Taking the Measure of National Bandwidths: evolving patterns of the international digital divide for 1986 - 2013

Martin Hilbert
University of California, Davis
hilbert@ucdavis.edu

Abstract
The article measures different aspects of the evolution of the international digital divide between 1986 and 2013. The traditional approach of measuring digital development in terms of telecom subscriptions is increasingly becoming obsolete, especially as mobile phone penetration reaches population saturation levels. It shows that the bandwidth divide continues to be dynamic and very persistent, as it can closely be linked to the international income divide. Novel findings reveal that only 3 countries host 50% of the globally installed bandwidth potential (10 countries 75%); that the digital divide between high- and low-income countries in terms of kbps per capita has increased between 2001 and 2008, only decreasing below historic levels very recently during 2012-2013; and that the contribution of mobile to global bandwidth is continuously fluctuating with incessant technological innovation, reaching almost 40% in both 2001 and 2013, while falling to 16% in 2007.

1. Introduction
In the digital age, it is essential for any kind of public policy or private business strategy to distinguish unique characteristics of the evolution of the global infrastructure of Information and Communication Technology (ICT). This starts with a general macro-level outlook on the diffusion of this fundamental infrastructure. The article measures to historical evolution of the international telecom infrastructure in terms of the installed telecommunication bandwidth capacity between 1986 and 2013.

The motivation behind this effort consists in the fact that technological progress has rendered the traditional metric of international digital progress obsolete. Traditionally global digital development is assessed in terms of telecommunication subscriptions [1,2]. These statistics mainly stem from the United Nations International Telecommunication Union (ITU) [3]. ITU has undertaken an important and sustained effort over recent decades to collect this data from administrative registries of national telecommunication authorities in a harmonized manner. These same databases have shown that the number of mobile and fixed telecom subscriptions per person are increasingly reaching a certain level of global saturation, including 6.5 billion mobile phone subscriptions worldwide for 6.9 billion people in 2013. Since there seems to be a certain limit in how many telecom devices a person handles, the international divide in terms of subscriptions is quickly closing.

This does not automatically imply that the digital divide in terms of equal access to digitalized information is closing as well. Nowadays bandwidth is not uniformly distributed among subscriptions. Figure 1a shows this distinction was not relevant a few decades ago. The late 1980s showed a linear relationship between subscriptions and bandwidth, as there was only fixed line telephony around, all with the same bandwidth. Today’s digital infrastructure offers a myriad of different bandwidth options, which leads to an L-shaped non-linear relationship exhibit in Figure 1b. The Figure shows that ICT diffusion seems to hit an invisible wall at around 2-3 subscriptions per person. However, the digital divide continues at this point, just along a new dimension: the bandwidth dimension. “This implies that we have now moved into a second, more mature, and also more persistent stage of the digital divide. The first phase consisted of a universalization of the required technological infrastructure. The second stage consists of an endlessly evolving inequality of technological capacity based on this more and more universalized infrastructure” [4]. This second stage of the digital divide becomes extremely relevant for an age of big data, in which data (not mere access) have become a driver of growth and progress [5,6].

Despite this fact, even recent research continues to disregard this reality and still refers to the number of telecom subscription, especially for large-scale international statistical tests (for example [7] and [8]). This seems to have less to do with the fact that scholars are not aware, but that globally harmonized subscriptions data is readily available, and bandwidth and traffic data is not. Reminiscent of the famous drunk who is looking for the lost keys under a well-lit lamppost far away from the dark site where he dropped...
the keys, analysts continue to recur to the readily available and harmonized ITU database.

![Figure 1](image-url)  
Figure 1. Subscriptions per capita (fixed and mobile) vs. kbps per capita (voice and data in optimally compressed kbps of installed bandwidth potential). N = 100 countries. Size of bubbles: log population. (a) 1986 (b) 2013.

Changing this current practice faces two challenges. One the one hand, important efforts are currently underway to explore new ways to measure relevant aspects of this second stage of the digital divide. Leading authorities like the U.S. Federal Communications Commission or Ofcom in the UK have started to produce detailed annual reports at the national level [9,10]. At the same time, researchers from academia have started to explore a combination of metrics from private and public sector sources to evaluate the implications of the diverse bandwidth landscape for growth and competition [11]. These efforts usually go deep and explore different aspects, but are naturally demographically limited in scope and in the analyzed timespan, since globally harmonized long term time series are not available for more detailed metrics.

