Airline data for global city network research: reviewing and refining existing approaches

Ben Derudder*, Frank Witlox*, James Faulconbridge**, Jon Beaverstock***

*
Geography Department
Ghent University
E-mail: ben.derudder@ugent.be, frank.witlox@ugent.be

**
Geography Department
Lancaster University
E-mail: j.faulconbridge@lancaster.ac.uk

***
Geography Department
Nottingham University
E-mail: jonathan.beaverstock@nottingham.ac.uk

Abstract

Information on air passenger flows is potentially a prime data source for assessing spatial patterns in the global city network, but previous analyses have been hampered by inadequate and/or partial data. The ensuing analytical deficiencies have reduced the overall value of these analyses, and this paper examines how some of these deficiencies may be rectified. First, we review the rationale for using airline data to analyse the global city network. Second, we assess the data problems encountered in previous research. Third, we elaborate on the construction of datasets that may circumvent some of these problems. The proposed refinements include the omission of the hub function of major airports and ways to extract relevant business flows from the data.

Keywords: global city network, transnational mobility, business travel, airline data

I. Introduction

It has become commonplace to underline that recent developments in (the geographies of) transport and communication infrastructures have had a profound impact on the spatial organization of an increasingly globalized society (e.g. Black, 2003; Rodrigue et al., 2006; Dicken, 2007). One of the most commonly cited evolutions in this context is the alleged demise of the relevance of 'territoriality' in favour of 'networks', an evolution which leading sociologist Manuel Castells (1996, 2001) famously described as a transition from an international economy organized around 'spaces of places' to a global economy organized around 'spaces of flows'. Although there is a great deal of debate on the actual significance and implications of this shift, there can be little doubt that the spectacular growth of border-crossing mobility – for the largest part through air transport – is increasingly producing new spatial patterns of economic and social life. This had led some researchers to suggest that radical new ways of structuring our thinking about spatial patterns are required. Sheller and Urry (2006), for instance, attempt to capture these trends by devising a 'new mobilities paradigm', which is concerned with the patterning, timing, and causation of the face-to-face copresence so greatly facilitated by the contemporary surge in mobility. Another major strand of research focuses on potential spatial frameworks for capturing these trends, whereby a so-called 'global city network' (GCN) appears to be a likely candidate to replace the inter-state system for organizing our knowledge about the world (Taylor, 2004). Global cities are hereby essentially defined as key points in the organization of the global economy, and increasingly derive their functional importance from their mutual interactions rather than with their proper hinterlands[1]. In this paper, we will focus on one particular aspect of the interrelation between transnational mobility and this networked geography, i.e. the relevance of data on air passenger flows for revealing the material spatiality of this GCN.

The paper consists of three main parts. The first section presents a general introduction to the GCN literature, with a specific focus on the position of airline-based studies within this research domain. The relevance of air transport for GCN research may seem deceitfully obvious: air transport is all about connections between cities, while airline data are comparatively easy to obtain. Our intention here, however, is to provide a somewhat deeper understanding of the relevance of airline data by situating this information source within the GCN literature at large. The second section shows how previous airline-based studies have quasi-systematically been hampered by inadequate and/or partial data. The third section, then, presents some possible alternatives to the problem of inadequate data. It is not our intention to provide yet another empirical analysis of global city-formation based on 'better' data, but rather to provide a conceptual overview of how airline-based analyses of GCNs may collectively be improved in future research. To this end, we discuss some alternative data sources and propose some data manipulations that, taken together, may advance our understanding of the empirical association between air transport and GCNs. In a short conclusion, we briefly discuss the main implications of this paper and outline some avenues for further research.

[1] Some researchers explicitly differentiate between the terms 'global city' and 'world city' (e.g. Sassen, 2001), and in some cases such distinction is indeed no less than crucial (see Derudder, 2006). However, in the context of the present paper, these conceptual details are of lesser importance, and we will therefore consistently use the generic term 'global city' to address the literature at large.

II. The position of air transport-based studies within global city network research

II.1 GCN research: basic assumptions and main critiques

The contemporary GCN literature can be traced back to two interrelated papers by Friedmann and Wolff (1982) and Friedmann (1986). Both texts framed the rise of a global urban system in the context of a major geographical transformation of the capitalist world economy. This restructuring, most commonly referred to as the 'new international division of labour', was basically premised on the internationalization of production and the ensuing complexity in the organizational structures of multinational enterprises (MNEs). This increased economic-geographical complexity, Friedmann (1986) argued, requires a limited number of control points in order to function, and global cities were deemed to be such points. The publication of Saskia Sassen's (1991) *The Global City* in 1991 marked a shift of attention to global inter-city flows resulting from the critical servicing of worldwide production rather than to its formal command through corporate headquarters of MNEs. Sassen's approach focuses upon the attraction of advanced producer service firms (providing professional, financial and creative services for businesses) to major cities with their knowledge-rich environments and specialist markets. In the 1980s and 1990s many such service firms followed their global clients to become important MNEs in their own right. These advanced producer service firms thereupon created worldwide office networks covering major cities in most or all world regions, and it is exactly the myriad of interconnections between service complexes that gives, according to Sassen (1991, 2001), way to GCN formation.

