(Pre-print final draft. The paper was published in: Information Economics and Policy, Volume 13, Issue 3, September 2001, pp. 339-349).

STATISTICS FOR THE INFORMATION AGE

Hans-Jürgen Engelbrecht*

Department of Applied and International Economics, Massey University,
Palmerston North, New Zealand

ABSTRACT

The rapid changes in the telecommunications sector shorten the useful lifespan of many of the established telecommunications traffic indicators, creating the need for constant re-assessment and improvement of existing indicators and the development of new ones. This paper first reviews two prominent publications by TeleGeography, Inc. that provide telecommunications traffic and Internet statistics. It then lists other data sources and discusses several prominent measurement efforts that highlight the need for internationally agreed upon measurement standards. Finally, it is pointed out that many of the economic and social statistics that are needed to help interpret the impact of the information revolution, as well as the techniques used to analyse them, are also in dire need of improvement.

* Tel.: +64 6 350 5799 ext. 5968; fax: +64 6 350 5660; *E-mail address:*

H.Engelbrecht@massey.ac.nz.

The author thanks TeleGeography Inc. for providing copies of TeleGeography 2000 and 2001, and Hubs and Spokes. All web addresses cited below were last checked on 21 February 2001.

1. Introduction

In order to conduct research into economic and policy aspects of information and (tele)communications, researchers have to have data that can be used in serious econometric and other analyses of these issues. It is a feature of the 'information age' that the more information there is, and the more knowledge is created and accumulation, the more difficult it is to measure what is going on. The current phase in the development of appropriate indicators is one of experimentation. There are plenty of new indices that try to measure some aspects of the emerging trends. They are often devised for particular purposes, but are not necessarily comparable or meaningful. It remains to be seen how many of these will stand the test of time.

This article reviews two publications by TeleGeography Inc., a Washington D.C. based research and publishing company and an independent subsidiary of Band-X Ltd., that provides what is probably the most authoritative compilation of international telecommunications traffic flows and commentary. Its annual flagship publication, TeleGeography, has been published since 1989. It focuses mostly on the international geography of telecommunications and is an indispensable reference work for anyone analysing or having to make strategic decisions in the quickly changing telecommunications sector. Given its price, it seems primarily aimed at a commercial, i.e. non-academic audience.¹ However, it should also be of great interest to other researchers.

The following section discusses TeleGeography2001 (TeleGeography, 2000a), as well as its recent offshoot, Hubs and Spokes (TeleGeography, 2000b), which focuses on the Internet. Section 3 lists some other data sources and briefly discusses several other prominent measurement efforts that highlight the need for internationally agreed upon measurement standards. Section 4 emphasises that telecommunications indicators need to be combined with a wide variety of other statistics in order to interpret their meaning and significance. However, many of these 'conventional' statistics are also

¹ At the time of writing the printed edition of TeleGeography 2001 sold for US \$ 1295, the electronic edition (single user licence) for US \$ 2495, and the combined printed and single user licence for US \$ 2990. The publication of both TeleGeography and Hubs and Spokes was partly supported by grants from telecommunications or related corporation. It is stated that these grants were given without preconditions.

in dire need of improvement. Finally, it is not just the data that need to be improved, but also the techniques used to analyse them.

2. The Geography of Telecommunications and the Internet

2.1 TeleGeography

Using text, tables, figures, charts and maps, TeleGeography provides an annual overview and commentary (in the form of topical essays or primers) on recent and likely future developments in the telecommunications sector. Route-by-route international telecommunication traffic flow data (in minutes) are provided for many countries in the world. The list below indicates some of the breadth of data and detail available in TeleGeography 2001 (published in October 2000), with some comments on differences from the previous year's edition.²

Following a brief executive summary, the main part of the publication has sections on telecommunications carriers, pricing, facilities, Internet backbones, telecommunication traffic analysis, country traffic statistics, and reference blue pages. The sections mostly consist of an introduction, sometimes containing several topical essays on recent industry developments, of which there are fifteen in total, followed by statistical tables and charts (over 250 in total):

Carriers: The growth of international services competition, i.e. the number of authorized international carriers by country (annually for five years); market shares of international carriers by country for up to 10 years, the top 40 international carriers, their traffic and revenue, the top US carriers; a primer on the new business of bandwidth exchanges, introducing the currently very fluid picture of business practices in this area. A feature of the previous report, not contained in the latest report, was a large section on global carrier ownerships, showing affiliates and acquisitions current to mid-October 1999.

