Digital urban network connectivity: Global and Chinese internet patterns

Tranos, Emmanouil; Kourtit, Karima; Nijkamp, Peter

DOI:
10.1111/pirs.12097

Digital Urban Network Connectivity: Global and Chinese Internet Patterns

Emmanouil Tranos

GEES, University of Birmingham

School of Geography, Earth and Environmental Sciences

University of Birmingham, B15 2TT, UK

e.tranos@bham.ac.uk

Karima Kourtit, Peter Nijkamp

Dept. of Spatial Economics, VU University, The Netherlands

Please cite as:
Tranos, E., Kourtit, K., & Nijkamp, P. (2014). Digital urban network connectivity: Global and Chinese internet patterns. *Papers in Regional Science, 93*, 409-428.

Abstract
Cities are not only connected through conventional infrastructure, but also through digital infrastructure. This paper tests whether digital connectivity patterns follow traditional ones. Using a generalized spatial interaction model, this paper shows that geography (and distance) still matters for an extensive set of world cities. With a view to the rapidly rising urbanization, the attention is next focused on the emerging large cities in China to test the relevance of distance frictions − next to a broad set of other important explanatory variables − for digital connectivity. Various interesting results are found regarding digital connectivity within the Chinese urban system, while also here geography appears to play an important role.

Keywords: Digital Networks, Internet connectivity, World Cities, Death of Distance, Gravity Model

Running title: Digital Urban Network Connectivity

Acknowledgements: The authors would like to thank Dr. Daniel Arribas-Bel for his advice on the python implementation.

1. Prelude

Cities have over the past centuries become the cradle of political power and economic progress. And their position has become increasingly stronger in the course of history. A highlight in the history of urban systems in our world was formed by the year 2007. This year meant an important milestone in the long record of urbanization: for the first time in human history, the city took over the 'power' from its hinterland, since as of that year more than 50 percent of the world population was registered to live in urban areas. The 21st century is by some people nowadays even called *'the urban century'* (UN 2010). Surprisingly – and in contrast to this recent development – only a few centuries ago less than 20 percent of the population on our earth lived in cities. This structural urban development is still continuing, with urbanization rates exceeding 70 percent in various European countries and elsewhere (see for details e.g., Mega 2010). Although there are nowadays also signs of shrinking cities, these are rather an exception. Mulligan and Crampton (2005) find that city size evolution is tied to overall national population size, although the growth rate of any city depends on a complex myriad of private and public forces (e.g., amenities, social capital, industrial diversity, connectivity).

It is also noteworthy that cities are becoming nodal points in complex, multi-layer and often global networks. Traditionally, such nodal points were production and consumption centres that were connected by means of physical infrastructures through which material flows of people and goods could be transported. Infrastructure acted essentially as the backbone of an interconnected economy, with a fractal representation at various geographical scale levels (see, e.g., Batty and Longley 1994).

The views on the position of global, international and connected (or network) cities have significantly changed over the years. Cities increasingly act in a system of connected networks that serve as strategic alliances for the development of our world (see for an extensive urban network analysis Neal 2012). In this perspective, urban agglomerations are not necessarily a source of problems, but offer the strategic economic platform for creative solutions and new opportunities on a world-wide scale. Since the seminal work of Jane Jacobs (1961), we know that social capital (e.g., in the form of bonding and networking) and human capital (e.g., in the form of creative entrepreneurship or self-employment) are essential for smart and booming urban economies (see Nijkamp and Kourtit 2012).

In addition to a strategic re-profiling of urban areas into integrated network cities, we also observe gradually a new transformation of urban agglomerations into (regional, national or even global) spatial-economic network constellations. World-wide, urban areas are becoming centripetal and centrifugal nodes in complex multi-layer networks (Taylor 2001; Taylor 2004; Taylor et al. 2002), in which regional and national borders will play a less prominent role. This new development may turn into an urban network revolution in the history of human settlements. This may lead to the emergence of hierarchical networks or interconnected global networks of urban agglomerations. Such city networks will definitely become a source of creative and strategic research and policy action on the future of metropolitan areas.

It is thus plausible that cities in our age will likely turn into complex connected networks ('network cities in city networks') and will accommodate a rising share of socio-economic activities of a nation as a result of proximity and density externalities. These spatial urban constellations have been studied in the urban science literature from a variety of different analytical perspectives (see Nijkamp 2008 for a

review). Urban agglomerations and networks will become the cornerstones of global interaction and evolution.

In parallel with a rapid rise in urbanization rates, another megatrend has emerged, viz. the transition to a 'digital economy', thanks to the introduction and large-scale penetration of Information and Communication Technologies (ICT) in all sectors of the economy. This phenomenon has induced an intense debate on the spatial consequences of these modern technologies, which has led to various metaphors such as 'tele-cities' or 'electronic cottages'. For a critical review, see, inter alia, Cohen and Nijkamp (2006) and Cohen et al. (2002). In a more challenging way, the above-mentioned debate can be summarized under the heading of the validity of the 'death of distance' hypothesis (Cairncross 2001). Despite the fuss related with this hypothesis, it has also been increasingly questioned (see Wang et al. 2003). A test of this hypothesis calls for solid empirical research (see also Gorman and Malecki 2002) as only limited evidence can be found in the literature. For instance, we know that gravitational forces are still valid for the formation of personal social networks (Mok et al. 2010), for the purchase of virtual goods having no trading costs (Blum and Goldfarb 2006) and on the formation of Internet's physical infrastructure (Tranos and Nijkamp 2013; Tranos 2013).

The debate on the impacts of ICT on cities has already a respectable history. It follows – from a social-functional perspective – the Castells (1996) thesis on 'space of flows', while it has also strong roots in the regional-economic analysis of ICT impacts (see e.g. Cohen et al. 2002). ICT, just like ordinary infrastructure, provides the necessary spatial framework for the development and existence of urban systems at various levels. A limited number of studies regarding the impact of traditional and digital infrastructure in emerging economies – in particular China – can also be recorded. We refer here to Démurger (2001), who offers an impact assessment of (general) infrastructure and regional growth in China, to Ding and Haynes (2006), who study the leapfrogging implications of ICT in China, and to Ding et al. (2008), who researched the relation between telecommunications infrastructure and income convergence. Advanced infrastructure appears to be critical in all cases. An under-investigated issue however, is the question how cities and urban networks are related to digital spatial connectivity. And this will be the main research challenge in this paper. The central methodological task is to investigate whether in a spatial interaction model, where the spatial interaction refers to the digital infrastructural capacity (e.g. the digital links which form the Internet), the standard gravity model still holds. The main aim of our study is now to test whether the empirical connectivity pattern reflected in Internet infrastructural capacity leads to a statistically significant model based on standard gravitational forces in relation to a spatial interaction model. A second aim of our study is to analyze whether the digital connectivity in urban systems in rapidly emerging economies – in particular in China – can also be appropriately mapped out by a spatial interaction representation of Internet linkages. Also, when necessary, comparisons with other world regions (e.g. Europe) take place. In summary, the purpose of our article is to explore the symbiotic relation between urban systems and digital infrastructural networks, both world-wide and within China.

China – as an important player in both Asia and world-wide – is an interesting case because of its rapid economic and technological growth. Many Asian countries –including China− have in the past decade exhibited a surprisingly high economic growth. At the same time, many of these countries have

shown very high urbanisation rates, to the extent that many large to very large cities can be found in Asia nowadays. The unprecedented rise in megacities in Asia is partly caused by their indigenous growth mechanisms and partly by their high-quality connectivity. A recent study of the UN (2010) demonstrated that from the world's largest cities (30 in total), 17 such megacities are located in Asia. Urban agglomeration externalities are apparently so powerful in this region, that an unprecedentedly strong urbanisation megatrend is emerging (Kusakabe 2012; Morichi and Raj Acharya 2013).

China has not only witnessed a formidable economic growth in the past decades, but also a surprising rise in its urbanization degree, with many new megacities. For that reason, infrastructure policy is of paramount importance in this country. But physical connectivity is not sufficient; digital connectivity in an information society driven by digital infrastructure is equally important. And therefore, a closer analysis of the geographical structure and intensity of the usage of digital technologies in China (in particular, the Internet), is a challenging research task (see also the work of Derudder et al. 2013). In our empirical study we will address in particular the digital connectivity among major urban agglomerations in China, including not only the big megacities such as Shanghai or Beijing, but also many lower ranked – but often still multi-million – cities (Zhen et al. 2013).