A number of fundamental assumptions of GCN research have been criticized from different quarters. A main conceptual limitation of this literature has been the concentration on a relatively few large metropolitan centres to concomitant neglect of all other cities. The most trenchant critique along these lines is by Robinson (2002, p. 536), who complains that "millions of people and hundreds of cities are dropped off the map of much research in urban studies." This exclusion is from two 'maps': (i) the geographical map of world cities wherein most cities in the 'South' are missing; and (ii) the conceptual map of global cities which focuses on a narrow range of global economic processes so that myriad other connections between cities are missing. However, *all* cities experience contemporary global processes, and globalization can therefore not be construed as affecting just a few privileged cities. Subsequently Robinson (2005, p. 760) has conceded that the GCN literature now covers "a much wider range of cities around the globe" thus lessening the exclusion from the map. This attempt to broaden our understanding of the global city network has seen the postulation of such ideas as 'globalizing cities' (Marcuse and van Kempen, 2000) or 'cities in globalization' (Taylor et al., 2007).

Empirical GCN research, in turn, the topic on which the present paper focuses, has long remained underdeveloped because of the lack of appropriate data, a problem which Short et al. (1996) referred to as 'the dirty little secret of world cities research'. This empirical poverty can, for instance, clearly be read from Castells' (1996, p. 469) book, which is part of a trilogy that is above all an attempt to reformulate social studies for a global age in which "networks constitute the new social morphology of our societies." However, when it comes down to providing a basic cartography of this global network society, Castells' argument falls short of the conceptual shift he advances: the only actual evidence he comes up with in the chapter on the 'space of flows' consists of some limited inter-city information gathered from Federal Express. One can therefore only conclude, as Taylor (2004, p. 35) has recently done, that "the evidence [Castells] marshalls is mightily unimpressive." This gap between theoretical sophistication and evidential poverty is however not a lacuna specific to Castells' book: it has been a *structural* feature of research on the GCN, because data for assessing such urban networks are in general insufficient or even totally absent.

II.2. Solutions to the empirical problem

The basic reason for this problem of evidence is that standard data sources are ill-suited for GCN analyses (Taylor, 1997, 2004). To get an evidential handle on big issues, researchers normally rely on the statistics that are available, that is to say, already collected. But such collection is carried out usually by a state agency for the particular needs of government policy rather than for social science research. The result is that such data that are available have an attributional bias (measurements *of* administrative areas rather than *between* administrative areas) and are limited to national territories. Where official statistics extend beyond a state's boundaries they will still use countries as the basic units (e.g. trade data). Thus there is no official agency collecting data on, say, the myriad flows between London and New York. The major result has been that "few of the available data reveal anything about the flows and interdependencies" that are at the heart of this body of literature (Knox, 1998, p. 26), which leads Alderson and Beckfield (2004, p. 814) to note that in the past relatively few of the empirical GCN studies "utilized the sorts of relational data necessary for firmly establishing such rankings empirically."

These data problems have put researchers to work in recent years, and we have therefore witnessed a proliferation of empirical studies that explicitly seek to rectify this situation. Researchers have hereby relied on a wide variety of data, albeit that some information sources have come to dominate the empirical research as a whole (Derudder, 2006), i.e. (i) information on corporate organization (e.g. data on ownership links between firms across space) and (ii) information on infrastructure networks (e.g. data on the volume of air passenger flows across space). The success of both approaches can, of course, be traced back to their commonsensical appeal: the corporate organization approach acknowledges that well-connected cities derive their status in large part from the presence of key offices of important firms, while the infrastructure approach recognizes that well-connected cities are typified by the presence of vast enabling infrastructures. Put simply: the most important cities harbour the most important airports, while the extensive fiber backbone networks that support the Internet have equally been deployed within and between major cities, hence creating a vast planetary infrastructure network on which the global economy has come to depend almost as much as physical transport networks (Rutherford et al., 2004).

Table 1 summarizes the approaches developed in the empirical GCN literature through an overview of some key studies in this research domain. The table acknowledges that the basic bifurcation between corporate organization and infrastructure needs to be deepened on the basis of the exact types of firms and infrastructures, and equally shows that all this is in practice somewhat more complicated because of the presence of a limited number of studies that (i) make use of other types of data (e.g. Taylor's (2004b) analysis of non-governmental organizations) and/or (ii) combine indicators from both approaches (e.g. Beaverstock et al., 2000b). In the next section, we focus on empirical GCN studies that utilize data on international air transport flows to map GCNs.

Table 1 about here

III. Data issues in airline-based GCN studies

- Preamble

The starting point of airline-based GCN studies is the rather commonsensical observation that interactions between global cities are in large part facilitated and defined by transnational air transport flows. Following Keeling's (1995) initial contribution, there have been a large number of empirical researches that draw upon airline data to devise a mapping of the GCN (e.g. Cattan, 1995, 2004; Short et al., 1996; Kunzmann, 1998; Rimmer 1998; Shin and Timberlake, 2000; Smith and Timberlake, 2001, 2002; Matsumoto, 2004, 2007; Zook and Brunn, 2005). In principle, the most important advantage of this approach over researches carried out in the corporate organization approach is that airline statistics feature *tangible* inter-city relations. However, in hindsight, and in spite of the remarkable success of this type of research, it can be noted that most authors have simply asserted the relevance of publicly available airline data, although these are – in our view – downplayed by a number of structural problems (for earlier, but partial assessments, see Taylor, 1999; Beaverstock et al., 2000). In this section, we will provide a systematic overview of these data problems, which will in turn be used in the next section to show how this baleful situation may be rectified.