Pricing: Overview of international pricing trends, i.e. cost components of an international call to/from the US on selected routes, including a list of international

² For further information, see <http://www.telegeography.com/>.

settlement rates; a primer on interconnection accounting, which discusses the different methodologies used to set interconnection rates, data on retail prices for a three minute call by country, price trends etc.; a long essay and critical discussion of network-to-network payments for Internet Telephony and other IP traffic streams written by George Staple, the founder of TeleGeography Inc.

Facilities: An essay on Metropolitan Area Networks (MANs), which are seen as one of the most dramatic areas of development in telecommunications over the next few years; regional data on submarine cable systems and their capacity, international circuit usage data by U.S. carriers (the latter provide a rough proxy for the balance of voice and data traffic on international network connections to the US), charts and tables on international communication satellites.

Internet Backbones: An essay on International Internet Backbones, followed by the results of a study completed in October 2000 of over 300 International Internet Service Providers (IISPs) (maps and tables presented focus on bandwidth, not traffic as such). Tables and charts on inter-regional Internet bandwidth, the top 50 International Internet routes, the top 50 international Internet Hub cities and major international backbone routes (the latter both worldwide as well as by region). In contrast to the previous edition, data on many African countries have now been included.

Traffic Analysis: An overview of international traffic trends shows the dramatic effect of the proliferation of telecom competition, in many cases only initiated in 1998, on the market shares of incumbents. This is followed by an analysis of alternative telecommunications traffic arrangements (i.e. 'bypass' and 'refile' traffic), and a report on the first annual TeleGeography survey of international Voice-over-Internet-Protocol (VoIP) traffic, indicating how much international VoIP traffic there is, and where it is going. The latter is reported in form of the top 20 U.S.-originated VoIP Routes in 1999 and 2000. Next, an analysis of data on international mobile traffic is provided. Three international call quality metrics are reported by country. The changes in just one year, i.e. from 1999 to 2000, are intriguing (the 1999 data are reported in TeleGeography 2000). Several reasons are discussed for the reported fall in call quality in many cases, e.g., the dramatic rise in mobile phone use.

Traffic Summary: Figures, charts and tables on global traffic trends, including aggregate data by region and country, mainly for 1999. International traffic indicators include, e.g. outgoing and incoming traffic (in million minutes), telecom traffic

minutes per capita, number of main lines, minutes per mainline. Data on international traffic by route, traffic imbalances on selected routes etc.

International Traffic Statistics: More than 130 pages of charts and tables showing the largest telecommunications routes for more than 120 countries and territories, measured by outgoing traffic (millions of minutes and percentages), mostly for 1999. National traffic balances for up to three years (1996 –98). When using the data for analysis, the reader should make sure to carefully read the methodology page. Despite TeleGeography's effort to ensure comparability, problems remain which are more severe for some countries than others. Also, the number of countries included in Telegeography is not the same each year. About 30 countries, mainly from Africa and Asia, were added in Telegeography 2001. Data for a small number of countries were either dropped or not updated from the previous edition.

Global Reference Blue Pages: Tables showing national telecommunications indicators and telephone traffic by country/territory (GDP, Population, number of main lines, lines per 100 people, number of cellular users, number of international carriers, number of Internet hosts (all for 1999)), outgoing and incoming minutes, traffic balances (for 1998 and 1999). World and North American dialing codes in map and list form. Lastly, there is a one page primer on bits.