Methodologically, apart from econometrics to understand the impact of gravitational forces in the evolution of the digital infrastructure, concepts and techniques from the network analysis field are also utilized. The latter has become an advanced research area, starting form Euler's well-known Königsberg bridge system to small-world or scale-free networks. In the recent geography and regional science literature much attention has been paid to spatial linkage analysis and spatial interaction models, in which also the interwovenness of cities has been studied extensively (including hierarchical organization of cities, e.g. in the context of central place theory or Zipf's law). With the introduction of digital technology, new types of connected networks have emerged, often of a hub-and-spokes nature, with various degrees of user intensity on different edges. Common characteristic of some these networks is their spatial reflection. The methodological novelty of this paper is the utilization of network analysis with econometrics to understand whether the gradually emerging global network connectivity pattern is also replicated at the level of an upcoming economic region such as Asia, and in particular whether such a pattern is showing up in a rapidly emerging country like China.

The present paper is organized as follows. After this introductory section, Section 2 will concisely describe the database for the models to be used, while Section 3 will be devoted to a modelling experiment on digital connectivity among world cities. China is a rising player in a global internet connectivity system, and therefore, in Section 4, our study will zoom in on the Chinese urban network, also in comparison to Europe . An explanatory causal econometric model for digital connectivity in China – and its results – will be presented in Section 5, while Section 6 offers concluding remarks.

2. Data for digital connectivity

The main database used for this paper has been derived from the DIMES project. This is "a distributed scientific research project, aimed to study the structure and topology of the Internet"

(DIMES 2010). It is based on 3–6 million traceroute[1] measurements made daily by a global network of more than 10,000 agents, who are voluntarily participating in this research project (see Carmi et al. 2007; for a description of the DIMES project, see also Shavitt and Shir 2005). One of the outcomes of the DIMES project is an extensive database with geo-located IP (Internet Protocol) links discovered by the DIMES volunteers. It contains all the IP links between any two cities discovered by the agents. Although overlapping connections between any two regions are included in the database, there is no information on the bandwidth of these links. However, this is still an infrastructural measure, as the IP links represent physical (overlapping) data links between cities, which follow the IP protocol[2].

Some important notes should be made at this outset. Firstly, the DIMES project only includes IP links which have been captured by its agents and thus, only a small fraction of the total Internet. By sending data packets from the agents' locations to known destinations, DIMES researchers record the different IP links used by its agents, completing the largest available data set for geo-coded IP links (Tranos and Nijkamp 2013).

Secondly, there is a common limitation faced by any study focusing on the Internet from a spatial perspective: the Internet has been built as a logical network and its links are defined in topological and not in geographic terms. Therefore, the architecture of Internet destinations (IP addresses) has little to do with geographical locations (Dodge and Zook 2009). To geo-locate the above system, the DIMES project geo-codes the different IP addresses using IP registration tables. A potential accuracy issue needs to be highlighted here. It is not uncommon that IP addresses are owned by expert firms, which lease these IP addresses to content providers (Dodge and Zook 2009). This might result in a possible mismatch between some physical locations of IP addresses and the content location. However, this does not create any bias here, as the focus of this paper is on the physical infrastructure of the Internet (Tranos and Nijkamp 2013).

Different subsets of this global DIMES dataset are used for this paper. For the analysis in Section 3, which focuses on digital connectivity and world cities, the IP links among a sample of 34 world cities are utilized for the year 2010. This analysis is limited to a cross section of these 34 cities for one year because of the scarce data on world city characteristics. More details about this data are provided in the next section. After the global analysis of linkages between cities world-wide, in Sections 4 and 5 the analysis turns to the Chinese urban system, which includes IP links only between Chinese cities. This analysis follows panel specifications and includes the period 2007-2011.

In a nutshell, the DIMES dataset is, at least to the best of our knowledge, the richest available geographical data source for the Internet infrastructure. Despite the above limitations, the scattered locations of the agents and the size of the DIMES experiment secure the robustness of this data set, especially considering the general lack of geographic data on the Internet infrastructure.

---

[1] Traceroutes are specific programs, which map the route that a data packet follows through different nodes in order to reach its final destination (Dodge and Zook 2009).

[2] These links function at level 3 of the OSI model. As noted elsewhere (Tranos 2013), the first three layers of the OSI model represent physical infrastructural capital, while the four highest layers reflect 'infratechnologies' (Tassey 1992; Tassey 2008).

3. Digital connectivity and global cities: a high-level analysis

Before focusing on the main object of our analysis, viz. the digital infrastructure of the Chinese urban system, the global nature of our dataset is utilized in order to provide an overall, global context for urban digital connectivity. As discussed in the previous section, ICT and the Internet support the globalization process and global cities are increasingly reliant upon digital infrastructure. Nonetheless, there are hardly any studies at a global level linking the global city characteristics with the digital infrastructure. The only exception at this scale is the work of Choi et al. (2006) who investigated the network structure of the Internet backbone networks among the most well-connected world cities[3]. In order to fill in this gap and to provide a broader understanding of the relation between world cities and the underlying first layer of the space of flows, this section employs simple spatial interaction models (SIMs) to investigate the pull factors for attracting such digital infrastructures in global cities. The conceptual model of this analysis is formulated in the following generalized version of a SIM, according to which the number of IP installed links between *i* and *j* (IP_{ij}) is affected by the characteristics of *i* (X_i) and *j* (X_j) as well as by bilateral characteristics between *i* and *j* (X_{ij}).

$$IP_{ij} = f(X_i, X_j, X_{ij}) \quad (1)$$

The main limitation for such an endeavor is data availability, as hardly any homogeneous urban data is available at a global, cross-country level. In order to overcome this difficulty, a unique dataset depicting urban characteristics of 34 world cities is utilized here produced by the Institute for Urban Strategies (2010). This Global Power City Index (GPCI) offers a balanced picture of the socio-economic performance and power of 34 world cities[4] from the perspective of attracting talent, business and investment to cities, complemented with information on perceptions of various classes of stakeholders. Based on 69 individual indicators compiled from secondary sources as well as from interviews with stakeholders, two sets of indicators have been produced for the year 2010[5]: city *function* indicators, which include a normalized score on variables focusing on urban accessibility, economy and the environment; and city *actors* indicators including variables on how managers, researchers, artists and residents perceive and score the performance of the city (see also Arribas-Bel et al. 2013).

Apart from the variables derived from the actors and functions data, the impact of variables related with the spatial organization of the world cities sample is also tested here. Firstly, physical distance between cities is expected to have a negative impact on the pair-level IP connectivity. As discussed elsewhere (Tranos and Nijkamp 2013), physical distance (*distance*) maintains its importance even in the frame of the digital infrastructure. Similarly, the spatial continuity (*continuity*) between the countries of the cities included in the analysis is expected to have a positive impact on digital connectivity. Moreover, we employ variables from the world trade literature (Mayer and Zignago 2005)

[3] For a US-centric study on similar issues see the work of Malecki (e.g. 2002) and the work of Tranos for a pan-European perspective (e.g. 2011; Tranos and Gillespie 2009)

[4] Mumbai was also included in the GPCI, but it is excluded from our analysis as no IP data was available from the DIMES project. The rest of the cities included in the GPCI are presented in Table 3.

[5] For a detailed review of the GPCI index see Institute for Urban Studies (2010). A table of descriptive statistics is presented in the Appendix.

and we expect variables such as common language (*language*) and past colonial (*colonial*) ties will affect the digital connectivity (Tranos and Gillespie 2011).

In order to utilize these variables, model (1) is expanded in the following log-log form:

$$\ln(IP_{ij}) = a_0 lnk + a_1 \ln(X_i * X_j) + b_1 \ln\left(distance_{ij}\right) + b_2 relational_{ij} + \varepsilon_{ij} \quad (2)$$

X_i and X_j are the variables reflecting the city-level attributes discussed above. Because the dependent variable reflects infrastructural capacity and not flows, there is no directionality involved and therefore instead of estimating the effect of *i* and *j* separately, their combined impact is estimated using the product of X_i and X_j. Since the city-level attributes are only available for year 2010, equation (2) is estimated cross-sectionally using ordinary least square (OLS). Table 1 presents the results of the actors variables and Table 2 the results of the function variables.