- The first problem: the lack of origin/destination data

In spite of profound differences between the most commonly employed air transport statistics – i.e. those provided by the Official Airline Guide (OAG), the International Air Transport Association (IATA) and the International Civil Aviation Organization (ICAO) –, these data sources are collectively hampered by at least three problems. A first major problem is the *lack of origin/destination information*. Standard airline statistics feature the individual legs of trips rather than the trip as a whole. Thus, in the case of a stopover, a significant number of 'real' inter-city links are replaced by two or more links that reflect corporate strategy rather than GCN relations. Furthermore, the lack of origin/destination information makes geographically detailed assessments of the GCN difficult, as direct connections become less likely as one goes down the urban hierarchy. And finally, the emergence of hub-and-spoke strategies in the airline sector have led to the rise of a set of urban areas that serve as interconnection points between different regions (Bowen, 2002; Derudder et al., 2007). As a consequence, there is a continuous shift towards a more polycentric organisation of airline passenger networks away from the traditional orientation on major cities. This implies that a number of secondary cities is rapidly gaining prominence in this new polycentric structure because of their role as transfer points rather than as origins and/or destinations in their own right, and the way in which airline statistics are commonly recorded tends to give disproportionate weight to such cities in GCN analyses.

An example of this first data problem can be observed in Keeling (1995), in which the mapping of the GCN is derived from an analysis of the dominant linkages in the global airline network. This map was created from a matrix of scheduled air services between 266 cities based on OAG data. However, this implies that only non-stop and scheduled direct flights between two cities were taken into account. As a consequence, the measures used by Keeling are not necessarily a reflection of actual inter-city relations. That is, such an analysis is likely to overstate the relational importance of cities that function as airline hubs, such as Amsterdam (KLM) and Frankfurt (Lufthansa), at the expense of cities such as Brussels and Berlin. Furthermore, direct links between, say, Brussels and Rio de Janeiro cannot be measured, as passengers are likely to go through São Paulo to make this trip.

- The second problem: state-centrism in data

The second obstacle to translating air transport statistics into GCN analyses arises because some of these data sources have incorporated *a subtle form of state-centrism*. Despite their global aspirations, most analyses are based on databases that contain information on international flows. This bias towards inter-state rather than trans-state flows tends to undervalue relations between cities that are situated in large and/or densely populated countries. Rimmer (1998, p. 460), for instance, has based his GCN analysis on data on 'international passengers'. This results in a downgrading of US world cities in particular because important connections such as Los Angeles–New York and Chicago–New York are not incorporated in this approach. As a consequence, Chicago only appears on one of Rimmer's maps as a 'fourth-level' link to Toronto, while Dublin appears on all maps because of its 'first-level' link with London. Of course, nobody would argue that Dublin is more important than Chicago as a global city; it only appears to be this way when one relies on international rather than transnational data. Another example can be found in Smith and Timberlake (2002, p. 123), who report a lack of information on the volume of air passenger traffic between Hong Kong and London. This admittedly important inter-city link did not feature in pre-1997 databases of the ICAO because flights between London and Hong Kong were considered to be 'national'. While the classification of the London–Hong Kong route and the downgrading of US cities are extreme examples, they clearly reveal how data on international passenger flows may hamper a GCN analysis[2].

- The third problem: general flow patterns

The third obstacle to the straightforward use of standard air transport statistics arises from the fact that such data source feature *general flow patterns*. Since airline statistics are unable to differentiate between specific flows within air passenger transport (i.e. the purpose of a passenger's travels), it is doubtful that air transport flows that are relevant in the context of the GCN can straightforwardly be deduced from such general data. A key example is the inclusion of major tourist destinations in previous analyses. For instance, in his mapping of the European urban system based on air passenger flows, Kunzmann (1998, p. 49) lists 14 airports that are secondary to the big three (London, Paris, and Frankfurt), including Munich, Milan, Madrid, and Palma de Mallorca. However, the high ranking of the latter merely reflects its role as one of the most popular holiday destinations in Europe; nobody would argue that it is a major global city (however conceived). While it is likely that most researchers would agree that destinations such as Palma de Mallorca should be omitted from the analysis, such data manipulation becomes increasingly difficult when non-global city processes intersect with global city-formation

[2] Smith and Timberlake (2002) were able to overcome the London–Hong Kong problem by estimating the importance of this link. The relegation of US cities was dealt with by the use of an additional data source that contained information on major routes in the US (namely, data provided by the Air Transport Association in Washington DC). While circumventing the most obvious gaps in the initial database has produced one of the most refined databases used to date, in general, this problem continues to affect the major Canadian, Chinese, and Brazilian cities (among others).