To summarize, like the sector it analyses, TeleGeography is not static, but innovative and dynamic. It does not just provide an annual update of major telecommunications traffic statistics by country, but in addition emphasises new telecommunications sector developments, accompanied by new statistics wherever possible.

2.2 Hubs and Spokes

Recently, TeleGeography Inc. has become more involved in the analysis and measurement of the Internet and intends to produce a series of reports on the topic. Hubs and Spokes (TeleGeography, 2000b) is the first report of its Internet Program. It explores the Internet and IP-based technologies, focussing on international industry structure, network topology, and traffic flows.

Following a brief introduction, the remainder of the publication has two main parts. Part one, an essay called Global Internet Primer, is a very useful introduction to the

Internet. It contains subsections on the history of the Internet and its architecture, its financial structure, Internet governance issues, an overview of the demand drivers that are behind the Internet's rapid growth, including a whole sub-section on VoIP, all illustrated by charts and tables.

Part two of Hubs & Spokes provides detailed statistics about the Internet. It includes the following sections:

International Internet Bandwidth: This contains a long table listing international ISPS by country, which is not contained in TeleGeography 2000 or 20001.

Connectivity: This section contains data also available in TeleGeography, i.e. inter-regional Internet bandwidth, largest international Internet routes, top 50 international Internet Hub cities, top 50 international backbone routes etc.

Exchanges: Internet exchange points by country (they are not contained in TeleGeography 2001).

International Internet Indicators: Network metrics, i.e. host count by top-level domain and by country (the tables are from TeleGeography 2000, but they are not contained in the latest edition of TeleGeography).

Blue Pages reference section (Internet and telephone country codes).

With Hubs and Spokes, TeleGeography Inc. continues its strategy of product differentiation.³ This first of what is planned as an annual guide to the global Internet does not contain that much value added in terms of Internet statistics compared to TeleGeography 2000. However, this seems to be changing, if TeleGeography 2001 is taken as an indication of future developments. For those interested specifically in the analysis of the global Internet instead of the wider telecommunications sector, Hubs and Spokes provides an excellent and so far relatively low cost guide to current issues and relevant statistics at a detailed level.⁴

2.3 Comments

³ Other TeleGeography publications focus on international bandwidth, new international carriers, time series traffic data, maps showing global telecommunication flows and infrastructure etc. For an up-to-date list and the cost of publications, see <http://www.telegeography.com/>.

⁴ The price of Hubs and Spokes is listed as US \$ 195.

Inevitably, the speed of change in the telecommunications sector shortens the useful lifespan of many of the ‘established’ telecommunications traffic indicators, requiring the constant re-assessment of existing indicators and the development of improved or new ones. For example, Kelly and Woodall (2000) point out that traffic statistics reported by public telecommunications operators will carry less weight in future because an increasing share of traffic is simply not being reported. The reasons given include: 1. More and more voice and fax traffic is bypassing public switched telephone networks, e.g., by using the Internet instead. 2. Many of the smaller carriers are not obliged to report traffic statistics to regulators. 3. Increasingly, traffic is not passing through the traditional accounting rate system.

It seems fair to say that TeleGeography has taken these points into account by, e.g., providing separate VoIP traffic data for the first time in TeleGeography 2001, without being currently able to fully address them. Their main country traffic statistics do exclude the increasing share of VoIP traffic, among other things. Also, bandwidth is used to proxy Internet traffic flows. As noted in TeleGeography 2001, bandwidth has recently been growing faster than Internet traffic, making the former a less useful proxy for the latter, though it probably still is the currently most useful indicator of the geography of the Internet and anticipated network use (Abramson, 2000). Meaningful Internet traffic statistics do not yet exist. Current proxies for Internet traffic, besides the number of host computers by country and domain name⁵, include: Web servers and secure web servers by country; country-by-country Internet user counts; languages on the Internet; content architecture.⁶

In short, there is currently no aggregate summary statistic of the total amount of telecommunications traffic and information flows taking place, and the chances of obtaining such a statistic seem lower than ever. There are many other mechanisms besides phone, fax and the Internet by which information and knowledge is

⁵ They are derived from surveys conducted by the Internet Software Consortium, in association with Network Wizard (see <http://www.isc.org/ds/>).