Table 1: World city IP connectivity and actor-based city characteristics

Dep. Var. IP (ln)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
distance (ln)	-0.427	-0.448	-0.478	-0.453	-0.485
	(0.0512)***	(0.0506)***	(0.0492)***	(0.0492)***	(0.0554)***
manager	0.382	0.138	0.372	0.383	0.319
	(0.0537)***	(0.0768)*	(0.0835)***	(0.0826)***	(0.0848)***
researcher		0.205	0.359	0.265	0.241
		(0.0468)***	(0.0517)***	(0.0574)***	(0.0577)***
resident			-0.569	-0.635	-0.608
			(0.0922)***	(0.0930)***	(0.0933)***
artist				0.229	0.265
				(0.0635)***	(0.0651)***
continuity					-0.512
					(0.250)**
language					0.463
					(0.169)***
colonial					0.166
					(0.230)
constant	-1.512	-0.482	3.519	2.472	2.989
	(0.896)*	(0.912)	(1.095)***	(1.122)**	(1.130)***
N	561	561	561	561	561
R-squared	0.198	0.225	0.275	0.291	0.305

* *p*<0.1; ** *p*<0.05; *** *p*<0.01; Standard Error in parentheses

The first observation is the consistent negative effect of distance on the formation of the digital infrastructure at a global level. Put simply, the closer two world cities are in our sample, the more digital infrastructure is installed between them. This distance decay effect remains significant and its magnitude increases even after the inclusion of other bilateral variables including spatial continuity. The researcher variable has a positive effect which appears to be significant throughout most of the

specifications. This is not surprising either, as digital infrastructure was always related with knowledge-intensive urban environments (Malecki 2002). The same applies for the managerial effect, which is also positive and significant. Indeed, the higher the (product of two cities') score on managerial issues is, the higher the connectivity between these two cities is. Moreover, a significant and consistent negative effect is detected for the (product of the) score of cities on residential issues. This effect can be interpreted as a city-size effect: the higher the size of a city and consequently the diseconomies of scale (low score on residential issues), the higher the digital connectivity the city shares with other cities. On the contrary, creativity appears to be a significant positive factor for attracting digital infrastructure. The (product of the) score of two cities according to artists is a positive predictor of the digital connectivity between these two cities. Regarding the other bilateral variables, spatial continuity has a negative impact, reflecting a minimum distance threshold for the intensification of the digital links.

Table 2 presents the estimation of (2) using the function variables. Again, interesting results can be derived regarding the distribution of the digital infrastructure among our sample of world cities. Firstly, spatial configuration appears to be important even at this scale, as apart from the distance decay effect which is present here too, accessibility also has a significant positive effect: the more accessible two cities are (in other words, the higher the product of the accessibility of two connected cities is) the more digital infrastructure will be installed among them. On the contrary, although the variable reflecting the score on the urban economy is positive, its effect is of low significance when the bilateral variables are included in the analysis. This is in accordance with previous research highlighting that such infrastructure is mostly attracted by knowledge-economy related indicators instead of mere market size (Tranos and Gillespie 2009). Next, the environment variable appears to confirm the above comments on the effect of the *resident* variable. The city-size effect, as reflected in diseconomies of scale and related low urban environmental quality, is a digital connectivity factor. Regarding the bilateral variables, the same effect as in Table 1 is observed here.

Table 2: World city IP connectivity and function-based city characteristics

Dep. Var. IP (ln)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
distance (ln)	-0.475	-0.325	-0.337	-0.347	-0.376
	(0.0529)***	(0.0505)***	(0.0506)***	(0.0507)***	(0.0569)***
accessibility		0.491	0.420	0.409	0.421
		(0.0469)***	(0.0567)***	(0.0567)***	(0.0575)***
economy			0.100	0.121	0.0791
			(0.0452)**	(0.0461)***	(0.0471)*
environment				-0.0828	-0.0856
				(0.0386)**	(0.0382)**
continuity					-0.533
					(0.244)**
language					0.556
					(0.164)***
colonial					0.227
					(0.229)
constant	4.084	-3.089	-3.391	-2.108	-1.498
	(0.447)***	(0.798)***	(0.807)***	(1.002)**	(1.023)

N	561	561	561	561	561
R-squared	0.126	0.269	0.275	0.281	0.301

* *p*<0.1; ** *p*<0.05; *** *p*<0.01; Standard Error in parentheses

In total, the above analysis has revealed many interesting patterns at the global level of the world city network. Digital infrastructure is clearly affected by spatial configuration – even at this scale. Thus, it can be argued that also a digital system such as the Internet is ruled by strong spatial forces. In addition, other factors related to the knowledge economy and city size appear to play also important roles in the geography of this system. The next step in our analysis provides the link between the global digital network and the Chinese urban system. China is gradually moving from a lower wage and mass-production oriented economy to a modern, global and high-tech dominant economy, in which the digital technology – especially in its emerging urban system, not only in the coastal area, but also in its central region – plays a critical role (Dai 2003). Development and usage of ICT – ranging from hardware production to software use – has become an essential element of China's innovation policy.

4. The Chinese digital urban network

Moving from the global level of analysis to the Chinese urban system, the first task is to understand the position of Chinese cities in this global system of world cities and then to comprehend the structure of the Chinese inter-urban network of the digital infrastructure. In order to do so, concepts and methods from the complex network analysis (CNA) field and the *science of networks* (Barabási 2002; Buchanan 2002; Watts 2003; Watts 2004) are utilized. This is a new analytical field which focuses on large-scale real-world networks and their universal, structural, and statistical properties (Newman 2003). CNA is the tool which enables us to explore connectivity patterns in the topological configuration of the Chinese digital infrastructure. The latter is an essential step in order to move on to the next part of our analysis, where the structure of the urban network in China is modelled.

Table 3 presents three different centralities for the sample of 34 world cities. Firstly, degree centrality represents the accumulated IP links for each city and for each year[6]. This is a digital *infrastructural* capital measure, in which Beijing shows up as the third most connected city in our world city sample for 2010. Shanghai, the other Chinese city in our sample, is placed on the 15th position. Although this is an important measure reflecting the accumulated IP connectivity, degree centrality does not provide any insights into the functionality of these cities in the overall network.

In order to shed light on how these cities are placed in the global IP networks, two more centrality indicators are introduced here: betweenness and eigenvector centrality. Common characteristic of these measures is that instead of focusing only on direct network neighbours, they incorporate indirect links. Freeman (1978-1979, p. 224) identified betweenness centrality as an "index of the potential of a point

[6] This is a 'weighted' degree centrality measure in the sense that if two regions *i* and *j* are connected by multiple links, all of these links will be added in the degree centrality of *i* and *j*. If it had been a 'binary' centrality measure, then the multiplicity of the links between *i* and *j* would have been neglected.

for control of communication" and Nooy et al (2005, p. 131) defined it as the "proportion of all geodesics between pairs of other vertices that include this vertex". Because of its focus on indirect links, betweenness centrality fits better with the structure of the Internet data packet transport systems, which is based on packet switching through indirect routes (Tranos 2011). Going a step further, eigenvector centrality does not consider all connections to be equal as links to more central vertices are more important than connections to less central ones. It is still important to have a large number of connections, but a node with fewer but more important connections will end up being more central than a node with more, but less important links (Newman 2008). Thus, eigenvector centrality considers the global structure of the network by incorporating both direct and indirect links (Bonarich 2007; Hanneman and Riddle 2005). In total, it could be said that while eigenvector reflects hub roles, betweeness centrality represents gateway functions.

These metrics are useful in understanding the city functionalities in such a global system. On the one hand, high betweenness centrality is related with a high number of networks passing through a node and therefore to hub urban roles. On the other hand, high eigenvector centrality reflects a power position of gateway cities, which can control access to cities of high connectivity (see the discussion in Neal 2011). Thus, Beijing's high performance in both metrics indicates the Chinese capital's importance in the global digital network as Beijing performs both hub and gateway roles. However, this is not the case for Shanghai, whose' functionality is lower than the expected one according to accumulated IP infrastructure, as is reflected in the degree centrality.

Table 3: World-cities centralities in 2010

Cities	**Degree**		**Eigenvector**		**Betweenness**	
	score	***rank***	**score**	***Rank***	**score**	***Rank***
London	100.00	*1*	100.00	*1*	79.57	*2*
Seoul	94.22	*2*	0.22	*22*	16.62	*13*
Beijing	**75.77**	***3***	**28.13**	***3***	**20.61**	***10***
New York	47.19	*4*	54.20	*2*	55.97	*3*
Frankfurt	46.64	*5*	25.85	*4*	43.20	*5*
Tokyo	41.95	*6*	0.11	*23*	18.86	*12*
Amsterdam	33.72	*7*	10.00	*11*	44.56	*4*
Madrid	29.67	*8*	8.22	*12*	32.16	*6*
Moscow	26.83	*9*	14.49	*7*	100.00	*1*
Toronto	22.28	*10*	11.12	*10*	10.58	*17*
Paris	21.95	*11*	18.29	*5*	26.40	*7*
Singapore	19.57	*12*	4.32	*15*	13.05	*15*
Chicago	17.73	*13*	15.17	*6*	14.45	*14*
Los Angeles	13.91	*14*	14.10	*8*	3.34	*23*
Shanghai	**13.23**	***15***	**0.44**	***21***	**6.35**	***21***
Sydney	13.05	*16*	0.03	*30*	23.12	*8*
San Francisco	12.10	*17*	12.95	*9*	6.79	*18*
Vienna	11.10	*18*	0.08	*25*	19.33	*11*
Milan	11.07	*19*	6.18	*13*	12.55	*16*