(however conceived). The rising importance of Miami in airline networks, for instance, can in part be traced back to its rise as control centre vis-à-vis the Carribean (Grosfoguel, 1995; Nijman, 1996; Brown et al., 2002). However, it is obvious that in the main it has been Miami's function as a retirement centre and major holiday destination that has fuelled its increasing connectivity in worldwide air transport connections. Since none of the commonly employed airline statistics are able to distinguish between tourist and business flows, there have been no clear procedures for estimating the amount of GCN-related traffic in overall air travel.

Having said this, it should be stressed that the 'problem' of the inclusion of tourism flows is not as straightforward as suggested here. For instance, it is clear that cities such as New York, London, Los Angeles, and Tokyo have become major tourist attractions in their own right. However, in his initial formulation of 'The World City Hypothesis', John Friedmann (1986, p. 74) maintained that the major driving forces behind world city formation were found in a limited number of rapidly expanding sectors. Although Friedmann identified world cities as major tourist destinations, it seems that tourism is merely an ancillary function, since "[m]ajor importance attaches to corporate headquarters, international finance, global transport and communications; and high level business services, such as advertising, accounting, insurance, and legal servcices. An ancillary function of world cities is ideological penetration and control. New York and Los Angeles, London and Paris, and to a lesser degree Tokyo are centres for the production and dissemination of information, news, entertainment and other cultural artefacts." In other words: it is clear that major world cities are major tourist destinations, but this is a secondary function at best in conceptual terms.

IV. Towards some solutions

In the previous section, we have argued that the key advantage of air transport-based studies (i.e. their focus on tangible inter-city flows) is downplayed because standard air transport statistics are collectively plagued by three deficiencies, i.e. (i) the lack of origin-destination data, (ii) the state-centric nature of the most important airline statistics, and (iii) the intersection with non-GCN processes. In this section, we show how future research may overcome these problems by making use of alternative datasets, combined with some specific manipulations of these data. We will show how the first two problems can be tackled by using a so-called Marketing Information Data Transfer (MIDT) database, while the third problem can be addressed – at least for Europe – by drawing on a dataset of the Association of European Airlines (AEA) which distinguishes between different travel classes.

IV.1 MIDT data

Our first alternative data source comprises a unique data set (for social science research) that provides information on individual passenger flows in 2001. This MIDT database is described in detail in Derudder and Witlox (2005) and Derudder et al. (2007), and reference should be made to these publications for further details. Here we produce a summary so that our line of reasoning can be followed. The MIDT database contains information on bookings made through so-called Global Distribution Systems (GDS) such as Galileo, Sabre, Worldspan, Topas, Infini, Abaccus and Amadeus (Shepherd Business Intelligence, 2005). GDS are electronic platforms used by travel agencies to manage airline bookings (i.e. the selling of seats on flights offered by different airlines), hotel reservations, and car rentals. Using a GDS-based database therefore implies that bookings made directly with an airline are often excluded from the system and therefore the data. However, in 1999, just two years prior to our data, 80% of all reservations continued to be made through GDS (Miller, 1999). Thus, although our information source may give a slightly biased picture of airline connections, there is no reason to assume that the overall pattern of reservations made by direct bookings differs fundamentally from that for reservations made through a GDS.

Using this MIDT database instead of standard data sources has two advantages in the context of GCN research. First, as the MIDT-database contains real origin/destination information, the overrating of the connectivity of airline hubs and first-tier global cities is minimized, which allows assessing flows between cities in the lower rungs of the GCN in more detail. Second, the MIDT-based database does not distinguish between national and international flows, and can therefore be used to construct a truly transnational inter-city matrix. The New York–Chicago link is appropriately treated in the same way as the New York–Toronto link, which further reduces the underestimation of second-tier cities in large and/or densely populated countries.

Through our cooperation with an airline, we were able to obtain a MIDT database that covers the period January-August 2001, and contains information on more than 500 million passenger movements. This database was used to construct an inter-city matrix detailing the total volume of passenger flows between cities. To achieve this, we first relabelled the airport codes as city codes. This was necessary to compute meaningful inter-city measures, because a number of cities have more than one major airport[3]. The particular airport used by a passenger is not important in this context because, for the measurement of the London-New York relation, it is

[3] It is, however, very difficult to devise a precise criterion for grouping different airports into a single 'city'. Newark's airport for instance, clearly serves as a major gateway for New York City, but things are not always clear-cut. In our research, we focused on the morphological characteristics of metropolitan areas and the way in which airports market themselves.

irrelevant whether one flies from Heathrow to JFK or from Gatwick to Newark. Having summed the directional information into a single measurement detailing the total volume of passengers, we created an inter-city matrix that focuses on the most important cities in the global economy. Our selection of cities is based on the global city list compiled by the Globalization and World Cities (GaWC) group and network. GaWC's selection of cities was somewhat arbitrary in that it was loosely based upon previous GaWC experience in researching global office networks of producer service firms. Capital cities of all but the smallest states were included plus many other important cities in larger states. A total of 315 cities were selected, the full list is available at http://www.lboro.ac.uk/gawc/datasets/da11_2.html. In our data, 9 of the initial 315 cities were excluded because they had no airport (e.g. Bonn and Kawasaki) or because the airport was not serviced in the period under consideration because of political instability (e.g. Kabul). This reconfiguration produced a 306 × 306 matrix that quantifies the passenger flows between important cities in the global economy. Table 2 and Figure 1 present an overview of the most important cities and their connections. Table 2 features the 20 most important inter-city relations in the dataset. Figure 1 depicts the connections between the 30 most important cities in terms of total passenger flows. The size of the nodes varies with the total number of incoming or outgoing passengers; the size of the edges varies with the number of passengers flying between two cities. For reasons of clarity, only the most important links are shown.