⁶For references, see Abramson’s (2000) survey of Internet globalisation indicators. Readers who want to explore the many efforts to analyse, map and visualise the geography of the Internet and the World-Wide-Web are referred to the comprehensive website maintained by Martin Dodge of the Centre for Advanced Spatial Analysis, University College London (see <http://www.cybergeography.org>), especially to the ‘Atlas of Cyberspaces’ and the ‘Geography of Cyberspace Directory’.

transferred between countries, e.g., business practises of multinational companies and foreign direct investment, joint ventures and licensing agreements, business travel and international conference attendance, education and training received abroad, migration of skilled people, importation of journals and books, purchase of patents and trademarks, the importation of capital goods and services, reverse engineering, etc.⁷ Moreover, these cover not only explicit or codified knowledge, but also tacit knowledge. The latter tends to be easily forgotten in the discussion of the information and knowledge-based economy.⁸

3. Some Other Data Sources and Indices

While TeleGeography and Hubs and Spokes are indispensable for industry analysts, regulators and policy makers who have to have a detailed understanding of telecommunications sector developments, economists and other social scientists on smaller budgets who are interested in commentary on telecommunications and Internet developments and quantitative indicators are advised to check out the following sources (the list is not exhaustive):

- The International Telecommunication Union's publications on telecommunications indicators (see <http://www.itu.int/ti/>).
- The OECD's Science, Technology and Industry Directorate publications, e.g., the bi-annual OECD Telecommunications Database and Telecommunications Outlook, which list many indicators, as well as the OECD's work on new Internet and electronic commerce indicators (see <http://www.oecd.fr/dsti/sti/index.htm>).
- UNESCO's publications on world communication and information (see, e.g., UNESCO, 1999).
- The World Bank, especially the publications of its Global Information and Communication Technologies Department (see <http://www.worldbank.org/ict/>) and its Information and Development Programme (see <http://www.infodev.org/>).

⁷ See the discussion in Engelbrecht (1997).

⁸ See Lamberton (1997) for a critical discussion of this issue.

However, given the perplexing complexities of the information age, a recent major report on the 'networking' revolution argued that prospects for statistically sound analysis are very poor, particularly in regard to Internet activity. Instead of basing analysis on any robust statistical observations, it has to be based on 'qualitative hypotheses' (Analysys, 2000). It is expected that the historically strong and simple correlation between economic output and networking activities will weaken, and a wider range of causal relationship than in the past will emerge (ibid., p. C4).

Despite this, there is no shortage of attempts to devise new quantitative indices, some of which are more questionable than others. Here I briefly discuss two attempts to derive aggregate quantitative indices that purportedly capture some of the essence of the information age. The first is the Herculean attempt by Lynn and Varian (2000) to measure the total stock of information in the world, the other is an attempt to derive an aggregate information society index.

Lynn and Varian (2000) try to measure world information production by estimating how much space (in bytes) it would take to store all information produced and communicated. The production of content by media type (paper, film, optical, magnetic) and some communication flows are translated into a common standard (i.e. terabytes), while trying to adjust for duplication of content, data compression etc. Lynn and Varian's estimates for the Internet are of the 'deep web', which includes the content of searchable databases and is estimated to be 400 to 550 times larger than the commonly defined World Wide Web (BrightPlanet, 2000). Their overall estimate indicates that the world produces between one and two exabytes (i.e. 10^{18} bytes) of new information each year, or roughly 250 megabytes for every man, woman, and child on earth, indicating a severe case of 'information overload'. However, data on the consumption of these mountains of information suggest that the total time spent in media access has changed little, i.e. less and less of the information produced is actually consumed.