Taipei	9.53	*20*	2.93	*18*	6.61	*19*
Zurich	5.93	*21*	0.09	*24*	0.00	*33*
Osaka	5.56	*22*	0.01	*33*	1.59	*26*
Cairo	5.32	*23*	0.03	*28*	1.57	*27*
Brussels	5.28	*24*	1.99	*20*	20.64	*9*
Sao Paulo	4.69	*25*	0.02	*31*	0.00	*34*
Kuala Lumpur	4.57	*26*	0.01	*32*	2.43	*25*
Bangkok	4.44	*27*	2.02	*19*	6.56	*20*
Vancouver	3.94	*28*	4.49	*14*	0.30	*31*
Boston	3.11	*29*	3.12	*17*	3.16	*24*
Copenhagen	2.94	*30*	3.84	*16*	1.16	*29*
Berlin	2.78	*31*	0.04	*26*	3.40	*22*
Geneva	1.70	*32*	0.03	*29*	1.36	*28*
Fukuoka	1.22	*33*	0.00	*34*	0.40	*30*
Hong Kong	0.02	*34*	0.04	*27*	0.01	*32*

Note: centrality measures are normalized with maximum value = 100

After highlighting Beijing's role as the main anchor point of the Chinese IP network with the global one, our focus will next turn to the Chinese inter-city digital network. Table 4 presents some basic network statistics for the Chinese digital infrastructure for the period 2007-2011. The size of the network during the first two years of the study period is less than thirty per cent of the network size during the last three years. Although this might reflect, to some extent, the Internet growth in China, much of this is mostly related to the data collection process and the increase of the DIMES project agents. Nonetheless, a change in the topology of the network can be observed. The first statistic under study is again degree centrality[7]. The large difference between the average and maximum values reflect the existence of some very well connected nodes, which perform hub roles in the network. Regarding the change over time, while the average weighted degree centrality among the connected cities increased almost four times during the study period, the maximum degree centrality increased more than twenty times. This is a first indication of the existence of a *cumulative causation* process, according to which the higher the degree of a node is, the higher the probability of a new link to be attached to this node is. This *rich get richer* phenomenon results to high inequality in terms of connectivity among the Chinese cities, which increases over time according to the Gini coefficient of the degree centrality. In the network literature, this cumulative process is identified as preferential attachment (Batty 2012) and it will be further analyzed below.

Table 4: Network statistics for China's IP network

Net. Statistics	**2007**	**2008**	**2009**	**2010**	**2011**
Nodes	219	224	781	780	784

[7] Just as before, this is a weighted degree centrality measure.

Links	420	432	1557	1606	1638
Av. degree centrality	56.555	90.951	144.095	257.805	249.704
Max. degree centrality	2830	4671	32660	57889	59490
Density	0.017	0.017	0.005	0.005	0.005
Av. network distance	2.515	2.71	2.944	2.884	2.825
RN av. network distance	*4.078*	*4.082*	*4.801*	*4.821*	*4.818*
Clustering coefficient	0.462	0.351	0.364	0.393	0.408
RN clustering coefficient	*0.012*	*0.011*	*0.006*	*0.004*	*0.004*
Gini coefficient	83.77	89.06	93.4	93.71	93.83

Note: RN is a random network with the same number of nodes and links

The outcome of the uneven distribution of the IP links among the Chinese cities is an *efficient* digital network. Indeed, despite the very low density of the IP network, which decreases over time, the average network distance is exceptionally short. In the CNA framework, distance does not refer to Euclidean distance, but to the number of nodes that separate any two nodes[8]. For the case of the Chinese digital infrastructure, any two cities are separated on average by two intermediate nodes, which results in a network distance less than 3. The latter is an indication of efficiency, as it reflects the ability of the network to transfer data flows with minimal routing.

The above qualities and the efficiency of the network can be attributed to the *small world* (SW) characteristics of the Internet infrastructure. The latter refers to a widely used network model, whose main characteristic is the existence of highly-connected clusters of nodes, which gain global connectivity via a few links that span the entire network, linking distant clusters (Watts and Strogatz 1998). This theoretical network model became popular because of its many real-world applications. The digital infrastructure in China resembles SW networks because of the short average distance – shorter than the ones observed in same size random networks (RN)[9] – and the high clustering coefficient[10] – higher than the ones discovered in same size RN.

[8] Because there are usually numerous different ways to connect any two given nodes (known as *walks*), research commonly focuses on the shortest path, known as distance (Nooy et al. 2005).

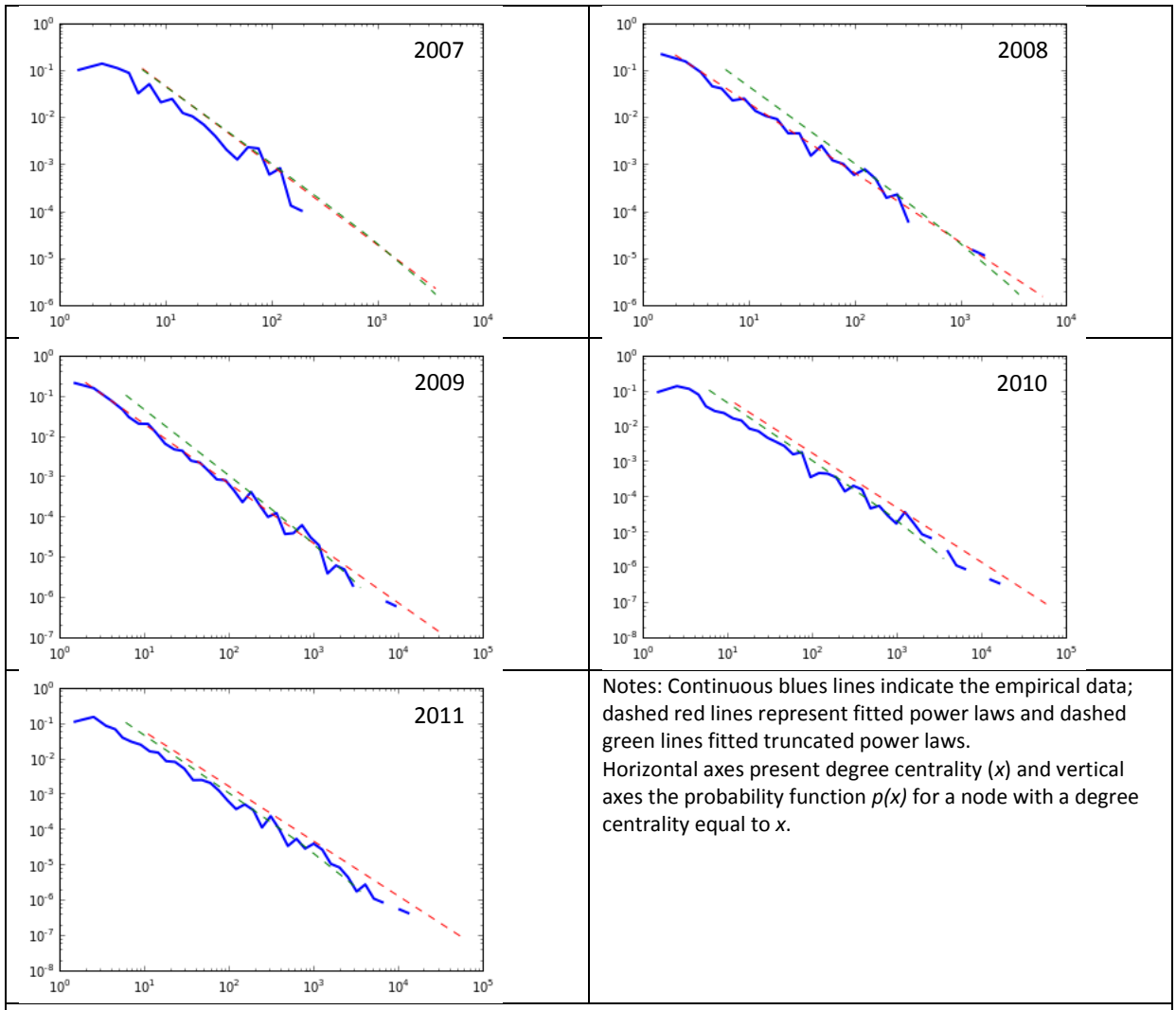
[9] RN were introduced by two Hungarian mathematicians Paul Erdös and Alfréd Rényi and refer to large scale networks with no obvious structure (Erdös and Rényi 1959). The distribution of vertices degree follows a Poisson distribution, which means that the majority of the vertices on the network have the same number of links and they are found nearby the average degree <k>; vertices that deviate from this are rare.