Table 2 about here
Figure 1 about here

With the problems discussed in the previous section used as a checklist, the MIDT-based matrix may be able to overcome two of the obstacles. To illustrate this, Table 3 presents a comparison of the 20 most connected cities in the airline network as identified in Smith and Timberlake (2001, 2002) and in the MIDT-database. It is important to stress that this comparison is only made for illustrative purposes: the list of 'most important cities' in Smith and Timberlake (2001, 2002) is constructed on the basis of a fully fledged centrality analysis, while the MIDT-list merely reflects a ranking based on the total number of passengers boarding on/off. Furthermore, the data in Smith and Timberlake (2001, 2002) refer to passenger flows in the year 1997, while the MIDT data feature statistics for the year 2001. As a consequence, the reader should bear in mind that we merely contrast both lists to point to some data-induced differences. First, because the MIDT-database contains origin/destination information, the overrating of the connectivity of airline hubs and first-tier world cities is minimized, which allows assessing the relational patterns in the lower rungs of the WCN in more detail (e.g. the downsizing of the importance of hub cities such as Amsterdam and Frankfurt in Table 3). Second, the MIDT-based database does not distinguish between national and international flows, and can therefore be thought of as a transnational dataset. The New York–Chicago link is appropriately treated in the same way as the New York–Toronto link, which further reduces the underestimation of second-tier cities in large and/or significant nation-states (e.g. there are 9 US cities in the top-20 for the MIDT-data, while only 5 US cities feature in Smith and Timberlake (2001, 2002), mainly at the expense of Southeast Asian cities).

Table 3 about here

IV.2 AEA data

The main problem with the MIDT dataset is that it remains largely impossible to discern GCN-related flows from other flows. The importance of the New York–Miami route and particularly the New York–Las Vegas route in an overview of the most important North American inter-city links, for instance, suggests the importance of non-GCN links in the data (Table 4). Flows related to obvious holiday destinations such as Palma de Mallorca and Cancun can easily be deleted from the database, but this manipulation only works for airports that are obviously *not* related to global cities[4].

Table 4 about here

To counter this problem, one may draw upon data such as that provided by the Association of European Airlines (AEA) to unveil the basic spatiality of the GCN (albeit in this particular case only within Europe). The AEA is a non-profit-making organisation that brings together 31 major European airlines. The organisation 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. The AEA gathers travel data among its member airlines, and brings this information together in a database that allows assessing the geography of airline networks serviced by major European airlines on a monthly basis. For each connection, the database features information on carrier, origin and destination (airport, city, country and region), and the total volume of passengers, freight and mail. The AEA database seems very suitable for research on GCNs because the information on passenger volumes distinguishes between economy class and business class bookings. As a consequence, the data allow for an exclusive focus on the geography of business class travel (which will very likely be a better measure of GCN connectivity) and an assessment of the way in which this geography differs from that produced by 'ordinary' air passenger flows[5].

Through the cooperation of one the member airlines of the AEA, we obtained this dataset for the period January 2002 – December 2005. The database includes flights within Europe as well as flights between Europe and other regions. For our purposes, only flights where both origin and destination are European cities were retained. Furthermore, since our interest is primarily in flows between cities, we converted the airport-to-airport-by-carrier database into a squared city-to-city database by aggregating, for each booking class, the number of passengers for all the airports of a given city, and for all the carriers of a given city-pair. The end result, then, is a non-directional connectivity matrix for each type of booking class[6].

The most important problem with this and other similar datasets is the disparities in business class bookings because of different strategies pursued by 'national carriers'. This bias relates to the fact that most airports are (still) dominated by one or two carriers, which may or may not have a specific approach towards business class travel. Some carriers, for instance, have recently chosen to remove business class from some of their short/medium haul routes (e.g. KLM and Brussels Airlines). On these routes, the seats are the same for all passengers, only the flexibility of the ticket and the food and beverage service differs. On the other side of the spectrum, some carriers have a relatively important business class component because they serve specific

---

[4] Another, more generic problem is that airline data cannot avoid undervaluing a second-tier city that is close to a major world city. For example, a passenger travelling from Rotterdam to New York is likely to depart from Amsterdam because of (i) the short distance between Rotterdam and Amsterdam (less than 50 miles) and (ii) the importance of Amsterdam's Schiphol airport.

[5] The AEA dataset equally includes information on first class passengers. However, the absolute volume of these flows is very small for intra-European flights, and furthermore restricted to a selected number of carriers. We have therefore chosen to add these flows to the business class category.

[6] Once again, a further data manipulation involved the removal of obvious holiday centres such as Palma de Mallorca from the data through the application of GaWC's global city list.

markets (e.g. carriers flying to London City Airport) and/or because of historic reasons. The latter is the case for travel to/from Scandinavian cities. SAS, the dominant regional carrier in Scandinavia, was the first European airline to introduce business class, and this is still reflected in a high proportion of business class seats on their flights. The net effect of this bias is that business class bookings for, say, Copenhagen and Stockholm will be somewhat overvalued when compared to, say, Brussels and Amsterdam, while the connectivity of cities that are located near other important business centres will be somewhat undervalued because business travellers can choose between different means of travelling.