What seems missing from Lynn and Varian (2000) is the geographical dimension (flows are converted into stocks), and some indication of how the estimates can be related to economic and social statistics, which is necessary for analysing their meaning and implications. There are other major limitations. Lynn and Varian's

estimates relate to explicit information. Tacit information is neglected. Also, even the 'deep web' does not account for all information. For example, private, intranet information hidden behind firewalls (i.e. a large part of the information created by private and public bureaucracies) is neglected (BrightPlanet, 2000, p. 5).⁹

The second example of an aggregate index constructed from a large number of variables is the 'Information Society Index', which is advertised as the world's first measure of 55 nations' abilities to participate in the information revolution (World Times/IDC, 2001). Now in its fifth year, it is made up of 23 computer, information, Internet and social infrastructure variables.¹⁰ The weighting given to these variables is difficult to verify and has changed over time, i.e. the reliability of the index is difficult to judge. In general, when such commercially produced aggregate indexes are used to rank countries internationally, they provide a false sense of precision. However, this does not prevent them from being widely reported in the news media.¹¹ Both examples indicate that internationally agreed upon measurement standards are urgently needed.

4. Concluding Comments

Researchers wanting to explain the meaning and significance of global communication and information flows have to analyse telecommunication and Internet statistics in conjunction with other relevant statistics, i.e. economic, business, social, political or institutional statistics at the micro or macro level of analysis. I have argued elsewhere (Engelbrecht, 1997, 1998) that this requires (a) a move away from the prevalent emphasis on the information *technology* economy back towards the information economy in the wider sense, as well as (b) adoption of a 'communication

⁹ In some respects Lynn and Varian's (2000) measurement project is reminiscent of Porat and Rubin's (1977) attempt to define and quantify information activities in an economy. Certainly the assumptions and compromises made in the re-working of existing data are at least as heroic in the former work compared to the latter.

¹⁰ While the aggregate index is reported free of charge, the accompanying background reports sell at a hefty price.

¹¹ Another recent index is the 'index of technological progress' used by Rodriguez and Wilson (2000) to assess the capacity of 110 countries to process information. It is more limited than the Information Society Index in that it is an index of IT products only. For a critical assessment, see the reviews by Heather Hudson and Nick von Tunzelmann in *Telecommunications Policy*, Vol. 24, Nos. 8/9, 2000 (available at <http://www.tpeeditor.com/contents/2000/24-8+9.htm>).

perspective' on the international information and knowledge system. The latter demands taking into account the interrelated aspects of information or knowledge creation, transmission and absorption. Such an approach can also serve as an integrating device for the various economic and non-economic literatures that exist on each of these aspects. The success of firms and countries in the information age depends on the appropriate mixture of the components underpinning the communication perspective (ibid.).

The developments in telecommunications and the Internet are intimately related to the 'digital economy'. Research into the latter is beginning to indicate that positive productivity impacts of IT capital are mainly due to the complementary innovations it facilitates, in particular organisational changes and investments in complementary organizational capital (i.e. the information economy in the wider sense). For example, Brynjolfsson and Hitt (2000, p. 25) argue that firm level data suggest that the recent productivity increases detectable in the US macro data have their roots in over a decade of computer-enabled organizational investments.

Many authors have pointed out that current data collection activities by statistical agencies are inadequate (see, e.g., Haltiwanger and Jarmin, 2000). To properly measure and understand the impact of telecommunications and the Internet on the economy and society not only requires new and widely agreed upon telecommunication and information flow indicators, but also an improvement in the measurement of 'conventional' statistics. Examples of the latter are improved output measures of non-goods producing sectors, quality adjusted price deflators, improved productivity measures and human capital stock data, and many more (see ibid.). In short, developments in telecommunications and in the digital economy are forcing statistical agencies to rethink the measurement of even the basic building blocks of the national accounts, i.e. outputs, inputs and prices (ibid., p. 31).

