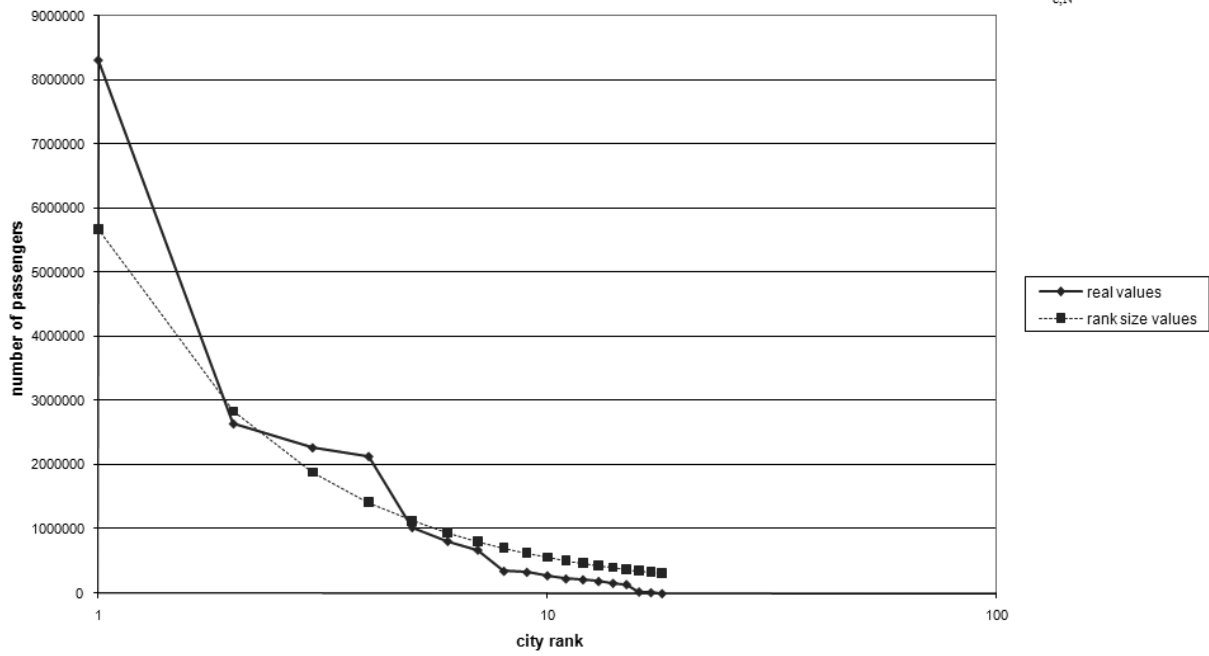


Figure 4: Countries, city level



United Kingdom, cities

$SD_N(DIT_i)$: 0.73
 $ODI_{c,N}$: 0.60