On the other hand, the challenge consists in showing why we should care globally about this new dimension of the digital divide [4,12,13]. Why is it important to undertake a sustained effort to produce, harmonize and analyze the global evolution of bandwidth? This article falls into this second group of research. As such, the article works with a rather rudimentary proxy for bandwidth and traffic, but is able to cast a wide global net to show a rough outline of ongoing dynamics for 172 countries [14], corresponding to 96 % of the world’s population and 99 % of the world’s Gross National Income (GNI).

The analyzed time series capture 27 years, which covers the entire transition from almost inexistent digitalization (less than 1 % of the global information stockpile was digital in 1986), to the full-blown digital age with almost all of it in digital format [15].

2. Methodology: a statistical challenge

The undertaking of creating global time series data faces two main statistical challenges, one related to the creation of national statistics among many countries (in space) and the other one to the creation of normalized time series (in time).

2.1. Installed national bandwidth

The creation of national statistics requires three main inputs: the number of telecommunication subscriptions (fixed and mobile); the kind of access technology per subscription (like DSL, GSM, etc); and the corresponding bandwidth per access technology. The first two are provided mainly by the well-known ITU database [3], which we complement with other sources for a variety of data gaps (especially for the diffusion of fiber optics [16] and wearables and tablets [17,18]).

The third one is trickier. Up until roughly 2006/2007, it was more straightforward to assign a certain bandwidth performance to a specific access technology. For example, a digital fixed-line phone
provides a general bandwidth of 64 kbps, an ISDN BRI internet modem 128 kbps, and the voice-transmission of a GSM mobile phone also 128 kbps (all uncompressed). After the introduction of global broadband solutions like DSL and cable modem, and 3G mobile telephony, the direct assignment of bandwidth to specific technologies becomes less viable.

We chose to approximate the installed bandwidth by recurring to crowd-sourced data from NetIndex [19], which allows us to maintain a very wide geographical focus. NetIndex has gathered the results of end-user-initiated bandwidth velocity tests per country per day since 01/01/2008. This crowd-source method results in very large samples. For example, an average of 179,822 tests per country per day were gathered in 2010 through Speedtest.net and Pingtest.net. The resulting database is seen as “the best of the currently available data sources for assessing the speed of ISP’s broadband access service” [20]. We consult both upload and download test and add both in our assessment of broadband capacity. We assume that the bulk of users of these tests use broadband connections (i.e. DSL, cable modem or fiber optics for fixed, and 3G, 4G for mobile) For details see [21].

After verifying national details with a large variety of commercial data from national telecom operators (especially for fiber optics solutions, which are more influential for our results), we created national averages for fixed broadband speeds. We followed the same strategy for fixed and mobile broadband, while we complemented the mobile speed-test data from NetIndex [22] with the quarterly reports from Akamai [23].

The sum of the product of the number of subscriptions and their respective broadband performance provides the installed national bandwidth potential. It is important to point out that this metric does not measure actual traffic flow, but works with the number the end user gets when performing an online speed test. Hence it merely refers to an installed potential [24]. It is important to point out that the network in its entirety would collapse if all users would demand their installed bandwidth capacity simultaneously (or 24 hours a day). Previous work that compared the installed bandwidth potential with actual traffic flows found that “the average user only uses its promised full bandwidth for effectively nine minutes per day” [25]. Users might sit in front of a computer for hours, but the full bandwidth is on average only filled with traffic during a much smaller proportion. Since actual traffic is a relative proportion for all countries, this issue less severe for purposes of international comparison. Relative comparisons would turn out equivalent if it were assumed that usage intensity within provided bandwidth would not differ among societies. This is of course a simplification, since traffic prices and cultural habits differ among countries and influence bandwidth usage intensity. However, the indicator of installed bandwidth capacity gives good first idea of the global situation and trajectories for the purpose of relative comparisons.