However, it gets worse. There is also a need to develop more appropriate techniques to analyse the impact of the developments in telecommunications and the digital economy. In a recent anniversary issue of the Journal of Econometrics, Phillips (2001, p. 23) commented on the challenges facing econometric research on productivity issues: "The changes in the nature of economic activity now being forged by the

rapidly emerging internet economy certainly seem unprecedented...How do we model trends when such fundamental changes are taking place?...Simply put, the inadequacy of our modeling apparatus in the face of an issue of such great importance is staggering.” Durlauf (2001), in the same issue, sees the need for a new growth econometrics. While both authors’ comments mainly pertain to macro-level analysis, similar issues arise with micro-level econometric analysis.

In short, the issue of how to measure the geography of telecommunications and of the Internet (i.e. the stocks and flows of information instead of information technology as such), and how to analyse them, should provide inexhaustible opportunities for further research. The TeleGeography publications are likely to be important resources and guideposts in these endeavours.

References

Abramson, B. D., 2000. Internet globalisation indicators, Telecommunications Policy 24, 69-74.

Analysys, 2000. The Network Revolution and the Developing World. Analysys Report Number 00-216, Final Report for World Bank and infoDev. Available at <http://www.infodev.org/projects/400networkingrevolution/>.

BrightPlanet, 2000. The Deep Web: Surfacing Hidden Value. Available at <http://www.completeplanet.com/Tutorials/DeepWeb/index.asp>.

Brynjolfsson, E., Hitt, L., 2000. Beyond computation: Information technology, organizational transformation and business performance, Journal of Economic Perspectives 14, 23-48.

Durlauf, S., 2001. Manifesto for a growth econometrics, Journal of Econometrics 100, 65-69.

- Engelbrecht, H.-J., 1997. The international economy, knowledge flows, and information activities, in: D. Lamberton (Ed.), *The New Research Frontiers of Communications Policy*. North-Holland, Amsterdam, pp. 19-42.
- Engelbrecht, H.-J., 1998. A communication perspective on the international information and knowledge system, *Information Economics and Policy* 10, 359-367.
- Haltiwanger, J., Jarmin, R. S., 2000. Measuring the digital economy, in: E. Brynjolfsson and B. Kahin (Eds.), *Understanding the Digital Economy: Data, Tools, and Research*, MIT Press, Cambridge, Mass., pp.13-33.
- Kelly, T., Woodall, M., 2000. Telecom traffic indicators, *Telecommunications Policy*, 24, 155-159.
- Lamberton, D., 1997. The knowledge-based economy: A Sisyphus model, *Prometheus* 15, 73-81.
- Lyman, P., Varian, H., 2000. How much information?, *School of Information Management and Systems*, University of California at Berkeley. Available at <http://www.sims.berkeley.edu/how-much-info/>.
- Phillips, P., 2001. Trending time series and macroeconomic activity: Some present and future challenges, *Journal of Econometrics* 100, 21-27.
- Porat, M., Rubin, M., 1977. The information economy, in nine volumes. Office of Telecommunications, Spec. Publ. 77-12, US Department of Commerce, Government Printing Office, Washington, DC.
- Rodriguez, F., Wilson, III, E. J., 2000. Are poor countries losing the information Revolution?, *InfoDev Working Paper*, The World Bank, Washington, D.C., May. Available at <http://www.infodev.org/library/working.htm>.
- TeleGeography Inc., 1999. *TeleGeography 2000*. Washington, DC.

TeleGeography Inc., 2000a. TeleGeography 2001. Washington, DC.

TeleGeography Inc., 2000b. Hubs and Spokes: A TeleGeography Internet Reader. Washington, D.C.

UNESCO, 1999. World Communication and Information Report 1999-2000. Paris.

World Times/IDC, 2001. 2001 Information Society Index. Available at <http://www.worldpaper.com/>.