[10] The clustering coefficient C_i of node i is the ratio between the number of edges E_i that exist among its nearest neighbours (nodes which are directly connected with node i) and the maximum number of these edges, where k_i is the number of nodes in clique: $C_i = {2E_i}/{k_i(k_i - 1)}$. A clique could be understood as a completely connected sub-

Apart from the latter, an essential element of the SW networks is the distribution of nodes' degree centrality, which distinguishes this network type from another well-established network model known as *scale free* (SF). SF networks share the above characteristics with SW networks, but the degree distribution of their nodes follows a power law, contrary to the exponential functions which distinguish SW networks. The distinct distributions reflect the difference between these two types of networks in terms of nodes' heterogeneity: while the power-law degree distribution of the SF networks reflects hub and spoke structures with a very few super-connected hubs which hold the network together and a vast majority of less-connected nodes (Barabási and Albert 1999), the exponential degree distribution of SW networks resembles highly-connected cliques and less heterogeneous nodes in terms of degree centrality. Power law distributions appear to be the outcome of the main generative mechanisms behind SF networks: growth and preferential attachment. Real-world networks tend to evolve and grow over time, but the new links are not distributed randomly in the network. On the contrary, the likelihood of a node to attract new links depends on the node's degree centrality (ibid).

The continuous lines in Figure 1 present the probability density functions (PDF) for the node degree distribution in China during the period 2007-2011. As it can be seen by the log-log plots, the Chinese inter-city IP network follows a heavy tail distribution and the PDFs can be approximated by straight lines, which is a first indication of a heterogeneous SF distribution. In order to test this visual observation, we follow the methodology suggested by Clauset et al. (2009) and the python code implementation by Alstott et al. (2013). The first step is to identify a minimum degree centrality value (x_{min}), above which the SF distribution applies. It is common that power laws and other heavy-tailed distributions do not apply for small values of the data (in this case, for nodes with low degree centrality). In order to identify the x_{min} Clauset et al. (2009) suggest the use of the Kolmogorov-Smirnov (KS) statistic, which is the maximum distance between the real data and a SF fitted model. KS is calculated for each unique value in the dataset and then x_{min} is defined as equal to the value which has resulted the smallest KS statistic. Then, power laws are estimated for $x \geq x_{min}$ using *maximum likelihood*. Nevertheless, it might be the case that there are other alternative distributions that fit the empirical data better than the estimated power law. Therefore, the normalised log likelihood ratio test is implemented: a positive sign of this indicator advocates towards the superiority of a power law, while a negative sign towards the superiority of the alternative distribution which is tested, given a p-value less than 0.1 (ibid).

graph. In this case, the clustering coefficient for the entire network is calculated as the mean of the all the node clustering coefficients.



Notes: Continuous blues lines indicate the empirical data; dashed red lines represent fitted power laws and dashed green lines fitted truncated power laws.
Horizontal axes present degree centrality (*x*) and vertical axes the probability function *p(x)* for a node with a degree centrality equal to *x*.

Figure 1: Probability Density Functions for the Chinese cities' degree centrality

Table 5 presents the results of this analysis. Apart from estimating the x_{min} and the power law exponent (equation 3 bellow), two alternative distributions are also tested: exponential and truncated power law (equations 4 and 5 as well as dashed lines in Figure 1). The theoretical justification for this choice lies in the bellow lines: (a) exponential distributions, apart from reflecting more homogenous nodes in terms in degree centrality, are related with SW networks, which are also common among real world network; (b) following Amaral et al (2000), it is also not uncommon for real world networks to be characterised (for every $x \geq x_{min}$) by a power law distribution followed by an exponential cut-off. This exponential cut-off reflects various constraints which prevent the evermore implementation of the preferential attachment mechanisms.

$$p(x) \propto x^{-a} \qquad (3)$$

$$p(x) \propto e^{-\lambda x} \qquad (4)$$

$$p(x) \propto x^{-a} e^{-\lambda x} \quad (5)$$

Table 5 presents the estimation of the x_{min} and the power law exponent based on this value. Based on the normalised log likelihood ratio, it can be safely said that power laws describe better the digital inter-city network than exponential laws and therefore the Chinese digital network does not fit with the SW definition. However, this is not the case with the truncated power law, which appears to be superior to the power law. Following the typology offered by Amaral et al. (2000), the Chinese inter-city IP networks can be characterised as broad scale networks. These are heavy tailed networks, with a heterogeneous distribution of connectivity. Hierarchy is a strong element of these networks and very few nodes (cities in our case) perform important essential hub roles. In spatial terms, this can be interpreted as an agglomeration effect of the digital connectivity in a limited number of cities which act as the main hubs for the Internet infrastructure in China. Nevertheless, constraints prevent the absolute implementation of preferential attachment mechanisms reflected in clear SF distributions. Such constraints might be related with the maturity level of a network and the inadequate growth rate or with limitations related with the ability of the hubs to keep attracting new links. These limitations can be the outcome of economic or physical constraints because of a threshold in the capacity of a city to attract and facilitate new IP links.

In order to understand the importance of such constraints, we compare the Chinese digital network with the equivalent one for the European cities. Using the data from Tranos and Nijkamp (2013), the same methodology is applied and the results are presented in Table 5. Moreover, the PDFs for the European network for the years 2005 and 2008 are presented in Figure 2. The European IP network exhibits the same characteristics with the Chinese as truncated power laws approximate better the node degree distribution. The main difference can be found in the estimation of x_{min}. Indeed, while x_{min} for China varies between 2 and 12, for Europe varies between 4290 and 5932. This means that the truncated power law in Europe represents cities with much higher IP connectivity than in China. In quantitative terms, a truncated power law represents 43 per cent of the interconnected cities in China in 2011 and only 20 per cent in Europe in 2008. The remaining – left – part of the distribution ($x \leq x_{min}$), is characterised by a more homogeneous distribution, which resembles an exponential one[11]. This observation brings into light mechanisms which are present in the European digital network, but not in the Chinese one. These mechanisms may be related to national policies and national borders which prevent the formation of a SF degree distribution in the least connected part of the European urban system. These mechanisms are juxtaposed with the more centralized approach of the Chinese infrastructural planning system. Therefore, the least connected European cities are more 'protected' as they enjoy a higher level of digital connectivity than the equivalent of the Chinese digital network. In total, the Chinese inter-city digital network is more centralised as almost 50 per cent of the total IP connectivity in China is agglomerated in only three cities in 2011, Beijing, Shanghai and Guangzhou. On

[11] A replication of the above analysis for the European subnet of $x \leq x_{min}$ indicated that exponential laws are superior to power ones. The results can be provided upon request.

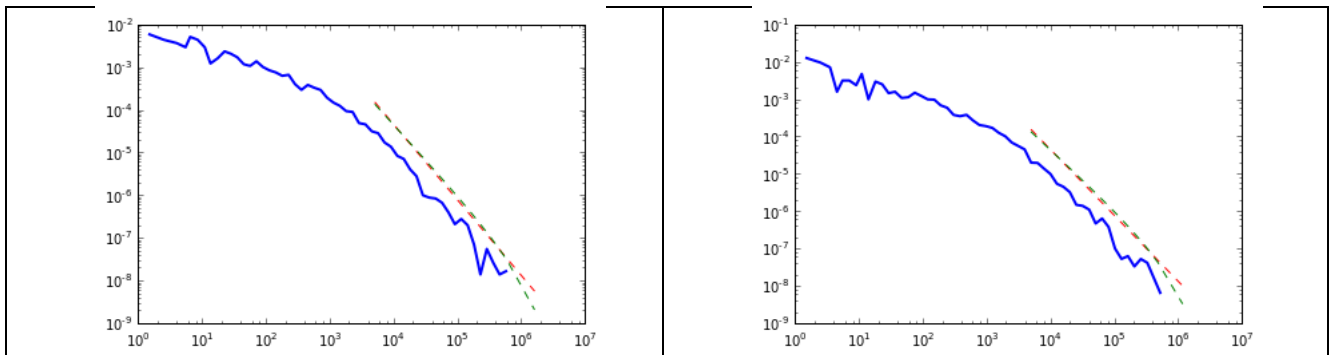
the contrary, the three most connected cities in Europe are responsible for only 8.5 per cent of European cities total IP connectivity.

The structural differences between the digital network in China and Europe reflect, to a certain extent, spatial configuration differences. Moving a step forward, the next section will explore the mechanisms behind the formation of this complex network in China.