### 2.2. Normalizing information time series

The creation of meaningful time series for communication capacities hinges on temporal normalization on technological progress in compression algorithms. Lossless compression allows to send the same amount of information with less binary symbols [26]. Compression is omnipresent in the digital age and represents the core of solutions like GSM, CMDA, JPEG, MPEG, etc. It has been shown that during the last three decades, the amount of information transmitted through the same hardware channel has been significantly increased through the ever more sophisticated use of compression algorithms, which shows that compression is an important driver of the global information explosion [27].

Achievable compression rates vary significantly among different kinds of content, depending on the amount of redundancy in the source. For example, video content is usually more compressible than alphanumeric text. Therefore, the creation of the average compression rates at certain points in time requires two main inputs: the kind of content flowing through fixed and mobile network; and achievable compression rates for different kinds of content. We estimate the amount of content by distinguishing between text, images, audio and video, for fixed and mobile, upstream and downstream, according to five world regions [28,17]. For details see [21]. We estimate the corresponding average state of the art of compression rates per kind of content every seven years, for 1986, 1993, 2000, 2007 and 2014 and interpolate between them. This spacing gives enough room to identify the dominating compression technology at a given point in time. We also estimate the optimally achievable compression rate, which approximates the entropy of the source [25,26].

We then normalize the content in time, acting as all content would always be optimally compressed. Since the entropy of the source is a constant and since today’s compression algorithms have gotten quite close to it, this gives us a stable ground to stand on while evaluating the incessantly moving technological frontier in compression. The result is reminiscent of what economists do when normalizing on inflation rates. It allows us to create meaningful time series that quantify comparable amounts of information through
time, not merely the quantity of more or less efficiently compressed binary symbols [21,25]. The resulting unit of measurement are optimally compressed bits, which we represent as kilobits per second (kbps) for telecommunication solutions.

Figure 2. Global shares of technologies: (a) subscriptions, (b) installed bandwidth potential, (c) installed bandwidth potential per subscription

3. Results: bandwidth divides

Figure 2 visualizes the source of the discrepancies between the tradition accounting of subscriptions, versus the accounting of bandwidth potential measured in optimally compressed kbps. Both, the accounting of subscriptions (Figure 2a) and bandwidth capacity (2b) evidence the elimination of the dominance of fixed line telephony, which was the dominating form of distance communication in the late 1980s. In terms of the number of technological devices, this dominance was clearly replaced by mobile telephony, especially narrowband 2G and 2.5G phones and in more recent years, broadband smart phones. However, in terms of telecommunication capacity, it shows that fixed-line broadband plays the dominant role. Representing less than 9 % of the world’s subscriptions, DSL, cable modem and fiber optics contribute 60 % to the global bandwidth potential.

It is interesting to observe that the relation between subscriptions and bandwidth is neither linear, nor stable. For example, while we detect a monotonically increasing share of mobile phone subscriptions (Figure 2a), the contribution of mobile to global bandwidth is continuously fluctuating with incessant technological innovation, reaching almost 40 % in both 2001 and 2013, while falling to 16 % in 2007 (Figure 2b).

The increasing importance of mobile broadband in recent years is noticeable, as is the most recent contributions of tablets and wearables. Especially the latter have the potential to once again shift the picture of the global telecommunication landscape in the short-term future. Figure 2c shows the corresponding bandwidth averages per subscription.

3.1. The divide between world regions

Figure 3 looks at the global total in terms of global income groups (following the classification of the World Bank of 2015 [29]). The last three decades show a gradual loss of dominance of today’s high income countries. High income countries contributed some 85-86 % of the global subscriptions and installed bandwidth potential in 1986, but in 2013 merely 30 % of subscriptions (Figure 3a) and 66 % of bandwidth (Figure 3b). This also shows that global deconcentration in terms of subscriptions was twice as strong as in terms of bandwidth. The difference stems from the simple fact that not all subscriptions are equal in their communicational performance.

Comparing these results with the global shares of population and Gross National Income (GNI) (Figure 3c and 3d), it becomes clear that the diffusion dynamic of the number of subscriptions follows existing patterns in population distribution. Especially the diffusion of mobile phones during recent decades has contributed to the fact that both distributions align. The match between the number of subscriptions and population shares is not 1-to-1, but is close in 2013. For example, upper middle income countries host 34 % of the world’s population and 37 % of telecom subscriptions. This match with population shares is not evident for bandwidth capacity in kbps. Bandwidth follows the signature of economic capacities much
closer. After only a few decades, both processes align impressively well. For example, in 2013 upper middle income countries host 23% of the world’s GNI and 28% of global bandwidth potential. This shows that the digital divide in terms of data capacity is far from being closed, but is rather becoming a structural characteristic of modern societies, which could turn out to be as persistent as the existing income divide [4].