Germany, cities

$SD_N(DIT_i)$: 0.48
 $ODI_{c,N}$: 0.54

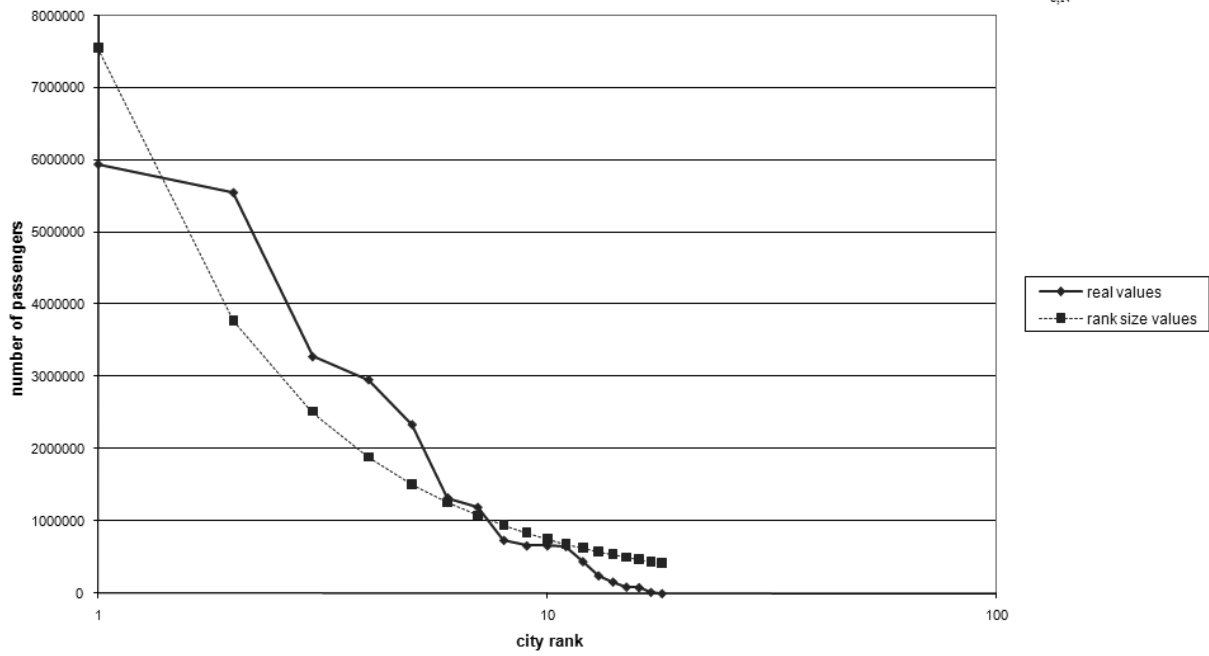
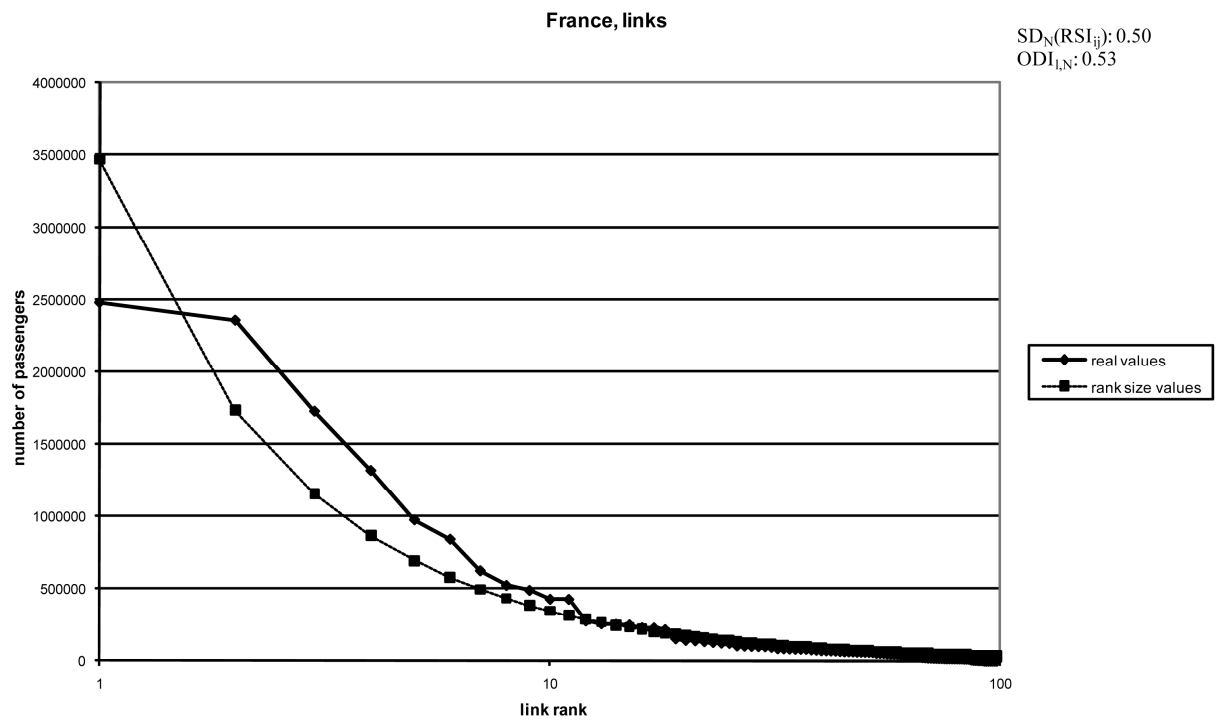