Table 5: Power laws and alternative node degree centrality distributions

	year	x_{min}	Power law exponent	Exponential		Truncated power law	
				R	***p***	***R***	***p***
China	2007	6	1.687	3.412	0.001	-1.253	0.202
	2008	2	1.478	4.607	0.000	-2.221	0.009
	2009	2	1.480	6.313	0.000	-2.564	0.000
	2010	12	1.548	4.886	0.000	-2.015	0.007
	2011	11	1.549	4.844	0.000	-1.963	0.008
Europe	2005	5077	1.769	5.256	0.000	-2.371	0.002
	2006	5932	1.848	4.243	0.000	-1.702	0.029
	2007	4290	1.830	4.870	0.000	-1.981	0.011
	2008	4916	1.763	4.275	0.000	-2.444	0.001

Note: *R* is the normalised log likelihood ratio and *p* its p-value



2005 2008

Figure 2: Probability Density Functions for the European cities' degree centrality

Notes: Continuous blues lines indicate the empirical data; dashed red lines represent fitted power laws and dashed green lines fitted truncated power laws.
Horizontal axes present degree centrality (*x*) and vertical axes the probability function *p(x)* for a node with a degree centrality equal to *x*.

5. The determinants of digital connectivity within Chinese urban system

After gaining a structural understanding of the IP network in China, this section aims to shed more light on spatial factors affecting the structure and the evolution of the Internet infrastructure in China. Using model (1) as the starting point, the effect of a set of explanatory variables, which reflect space and time dimensions of the digital infrastructure network, is tested in this section. Going a step further, the modeling results are juxtaposed with corresponding results from Europe. This comparison will increase the robustness of our analysis and enable potential comparisons.

More specifically, model (1) can be further expanded in the following log-log form:

$$\ln(IP_{ijt}) = a_0 lnk + a_1 \ln(distance_{ij}) + a_2 external_{ijt} + a_3 region_{ij} + a_4 periods_{ij} + \varepsilon_{ijt} \quad (6)$$

The dependent variable (IP_{ijt}) represents the connectivity between any two connected cities (*i* and *j*) in China in year *t*. The temporal dimension represents the five years study period (2007-2011). Building upon the results of the global analysis, we expect that physical distance (*distance*) between *i* and *j* will have a negative impact on the installed infrastructure between *i* and *j*. Then, a number of other structural explanatory variables are tested here. Based on the above discussion about the importance of international digital connections, it is expected that the IP connectivity between two cities (*i* and *j*) will be positively affected, if both *i* and *j* have international gateway roles. To test this effect we introduce here a dummy variable (*external*) which is equal to 1 when both *i* and *j* have international IP links during year *t*. Then, we test the impact of spatial structure and the importance of provinces in the formation of the Chinese digital infrastructure. Thus, another variable is introduced to test the impact of intra-province links: variable *region* is equal to 1 when both *i* and *j* are located in the same province in China. Finally, the effect of the stability of the connectivity over time is also tested here. Although IP networks are physical networks, re-wiring is possible within such networks in order to meet demand (Gorman and Kulkarni 2004). To test this attribute, the effect of variable *periods*, which indicates the number of years that a link between cities *i* and *j* was present during the study period, is tested.

In order to take advantage of the bi-dimensional data on digital connectivity (IP links between *i* and *j* at year *t*), panel data specifications are adopted for the estimation of (6). Panel data models improve the researchers' ability to control for missing or unobserved variables (Hsiao 2003). Such an omitted-variable bias as a result of unobserved heterogeneity is a common problem in cross-section models.

While panel data introduces considerable gains, there are also methodological limitations to be addressed. For instance, two are the main options for estimating panel regressions: *fixed effects* (FE) and *random effects* (RE) models (Wooldridge 2003). The RE model would have been the preferred choice here because the first differentiation process of the within estimator (FE) would have resulted in the elimination of the time-invariant variables (*distance*, *region* and *periods*) (e.g. Brun et al. 2005; Etzo 2011). However, the efficiency of the RE model goes hand in hand with other limitations: the

consistency of RE estimators depends on whether the unobserved random effects are uncorrelated with the regressors. For instance, some of the explanatory variables might be endogenous by being correlated with omitted variables which affect the installation of IP links between cities (Baier and Bergstrand 2001). If this is the case, an instrumentation of the endogenous variables would be necessary in order to obtain unbiased estimators. However, such instrumentation is not an easy task given the complexity of the Internet infrastructure and the lack of prior empirical research in this area. Therefore, a two-way fixed-effects estimation is introduced here (see also Tranos and Nijkamp 2013). This specification is differentiated by the usual FE because it addresses unobserved effects at two dimensions (Baltagi 1995). Thus, the error term ε_{ijt} from (6) can be analyzed as following: $\varepsilon_{ijt} = \omega_{it} + \zeta_{jt} + \upsilon_{ij}$. In this case, ω_{it} and ζ_{jt} are the *i* and *j* as well as time-specific effects and υ_{ij} the remainder stochastic disturbance term. Thus, the two-way FE will address potential *i* and *j* time-specific effects (i.e. the time variant city level effects which are not observed and are not of our interest here) and enable the estimation of the structural effects (*distance, external, region, periods*).

Columns 1-3 in Table 6 present the estimation of (6) using different effects. The results are consistent and some interesting conclusions can be drawn. First of all, a distance decays effect is present in the distribution of the Internet infrastructure across the Chinese urban system. The impact of distance increases with the use of specific effects, and especially when the two-way effects are introduced. This is important, as the latter specification appears to be the most robust due to the lack of unobserved effects. Then, the importance of the Chinese cities, which act as gateways with the rest of the world, in attracting intra-China connectivity is reflected in the variable *external*. Indeed, when both connected cities share IP links not only with cities in China, but also with cities abroad, then the installed digital infrastructure between this pair of cities is expected to be higher. The positive sign of this variable remains unchanged across different specifications. However, the effect of this variable stops being significant when the two-way fixed effects are introduced. This is not surprising as the two-way effect probably mask the impact of the links between gateway cities because the gateway roles vary both among time and space. Another important structural factor for the development of the Chinese IP network is regional connectivity as the location of two connected cities within the same province has a positive impact on installed connectivity between them. Finally, an indirect assessment of the above discussed *cumulative causation* process is achieved with the use of the variable *periods*. The consistent positive effect indicates that the number of years a pair of cities remains connected during the study period is positively related with the amount of installed infrastructure among these cities. In other words, early, and consequently lengthier, participation of a city-pair in the Chinese IP network has a positive impact on the connectivity between these two cities.

Table 6: Determinants of digital connectivity within Chinese urban system

Dep. Var. IP (ln)	**(1)**	**(2)**	**(3)**	**(4)**
IP (ln, 1 year lag)				0.373
				(0.19)*
distance (ln)	-0.056	-0.208	-0.216	-10.075

	(0.03)*	(0.05)***	(0.06)***	(4.83)**
External	1.026	0.241	0.228	-4.021
	(0.06)***	(0.08)***	(0.29)	(2.50)
region	0.305	0.794	0.851	-40.843
	(0.09)***	(0.14)***	(0.19)***	(20.48)**
Periods	0.454	0.501	0.656	2.286
	(0.02)***	(0.03)***	(0.05)***	(0.87)***
time effects	Yes	Yes		Yes
i,j effects		Yes		
ij effects				
it and *jt* effects			Yes	
Constant	-0.324	1.752	1.844	70.298
	(0.25)	(2.29)	(4.31)	(34.51)**
Hansen				2.15
Diff. Hansen				2.14
AR(2)				-1.16
R^2	0.23	0.60	0.71	
N	3,281	3,281	3,281	1,810

* *p*<0.1; ** *p*<0.05; *** *p*<0.01; Standard Error in parentheses

In columns (1)-(3) the results of OLS estimations with various effects are presented. Column (4) presents the two-step robust system GMM estimation. *Hansen* is the Hansen test for overidentification restriction; *Diff. Hansen* is the Difference-in-Hansen test of exogeneity for the validity of additional moment conditions; and *AR(2)* is the Arellano-Bond test for serial correlation.

A more direct estimation of the cumulative causation process or, in network terms, of the preferential attachment, can be made with the introduction of a dynamic framework. Therefore, model (6) is expanded to the following form:

$$\ln(IP_{ijt}) = a_0 lnk + \gamma_1 \ln(IP_{ijt-1}) + a_1 \ln(distance_{ij}) + a_2 external_{ijt} + a_3 region_{ij} + a_4 periods_{ij} + \varepsilon_{ijt} \quad (7)$$

The main difference with (6) is the inclusion of the autoregressive term (IP_{ijt-1}). This creates estimation complications, as OLS and conventional fixed and random effects estimators result in biased and inconsistent estimates because of the correlation between the autoregressive term and the error term. To overcome this, the generalized method of moments (GMM) technique is introduced here. The latter approach refers to Arellano and Bond's (1991) suggestion of using first differencing for eliminating individual effects and then using all possible lags of dependent and independent variables as instruments for the endogenous variables (in our case only the lags of the dependent variable). Later on, Arellano and Bover (1995) and Blundell and Bond (1998) suggested first differencing not on the regressors, but rather on the instruments, a choice which results in increased efficiency (Roodman 2006). The latter approach, known as *system GMM*, is used here.