Another interesting insight from Fig 3b is that the evolution of bandwidth is also not monotone among countries. High income countries controlled a dominating share of 82% in both 1993 and 2007, but their share was as low as 71% in 2001, and most recently has come down to the historic low of 66%.

This non-monotonicity becomes even more evident when analyzing these tendencies as per capita ratios. Figure 4a shows that in 2003 high income countries had 11 times more bandwidth per capita than low income countries, which increased to a gap of 18:1 in 2007, to fall to 8:1 in 2013. Figure 4b shows this ratio as a continuous line and contrasts it with the monotone tendency of the more traditional metric of subscriptions per capita (so-called ICT penetration rates). The reason for this in fluctuating tendency of the divide in terms of bandwidth capacities is technological change combined with recurring patterns of unequal dynamic of technology diffusion. The decreasing divide during the period until 2000 is explained by the global diffusion of narrowband internet and 2G telephony. The increasing nature of the divide between 2001 and 2008 is due to the global introduction of broadband for fixed and mobile solutions. The most recent decreasing nature of the divide is evidence of the global diffusion of broadband. Every new technological innovation has the potential to increase the divide once again, as every innovation once again runs through the process of technology diffusion through social networks [30], which is never neither instantaneous, nor uniform, and therefore inevitably creates a divide.

Figure 4a shows another very important fact. The divide continuously increases in absolute terms as bandwidth increases. In 2003, the average inhabitant of high income countries had 11 times more bandwidth per capita than low income countries, which increased to a gap of 18:1 in 2007, to fall to 8:1 in 2013. Figure 4b shows this ratio as a continuous line and contrasts it with the monotone tendency of the more traditional metric of subscriptions per capita (so-called ICT penetration rates). The reason for this in fluctuating tendency of the divide in terms of bandwidth capacities is technological change combined with recurring patterns of unequal dynamic of technology diffusion. The decreasing divide during the period until 2000 is explained by the global diffusion of narrowband internet and 2G telephony. The increasing nature of the divide between 2001 and 2008 is due to the global introduction of broadband for fixed and mobile solutions. The most recent decreasing nature of the divide is evidence of the global diffusion of broadband. Every new technological innovation has the potential to increase the divide once again, as every innovation once again runs through the process of technology diffusion through social networks [30], which is never neither instantaneous, nor uniform, and therefore inevitably creates a divide.

Figure 4a shows another very important fact. The divide continuously increases in absolute terms as bandwidth increases. In 2003, the average inhabitant of high income countries had 11 times more bandwidth per capita than low income countries, which increased to a gap of 18:1 in 2007, to fall to 8:1 in 2013. Figure 4b shows this ratio as a continuous line and contrasts it with the monotone tendency of the more traditional metric of subscriptions per capita (so-called ICT penetration rates). The reason for this in fluctuating tendency of the divide in terms of bandwidth capacities is technological change combined with recurring patterns of unequal dynamic of technology diffusion. The decreasing divide during the period until 2000 is explained by the global diffusion of narrowband internet and 2G telephony. The increasing nature of the divide between 2001 and 2008 is due to the global introduction of broadband for fixed and mobile solutions. The most recent decreasing nature of the divide is evidence of the global diffusion of broadband. Every new technological innovation has the potential to increase the divide once again, as every innovation once again runs through the process of technology diffusion through social networks [30], which is never neither instantaneous, nor uniform, and therefore inevitably creates a divide.
big data world, in which the amount of data is becoming a crucial ingredient for growth [5,6].

Figure 4: Bandwidth potential per capita: (a) high income group versus rest of world. (b) ratio of high income countries versus rest of world for bandwidth and subscriptions.