One further aspect of this dataset requires explicit interrogation, i.e. the crucial issue whether business class bookings actually capture the spatiality of business travel (and therefore GCN flows). After all, business travellers do not necessarily travel in business class, while (ostensibly rich) tourists may well travel in business class because of enhanced comfort. Thus the importance of, say, New York's GCN flows may well be overestimated because of rich leisure travellers, while business flows to some short-haul destinations may well be underestimated because business travellers may opt not to travel in business class because of the short travel time. The crucial question, therefore, is whether measures of *business class travel* provide us with satisfactory proxies for assessments of *business-related travel*? At one level distortions are clearly present, but the crucial question here is whether the ensuing biases are so strong that they totally undermine an analysis of business travel on the basis of business class bookings.

To address this issue, we discuss two basic features of the AEA-dataset for the year 2005, which jointly suggest that data on business class travel does indeed allow an actual assessment of business travel. First, there seems to be a straightforward difference in seasonal intensity for both types of booking classes. Figure 2 gives an overview of the monthly fluctuations in air travel in 2005 for the entire AEA-database for both booking classes. The monthly variations in connectivity are gauged through z-scores so that inter-booking class comparisons are possible in spite of different passenger volumes[7]. The seasonality of air travel is obviously different for economy and business class bookings. The economy class curve increases from January to July/August, and then decreases again towards the end of the year. The business class curve, in contrast, reaches its lowest levels in major holiday periods (July/August and December/January). The major point here is that the contrasting curves in Figure 2 suggest that, in general, air travel in business class captures business travel.

Figure 2 about here

Second, the relative proportion of business class travel is higher for clear-cut 'business cities' such as Geneva and Düsseldorf. Table 5 contains two rankings of European cities according to their connectivity in the European airline network. The first ranking focuses on the *absolute* importance of business class travel, the second ranking focuses on the *relative* proportion of business class travel within a city's overall passenger volume. When taking on board that (i) the proportion of business class travellers to/from Scandinavian cities is higher because of the historical legacy of SAS's corporate strategies, and that (ii) the proportion of business class

[7] In statistics, the z-score or standard score is a dimensionless quantity derived by subtracting the population mean from an individual raw score and then dividing the difference by the population standard deviation. The resulting z-score indicates how many standard deviations an observation is above or below the mean. It allows comparison of observations from different normal distributions, in this case between economy class and business class flows.

travel to/from cities such as London, Paris, Frankfurt and Amsterdam is somewhat relegated because of their function as gateways for rerouting international air travel (Derudder et al., 2007), it becomes clear that business centres do in effect have a higher proportion of business class travel. Once again, this seems to validate our claim that business class travel does indeed measure business travel.

Table 5 about here

To conclude this section, we extend the basic rankings in table 5 by focusing on one further aspect of the geography of business travel. First, rather than restricting the discussion to the absolute and relative dimensions of business travel on a city-by-city basis, we can assess the actual spatiality of business flows between cities. To this end, Figure 3 depicts the most important business travel links in 2005 between the 25 most important European cities in terms of total volume of business passengers. Once again, the size of the nodes varies with the total number of incoming or outgoing passengers, while the size of the edges varies with the number of business passengers flying between two cities. For reasons of clarity, only the most important links are shown (> 50,000 passengers). In addition to a cohesive business network centred on Stockholm, Oslo and Copenhagen, it is clear that business travel to/from Frankfurt, London, Paris and Amsterdam is dominant. These cities are highly interrelated, while most business travel from/to other major cities is also primarily orientated towards these cities (e.g., each city has well-connected business class flows to London. Table 6, in turn, looks at the spatiality of the relative importance of business travel to/from the most important business centres in 2005 (London, Paris, Frankfurt and Amsterdam). The table summarizes the results of a least squares regression on the logarithms of the volume of economy and business class passengers to/from each of these cities[8], and lists all cities with a standardized residual with an absolute value larger than 1: large negative residuals indicate that a city has less business class travellers than expected on the basis of the number of economy class passengers, positive residuals point to relatively strong business class connections to London, Paris, Frankfurt, and Amsterdam. Overall, the table reveals that cities with large positive residuals are primarily business centres (e.g. Frankfurt and Zurich), while cities with negative residuals are cities that also have an important tourism component to them (e.g. Rome and Barcelona)[9].

Table 1 about here
Figure 3 about here

[8] The logarithms were used to tackle the heteroscedasticity in the data.

[9] Obviously, the latter cities are also business centres in that they are also highly connected in business class flows (see Figure 3), but the point here is that their business class component is coupled with major tourism flows, which reinforces our suggestion that data on business class travel can actually be used to assess the geography of business travel.

V. Conclusion

The gist of our argument has been that although statistics of international air transport flows are potentially a prime data source in empirical studies of GCNs, previous such studies have not been able to live up to their potential. This is because the conventional airline statistics provided by IATA, OAG, ICAO,… result in less-than-perfect inter-city matrices for further analysis. We identified three systematic data problems in this context, and used this as a starting point for exploring alternative datasets. We discussed two such alternatives (MIDT and AEA), which jointly open up possibilities for future research along these lines. However, this paper merely engaged in a straightforward overview of some of the major patterns in the data. In the short run, this leaves us with three major avenues for further research.