Column (4) in Table 6 presents the estimation of model (7). The main finding is some weak evidence for the existence of a preferential attachment mechanism in the evolution of the Chinese inter-

city IP network as the lagged value of the IP connectivity between two cities has a positive impact on the installed IP infrastructure between these two cities. This effect is significant and marginally below the 0.95 threshold (p-value = 0.051). This rather weak evidence fits well with the previous result on the truncated power law distribution and the disturbed preferential attachment process. Although this cumulative causation phenomenon does not continue endlessly, it still plays a significant role in the evolution of the digital inter-city network in China.

Another important finding is that despite the inclusion of the autoregressive term, the distance decay effect is still present. The same applies to the indirect measure of the cumulative causation. The magnitude of these two effects is much higher than those in the static models. Nevertheless, the consistency in qualitative terms signifies their importance as structural elements of the Internet infrastructure across the Chinese cities. This is not the case for the effect of the variables *external* and *region*. While the former stops being significant, the latter has a negative and significant effect, which does not agree with the static models.

A crucial point in the above process is the validity of the instruments adopted for the GMM estimations. Three test have been performed here (Jiwattanakulpaisarn et al. 2009). Firstly, orthogonality conditions of the instruments are tested using the Hansen test for overidentifying restrictions. Then, the validity of additional moment conditions in levels is tested with the Difference-in-Hansen test of exogeneity. Finally, the Arellano-Bond test for serial correlation is reported, the null hypothesis of which (no second-order autocorrelation in differenced residuals) verifies the validity of two or more order lagged variables as instruments. All of the reported tests support the validity of the instruments used in the system GMM estimation.

In total, spatial forces affect the structure of the digital infrastructure network. Physical distance and localization effects are valid for the case of China. What is more, the dynamic panel analysis confirms the preferential attachment or, in other words, the cumulative process in the IP distribution. Finally, the effect for the links between gateway cities reflects the structure of the Chinese digital network as only 10 per cent of the digitally connected cities (110 cities in 2011) share links with cities outside China. Apparently, these cities act as the main hubs of the Chinese Internet and are responsible for the SF nature of this network. The positive concentration of IP links between such nodal cities reflect the importance of these in cities in holding the Chinese Internet together (see Doyle et al. 2005).

6. Conclusions

The combination of network analysis with econometrics in the present paper has resulted in various insightful outcomes regarding the understanding of the symbiotic relation between urban and digital networks, both globally and in China. In addition to persistent urbanization trends, the urbanized world will also be a connected world, in which next to physical infrastructure also the digital infrastructure (in particular, the Internet) will play a central role.

Despite the scale of the analysis, similar spatial forces – both globally and nationally – have been identified as drivers of the digital infrastructural network. Physical distance, accessibility and localization effects appear to be important factors behind digital connectivity. Thus, it is fair to say that the 'death of distance' discussion is not generally valid in the frame of the Internet infrastructure. Moreover, factors related to the knowledge economy and city size are also important factors behind the allocation of the Internet's physical layer.

In addition, it has also become apparent that the digital infrastructure reflects specific attributes of urban systems. While both Chinese and European urban IP networks resemble broad-scale networks with truncated power laws, the Chinese network is significantly more centralized than the European one, highlighting the polycentric nature of the latter. Moreover, it can be argued that the more centralized digital network, which underpins the Chinese Internet, reflects the centralized and regulative planning system in China. It is important to note that, despite a rapid increase in global bandwidth, the major factor contributing to the slowness of Internet traffic in China is caused by strict governmental arrangements. The Internet traffic to and from cities outside China is carried only through international gateways located in the three mega-cities of Beijing, Shanghai and Guangzhou, so that the Chinese government can better monitor and control the information on the Internet (Dai 2003). Without extending international gateways in China, the increased development of bandwidth of the Internet of interconnected networks and core routers on the Internet will not be fully utilized in China (Dai 2003).

The modelling experiment developed in the present paper call also for new applications in other emerging economies, such as India, Brazil or Indonesia. All such countries follow different development cycles and Internet access and use are governed by varying institutional regimes. Clearly, the sample of world cities used in our investigation could be extended. And it would certainly be relevant to test the robustness of our findings by examining other – and perhaps more extensive – data bases. A potentially useful area for further research is the utilization of data regarding digital flows instead of digital capacity. Online social media could be great candidates for data extraction. However, such data is difficult to be extracted for issues related to privacy and commercial business strategies. Nonetheless, it seems plausible to expect that the results of such an extended analysis would be largely similar to our present results.

Finally, global connectivity among cities presupposes a world without strict borders. Open access is an important condition for a globally linked world, and communication policies should do their utmost to create effective legal frameworks for ensuring open access conditions in a digital world. From this perspective, the urban century' and the digital age would have to run in tandem.

Appendix: Descriptive Statistics for GPCI

	N	Minimum	Maximum	Mean	Std. Dev.
Function-specific variables					
economy	34	19,60	58,30	36,8529	8,77987

RD	34	1,30	76,40	23,1882	15,95417
Cultural	34	4,30	60,60	21,2382	12,81666
Livability	34	32,70	60,70	43,7118	6,59342
environment	34	22,80	71,40	54,8324	10,15064
accessibility	34	18,80	57,90	33,5559	8,24622
Actor-specific variables					
manager	34	15,80	34,40	24,5853	5,06396
researcher	34	2,70	37,00	13,0676	7,58830
Artist	34	4,60	24,80	11,0765	3,76242
visitor	34	8,30	31,40	19,3559	5,06642
resident	34	13,20	36,20	24,4206	5,40834

References

ALSTOTT, J., BULLMORE, E., PLENZ, D. (2013) powerlaw: a Python package for analysis of heavy-tailed distributions. *http://arxiv.org/abs/1305.0215v1*,

AMARAL, L. A. N., SCALA, A., BARTHELEMY, M., STANLEY, H. E. (2000) Classes of small-world networks. *PNAS,* 97**:** 11149–11152

ARELLANO, M., BOND, S. R. (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies,* 58**:** 277–297

ARELLANO, M., BOVER, O. (1995) Another look at the instrumental variables estimation of error-components models. *Journal of Econometrics,* 68**:** 19-51

ARRIBAS-BEL, D., KOURTIT, K., NIJKAMP, P. (2013) Benchmarking of world cities through Self-Organizing Maps. *Cities,* 31**:** 248-257

BAIER, S. L., BERGSTRAND, J. H. (2001) The growth of world trade: Tariffs, transport costs, and income similarity. *Journal of International Economics,* 53**:** 1-27

BALTAGI, B. H. (1995) *Econometric analysis of panel data*.Chichester, John Wiley & Sons

BARABÁSI, A.-L. (2002) *Linked: The new science of networks*.Cambridge, MA, Perseus

BARABÁSI, A.-L., ALBERT, R. (1999) Emergence of Scaling in Random Networks. *SCIENCE,* 286**:** 509-512

BATTY, M. (2012) Building a science of cities. *Cities,* 29**:** S9-S16

BATTY, M., LONGLEY, P. (1994) *Fractal Cities: A Geometry of Form and Function*.New York, MIT Press

BLUM, B. S., GOLDFARB, A. (2006) Does the internet defy the law of gravity? *Journal of International Economics,* 70**:** 384–405

BLUNDELL, R., BOND, S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics,* 87

BONARICH, P. (2007) Some unique properties of eigenvector centrality. *Social Networks,* 29**:** 555–564

BRUN, J.-F., CARRERE, C., GUILLAUMONT, P., MELO, J. D. (2005) Has distance died? Evidence from a panel gravity model. *The World Bank Economic Review,* 19**:** 99-120

BUCHANAN, M. (2002) *Nexus: Small Worlds and the Groundbreaking Science of Networks*.New York, W. W. Norton

CAIRNCROSS, F. (2001) *The death of distance 2.0*.London, Texere Publishing Limited

CARMI, S., HAVLIN, S., KIRKPATRICK, S., SHAVITT, Y., SHIR, E. (2007) A model of Internet topology using k-shell decomposition. *PNAS,* 104**:** 11150-11154

CASTELLS, M. (1996) *The Rise of the Network Society*.Oxford, Blackwell

CHOI, J. H., BARNETT, G. A., CHON, B.-S. (2006) Comparing world city networks: a network analysis of Internet backbone and air transport intercity linkages. *Global Networks,* 6**:** 81–99