Figure 5: Countries, link level



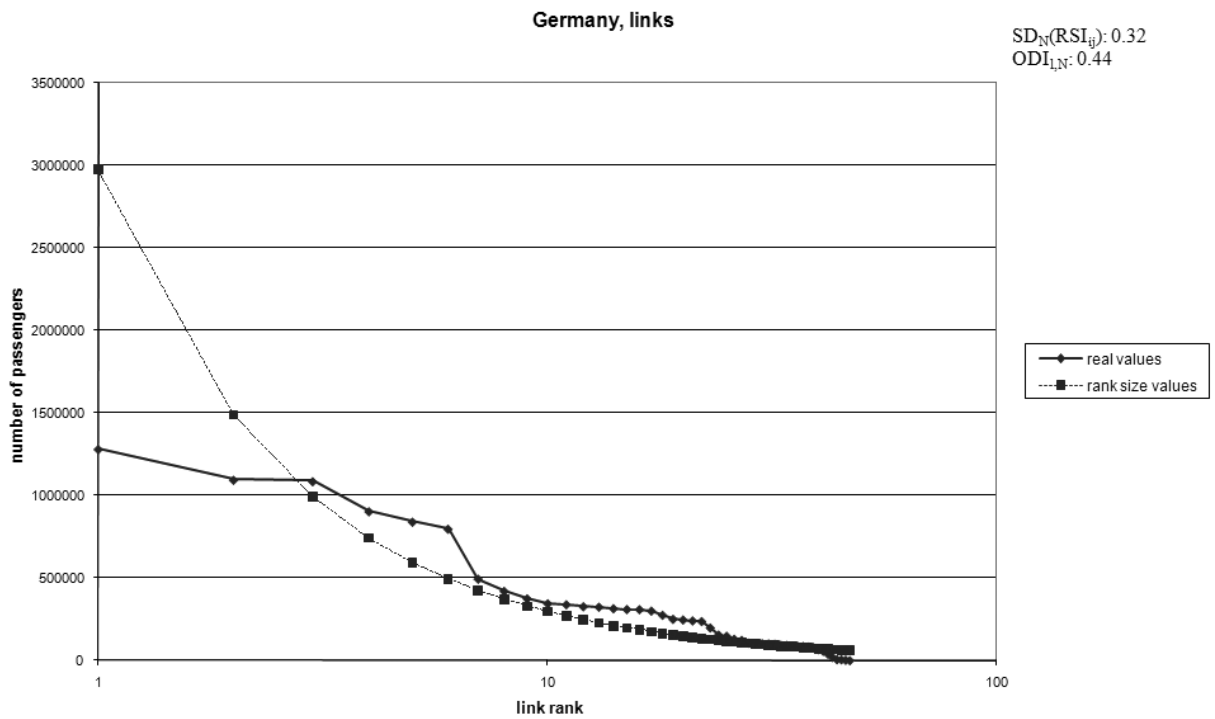
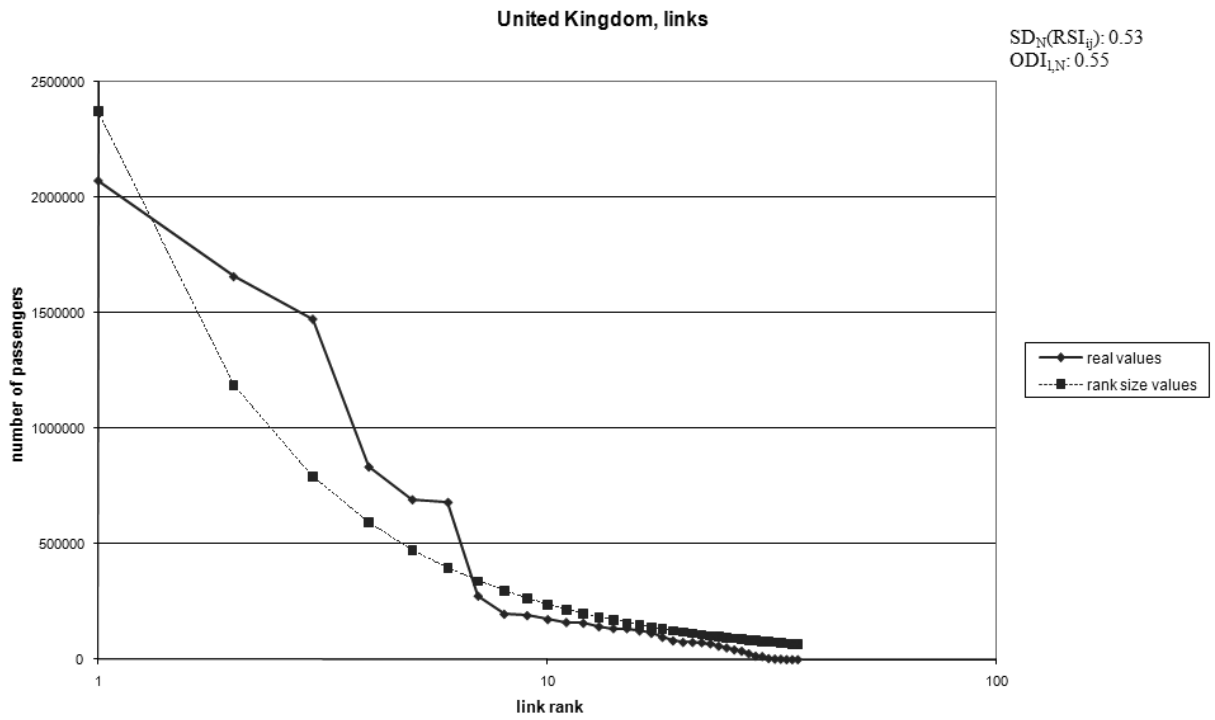