First, there is the question of how both data sources may inform one another. One obvious problem is that both datasets offer partial solutions (the MIDT data featuring general flow data, and the AEA data featuring the individual legs of a trip): it is only through their collective application that we can take full advantage of their potential. Leaving aside the obvious problem that the AEA dataset only contains information on European carriers (similar 'regional' datasets will be necessary for a global analysis), this will take the form of a normalization of MIDT data based on the proportion of business travel for particular cities/connections in the AEA data. Second, there is the actual analysis of the data. One possibility is to analyse the data along the lines suggested by Shin and Timberlake (2001). This involves a two-step analysis, i.e. (i) an overarching analysis of the overall connectivity of key cities and (ii) a detailed geographical dissection of this connectivity through the application of standard network analytical tools, such as clique and block analyses. Third, and this is perhaps the most important point, there is the obvious need for conceptually grounding this kind of empirical research. For one thing, our airline data review has shown that such information requires a great deal of assumptions to be used in the interpretation of the geography and functionality of business travel flows. Consequently, whilst data providing a general overview of the major patterns and trends in business travel is undoubtedly helpful, it is clear that a more nuanced and sophisticated understanding emerging from more qualitative approaches to the study of inter-city business travel is necessary (Faulconbridge et al., 2008).

VI. Acknowledgements

We would like to thank two anonymous reviewers for their helpful and constructive comments on an earlier version of this paper.

VII. References

Alderson, A.S. & Beckfield, J. (2004). Power and position in the world city system. *American Journal of Sociology, 109*, 811-851

Alderson, A.S. & Beckfield, J. (2007). Globalization and the world city system. Preliminary results from a longitudinal data set. (In P. J. Taylor, B. Derudder, P. Saey & F. Witlox (Eds.), *Cities in Globalization: Practices, Policies, Theories* (pp. 21-36). London: Routledge.)

Amiel, M., Melançon, G. & Rozenblat, C. (2005). Réseaux multi-niveaux: l'exemple des échanges aériens mondiaux de passagers. *M@ppemonde, 79*. Retrieved on 12/02/2007 from http://mappemonde.mgm.fr/num7/articles/art05302.html

Beaverstock, J.V., Smith, R.G. & Taylor, P.J. (2000a). World city network: a new metageography? *Annals of the Association of American Geographers, 90*, 123-134

Beaverstock, J.V., Smith, R.G., Taylor, P.J., Walker, D.R.F. & Lorimer, H. (2000b). Globalization and world cities: some measurement methodologies. *Applied Geography, 20*, 43-63

Black, W.R. (2003). *Transportation: A Geographical Analysis*. (New York: Guilford Press)

Bowen, J. (2002). Network change, deregulation, and access in the global airline industry. *Economic Geography, 78*, 425-440

Brown, E., Catalano, G. & Taylor, P.J. (2002). Beyond world cities. Central America in a global space of flows. *Area, 34*, 139-148

Castells, M. (1996). *The Rise of the Network Society. The information Age: Economy, Society and Culture, Volume 1*. (Oxford: Blackwell)

Castells, M. (2001). *The Rise of the Network Society. The information Age: Economy, Society and Culture, Volume 1 (2nd Edition)*. (Oxford: Blackwell)

Cattan, N. (1995). Attractivity and internationalisation of major European cities: the example of air traffic. *Urban Studies, 32*, 303–312

Cattan N. (2004). Le monde au prisme des réseaux aériens. *Flux, 58*, 32-43

Derudder, B. (2006). On conceptual confusion in empirical analyses of a transnational urban network. *Urban Studies*, *43*, 2027-2046

Derudder, B., Taylor, P.J., Witlox, F. & Catalano, G. (2003). Hierarchical tendencies and regional patterns in the world city network: A global urban analysis of 234 cities. *Regional Studies, 37*, 875-886

Derudder, B. & Taylor, P.J. (2005). The cliquishness of world cities. *Global Networks, 5*, 71-91

Derudder, B., & 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*, 2371-2388

Derudder, B., Devriendt, L. & Witlox, F. (2007). Flying where you don't want to go: an empirical analysis of hubs in the global airline network. *Tijdschrift voor Economische en Sociale Geografie, 98*, 307-324

Dicken, P. (2007). *Global Shift: Mapping and Changing the Contours of the World Economy* (5th Edition). (New York: Guilford)

Ducruet, C. & Lee, S.W. (2006). Frontline soldiers of globalisation: port-city evolution and regional competition. *Geojournal, 67*, 107-122

Dupuy, G. (2004). Internet: une approche géographique à l'échelle mondiale. *Flux, 58*, 5-19

Faulconbridge, J.R., Beaverstock, J.V., Derudder, B. & Witlox, F. (2008) Corporate ecologies of international business travel: examples from professional service firms. Manuscript under review, available from authors.