CLAUSET, A., SHALIZI, C. R., NEWMAN4, M. E. J. (2009) Power-law distributions in empirical data. *SIAM Review,* 51**:** 661-703

COHEN, G., NIJKAMP, P. (2006) e-Governance and Cyber Images of Cities. In: JOHANSSON, I. (ed) *Entrepreneurship and Development – Local Processes and Global Patterns.* University West, Uddevalla

COHEN, G., SALOMON, I., NIJKAMP, P. (2002) Information-communications technologies (ICT) and transport: does knowledge underpin policy? *Telecommunication Policy,* 26**:** 31-52

DAI, X. (2003) ICTs in China's development strategy. In: HUGHES, C. R., WACKER, G. (eds) *China and the Internet: politics of the digital leap forward.* Routledge, London

DÉMURGER, S. (2001) Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China? *Journal of Comparative Economics,* 29**:** 95-117

DERUDDER, B., TAYLOR, P. J., HOYLER, M., NI, P., LIU, X., ZHAO, M., SHEN, W., WITLOX, F. (2013) Measurement and interpretation of connectivity of Chinese cities in world city network, 2010. *Chinese Geographical Science,* 23**:** 261-273

DIMES (2010) DIMES Project. Tel Aviv

DING, L., HAYNES, K. E. (2006) The role of telecommunications infrastructure in regional economic growth in China. *Australasian Journal of Regional Studies,* 12**:** 281-302

DING, L., HAYNES, K. E., LIU, Y. (2008) Telecommunications infrastructure and regional income convergence in China: panel data approaches. *Annals of Regional Science,* 42**:** 843-861

DODGE, M., ZOOK, M. (2009) Internet-Based Measurement. In: EDITORS-IN-CHIEF: ROB, K., NIGEL, T. (eds) *International Encyclopedia of Human Geography.* Elsevier, Oxford

DOYLE, J. C., ALDERSON, D. L., LI, L., LOW, S., ROUGHAN, M., SHALUNOV, S., TANAKA, R., WILLINGER, W. (2005) The ''robust yet fragile'' nature of the Internet. *Proccedings of the Natioanal Academy of Science of the United States of America,* 102**:** 14497–502

ERDÖS, P., RÉNYI, A. (1959) On random graphs. *Publicationes Mathematicae Debrecen,* 6**:** 290–97

ETZO, I. (2011) The determinants of the recent interregional migration flows in Italy: a panel data analysis. *Journal of Regional Science,* forthcoming

FREEMAN, C. (1978-1979) Centrality in social networks conceptual clarification *Social Networks,* 1**:** 215-39

GORMAN, S. P., KULKARNI, R. (2004) Spatial small worlds: new geographic patterns for an information economy. *Environment and Planning B,* 31**:** 273-296

GORMAN, S. P., MALECKI, E. J. (2002) Fixed and fluid: stability and change in the geography of the Internet. *Telecommunications Policy,* 26**:** 389-413

HANNEMAN, R. A., RIDDLE, M. (2005) Introduction to social network methods. *Riverside, CA: University of California***:** Riverside (published in digital form at http://faculty.ucr.edu/~hanneman/)

HSIAO, C. (2003) *Analysis of panel data*.Cambridge, Cambridge University Press

INSTITUTE FOR URBAN STRATEGIES (2010) Global Power City Index. The Mori Memorial Foundation, Tokyo

JACOBS, J. (1961) *The Death and Life of Great American Cities*.New York, Vintage

JIWATTANAKULPAISARN, P., NOLAND, R. B., GRAHAM, D. J., POLAK, J. W. (2009) Highway infrastructure and state-level employment: A causal spatial analysis. *Papers in Regional Science,* 88**:** 133-159

KUSAKABE, K. (2012) *Gender, Roads and Mobility in Asia*.Bourton on Dunsmore, Practical Action Publishing

MALECKI, E. J. (2002) The economic geography of the Internet's infrastructure. *Economic Geography,* 78**:** 399-424

MAYER, T., ZIGNAGO, S. (2005) Market Access in Global and Regional Trade. *CEPII,* Working Paper 2

MEGA, V. (2010) Future Belongs to Cities. *Regional Symbiosis,* 18**:** 1-13

MOK, D., WELLMAN, B., CARRASCO, J. (2010) Does Distance Matter in the Age of the Internet? *Urban Studies,* 47**:** 2747-2783

MORICHI, S., RAJ ACHARYA, S. (2013) *Transport Developments in Asian Megacities*.Heidelberg, Springer

MULLIGAN, G., CRAMPTON, J. P. (2005) Population Growth in the World's Largest Cities. *Cities,* 22**:** 365-380

NEAL, Z. (2011) Differentiating Centrality and Power in the World City Network. *Urban Studies,* 48**:** 2733–2748

NEAL, Z. (2012) *The Connected City: How Networks are Shaping the Modern Metropolis*.New York, Routledge

NEWMAN, M. E. J. (2003) The structure and function of complex networks. *SIAM Review,* 45**:** 167-256

NEWMAN, M. E. J. (2008) Mathematics of networks. In: BLUME, L. E., DURLAUF, S. N. (eds) *The New Palgrave Dictionary of Economics.* 2nd ed. Palgrave Macmillan, Basingstoke

NIJKAMP, P. (2008) XXQ factors for sustainable urban development: a systems economic view. *The Journal of the Romanian Regional Science Association,* 2**:** 1-34

NIJKAMP, P., KOURTIT, K. (2012) The New Urban Europe: Global Challenges and Local Responses in the Urban Century. *European Planning Studies,* forthcoming

NOOY, W. D., MRVAR, A., BATAGELJ, V. (2005) *Exploratory Social Network Analysis with Pajek*. Cambridge University Press

ROODMAN, D. (2006) How to do xtabond2: an introduction to 'difference' and 'system' GMM in Stata. *Center for Global Development,* Working Paper No. 103

SHAVITT, Y., SHIR, E. (2005) DIMES: Let the Internet Measure Itself. *ACM SIGCOMM Computer Communication Review,* 35**:** 71-74

TASSEY, G. (1992) *Technology infrastructure and competitive position*.Boston, MA, Kluwer Academic Press

TASSEY, G. (2008) Modeling and measuring the economic roles of the technology infrastructure. *Economics of Innovation and New Technology,* 17**:** 615-29

TAYLOR, P. J. (2001) Specification of the world city network. *Geographical Analysis,* 33**:** 181-94

TAYLOR, P. J. (2004) *World city network*.London and New York, Routledge

TAYLOR, P. J., CATALANO, G., WALKER, D. R. F. (2002) Measurement of the World City Network. *Urban Studies,* 39**:** 2367-2376

TRANOS, E. (2011) The topology and the emerging urban geographies of the Internet backbone and aviation networks in Europe: a comparative study *Environment and Planning A,* 43**:** 378-392

TRANOS, E. (2013) *The Geography of the Internet: Cities, Regions and the Internet Infrastructure in Europe*.Cheltenham, UK, Edward Elgar

TRANOS, E., GILLESPIE, A. (2009) The spatial distribution of Internet backbone networks in Europe: a metropolitan knowledge economy perspective. *European Urban and Regional Studies,* 16**:** 423-37

TRANOS, E., GILLESPIE, A. (2011) The urban geography of Internet backbone networks in Europe: roles and relations. *Journal of Urban Technology,* 18**:** 35**–** 49

TRANOS, E., NIJKAMP, P. (2013) The death of distance revisited: cyber-place, physical and relational proximities. *Journal of Regional Science,* forthcoming

UN (2010) *World Urbanization Prospects*.New York, United Nations

WANG, Y., LAI, P., SUI, D. (2003) Mapping the Internet using GIS: The death of distance hypothesis revisited. *Journal of Geographical Systems,* 5**:** 381-405

WATTS, D. J. (2003) *Six Degrees: The Science of a Connected Age*.New York, W. W. Norton & Company

WATTS, D. J. (2004) The 'New' Science of Networks. *Annual Review of Sociology* 30**:** 243-270

WATTS, D. J., STROGATZ, S. H. (1998) Collective Dynamics of Small-World networks. *Nature,* 393**:** 440-442

WOOLDRIDGE, J. M. (2003) *Introductory econometrics: a modern approach*.Mason, Ohio, South-Western

ZHEN, F., WANG, B., XI, G., CHEN, Y. (2013) Research on China's Urban Network Based on the Relations between Micro-Blog Users: a Case Study of Sina Micro-Blog. In: SCHRENK, M., POPOVICH, V. V., ZEILE, P., ELISEI, P. (eds) *REAL CORP 2013.* Rome, Italy