Figure 6: Service classes, city level

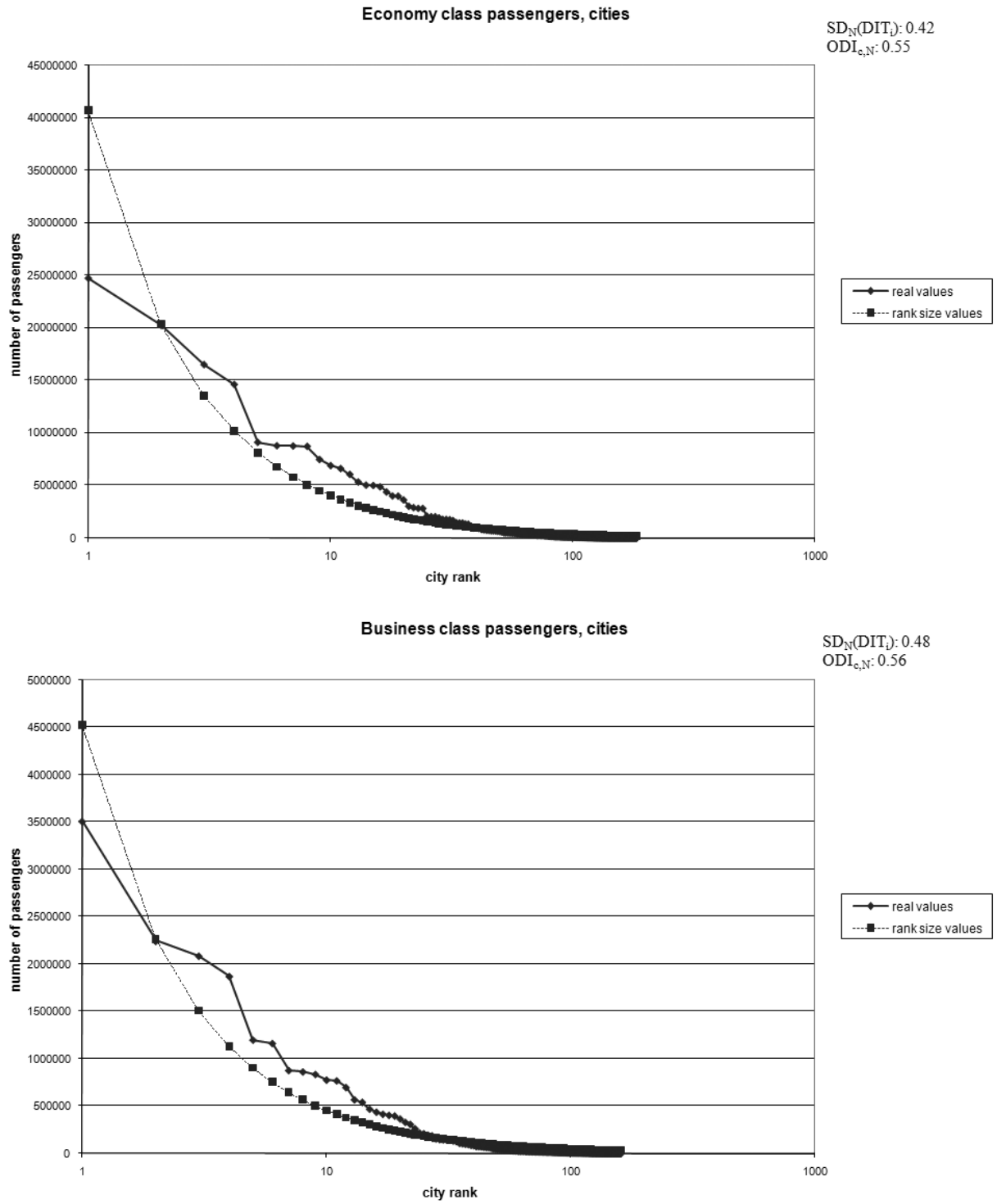


Figure 7: Service