Friedmann, J. (1986). The world city hypothesis. *Development and Change, 17*, 69–83

Friedmann, J. & Wolff, G. (1982). World city formation: an agenda for research and action. *International Journal of Urban and Regional Research, 6*, 309–344

Grosfoguel, R. (1995). Global logics in the Caribbean city system: the case of Miami. (In P.L. Knox & P.J. Taylor (Eds.), *World Cities in a World-System* (pp. 156-170). Cambridge: Cambridge University Press.)

Keeling, D.J. (1995). Transport and the world city paradigm. (In P.L. Knox & P.J. Taylor (Eds.), *World Cities in a World-System* (pp. 115-131). Cambridge: Cambridge University Press.)

Keeling, D.J. (2007). Transportation geography: new directions on well-worn trails. *Progress in Human Geography, 31*, 217–225

Knox, P.L. (1998). Globalization and world city formation. (In S.G.E. Gravesteijn (Ed.), *Timing Global Cities* (pp.21-31). Utrecht: Netherlands Geographical Studies.)

Kunzmann, K.R. (1998). World city regions in Europe: structural change and future challenges. (In F.C. Lo & Y.M. Yeung (Eds.), *Globalization and the World of Large Cities* (pp.37-75). Tokyo: United Nations University Press.).

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

Marcuse, P. & Van Kempen, R. (Eds.) (2000). *Globalizing Cities: a New Spatial Order?* (Oxford: Blackwell)

Matsumoto, H. (2004). International urban systems and air passenger and cargo flows: some calculations. *Journal of Air Transport Management, 10*, 241–249

Matsumoto, H. (2007). International air network structures and air traffic density of world cities. *Transportation Research Part E: Logistics and Transportation Review, 43*, 269-282

Nijman, J. (1997). Globalization to a Latin beat: the Miami growth machine. *Annals of the American Academy of Political and Social Sciences, 551*, 163-176

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

Rimmer, P.J. (1998). Transport and telecommunications among world cities. (In F.C. Lo & Y.M. Yeung (Eds.), *Globalization and the World of Large Cities* (pp. 433-470). Tokyo: United Nations University Press.)

Robinson, J. (2002). Global and world cities: a view from off the map. *International Journal of Urban and Regional Research, 26*, 531-554

Robinson, J. (2005). Urban geography: world cities, or a world of cities. *Progress in Human Geography, 29*, 757-765

Rodrigue, J.P., Comtois, C. & Slack, B. (2006). *The Geography of Transport Systems*. (New York: Routledge)

Rozenblat, C. & Pumain, D. (2007). Firm linkages, innovation and the evolution of urban systems. (In P.J. Taylor, B. Derudder, P. Saey & F. Witlox (Eds.), *Cities in Globalization: Practices, Policies, Theories* (pp. 130-156). London: Routledge.)

Rutherford, J., Gillespie, A. & Richardson, R. (2004). The territoriality of pan-European telecommunications backbone networks. *Journal of Urban Technology, 11 (3)*, 1-34

Sassen, S. (1991). *The Global City: New York, London, Tokyo*. (Princeton: Princeton University Press)

Sassen, S. (2001). *The Global City: New York, London, Tokyo* (2nd Edition). (Princeton: Princeton University Press)

Sheller, M. & Urry, J. (2006). The new mobilities paradigm. *Environment and Planning A*, *38*, 207–226

Short, J.R., Kim, Y., Kuss, M. & Wells, H. (1996). The dirty little secret of world city research. *International Journal of Regional and Urban Research, 20*, 697–717

Shin, K.H. & Timberlake, M. (2000). World cities in Asia: cliques, centrality and connectedness. *Urban Studies, 37*, 2257–2285

Smith, D.A. & Timberlake, M. (2001). World city networks and hierarchies 1979–1999: An empirical analysis of global air travel links. *American Behavioral Scientist, 44*, 1656–1677

Smith, D.A & Timberlake, M. (2002). Hierarchies of dominance among world cities: a network approach. (In S. Sassen (Ed.), *Global Networks, Linked Cities* (pp. 117-141). London: Routledge.)

Taylor, P.J. (2001). Specification of the world city network. *Geographical Analysis, 33*, 181-194

Taylor, P.J. (2004a). *World City Network: A Global Urban Analysis*. (London: Routledge)

Taylor, P.J. (2004b). New political geographies: Global civil society and global governance through world city networks. *Political Geography, 24*, 703-730

Taylor, P.J., Catalano, G. & Walker, D.R.F. (2002a). Measurement of the world city network. *Urban Studies, 39*, 2367-2376

Taylor, P.J., Catalano, G. & Walker, D.R.F. (2002b). Exploratory analysis of the world city network. *Urban Studies, 39*, 2377-2394

Taylor, P.J., Derudder, B., Saey, P. & Witlox, F. (Eds.) (2007). *Cities in Globalization: Practices, Policies, Theories*. (London: Routledge)

Townsend, A.M. (2001a). Network cities and the global structure of the Internet. *American Behavioral Scientist, 44*, 1697-1716

Townsend, A.M. (2001b). The Internet and the rise of the new network cities (1969-1999). *Environment and Planning B, 28*, 39-58

Zook, M. & Brunn, S. (2005). Hierarchies, regions and legacies: European cities and global commercial passenger air travel. *Journal of Contemporary European Studies, 13*, 203-220