classes, city level, log-log graph

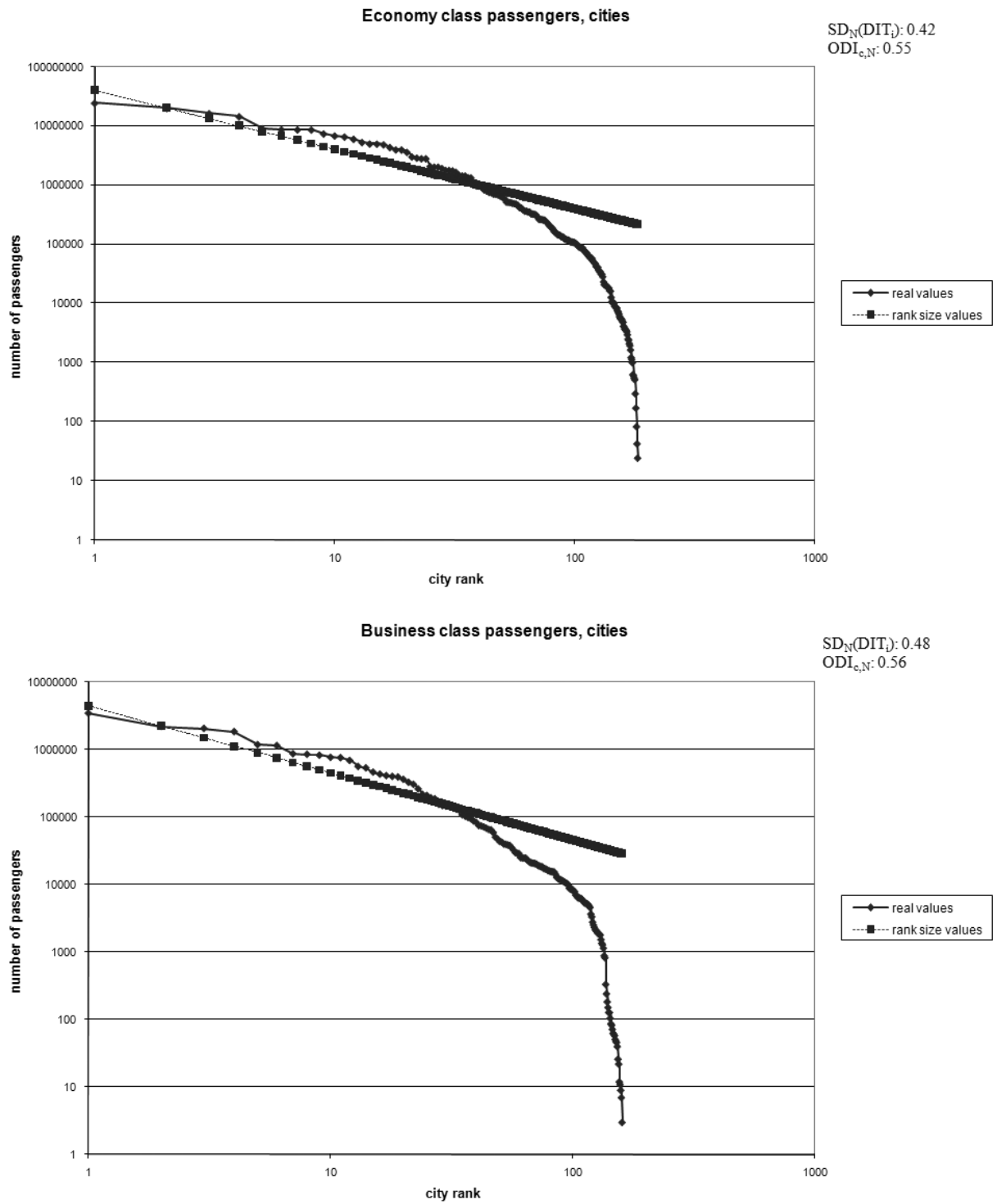


Figure 8: Service classes, link level, log-log graph

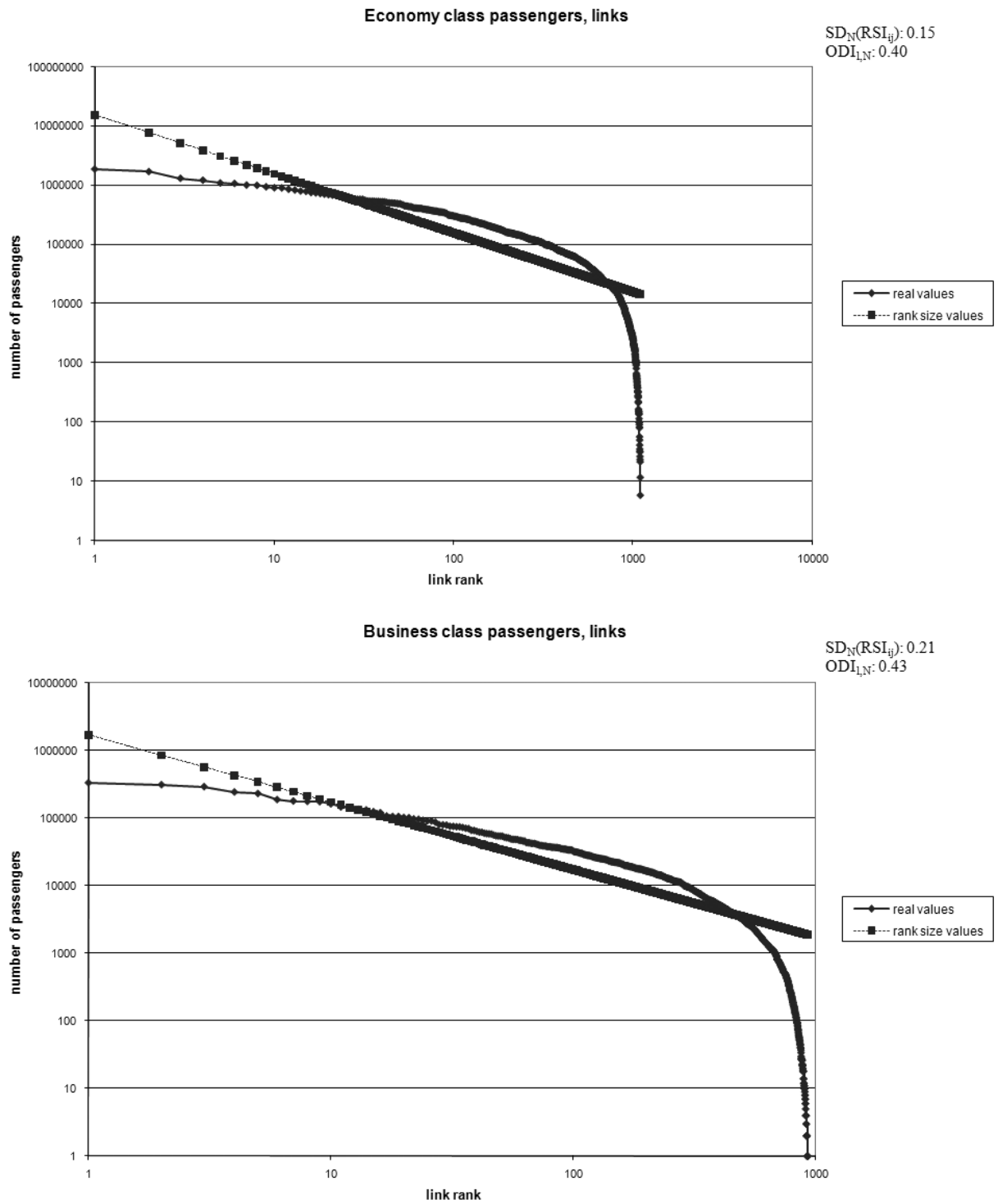


Figure 9: Service classes, link level