3.2. The divide between countries

We can now also look at the evolution of the divide country level and add an additional dimension to the logic behind Figure 1. Figure 5 adds the dimension of income per capita (GNI per capita) on a third axis of Figure 1b. It shows that countries do not require much income to evolve along the axis of subscriptions per capita. Rich and poor countries alike are able to increase this indicator. After having reached about 2 subscriptions per capita (roughly one fixed and one mobile solution), saturation sets in, regardless of income. However, the digital divide continues to evolve. No non-linear L-shaped saturation logic can be detected when analyzing the relationship between kbps/capita and GNI/capita. The relationship rather follows the trajectory of a diagonal linear one-to-one relationship between bandwidth and income. This reconfirms the previous finding that the digital bandwidth divide follows the logic of income, and therefore has the potential to become as persistent as the income divide.

Some interesting outliers are detectable in Figure 5. Countries like Switzerland (CHE), Norway (NOR) and Australia (AUS) have less bandwidth potential than their income level would suggest. The U.S. (USA), Canada (CAN), France (FRA) and Germany (DEU) also fall in this underperforming category. Others, like South Korea (KOR), Hong Kong (HK), Japan (JPN), Lithuania (LTU), Russia (RUS) and China (CHN) have more installed bandwidth that is to be expected with regard to their income. South Korea, Japan and Lithuania has long been identified as global best practices of government-led broadband role out [31,32,33]. This provides evidence for the feasibility of influencing global broadband development through pro-active public policy.

The rise of several Asian countries in this list provides evidence for another important trend in the global distribution of bandwidth capacity. Asia share of global bandwidth potential increased from 23 % in 1986 to 51 % in 2013. In other words, hosting 57% of the world’s population, Asia now also hosts more than half of globally installed bandwidth potential. Figure 6 shows that this leading role of Asian countries has undergone an internal change. In 2001, South Korea and Japan represented 16 % of global bandwidth potential, while Russia and China represented 15 %. By 2007, South Korea and Japan expanded its global share to 27 %, mainly driven by their early adoption of fiber optic broadband. Russia and China feel to 11 %. By 2013, the latter expanded their global share to 26 %, overtaking South Korea and Japan with 17 %. This implies that together these four countries capture 43 % of global bandwidth. On average, the installed bandwidth potential of South Korea and Japan sustained a compound annual growth rate of some 27 % over 1986-2013, while Russia and China grew 44 % per year.

Figure 7 shows that especially the last few years evidenced some important rearrangements with regard to which countries lead the pack in the global bandwidth race. Back in 1986, a group of traditionally highly developed countries were found in the global top ranks. The U.S., Japan, France and Germany hosted more than half of global bandwidth. By the end of the 1990s, China had already started to join the top ranks, hosting 7 % of global bandwidth. By the end of 1996. A decade later, by 2006, three Asian countries, Japan, South Korea and China occupied ranks 2, 3 and 4. The share of global bandwidth of the U.S. had almost shrunk to half its size from the late ‘80s by 2010, merely representing 15 %. At the same time, Russia is regains its strength as an important player in terms of telecommunication capacity. By 2013, China overtakes the U.S. for the first time, hosting some 5.1 petabits per second, versus 4.6 in the United States.
It is interesting to notice that the share of the remaining countries has stayed surprisingly stable at around one quarter of global bandwidth. This means that historically, a very small group of countries dominates global bandwidth. The top 10 countries provide roughly 75% of globally installed bandwidth potential. However, both Europe and North America were replaced at the top of these ranks by Asian countries. It is also interesting to notice that the top-3 countries usually captured a share of about 40 – 45%, but expanded its influence to 50% in 2013.

**Figure 5:** Different perspectives on three dimensions of the digital divide: subscriptions per capita; kbps per capita; GNI per capita. N = 100 countries for 2013. Size of bubbles: log population.

**Figure 6:** Installed bandwidth potential in kbps: the rise of Asia
4. Conclusions

International digital development has traditionally been measured in terms of the number of ICT subscriptions. Globally harmonized databases are readily available and are still used as the main proxy in analytical and statistical exercises. While global saturation in terms of subscriptions is rendering this indicator increasingly obsolete and meaningless, this article has shown that the digital divide in terms of bandwidth is actively evolving and showing lots of dynamics on the international level during recent years. While the digital divide in terms of subscriptions is rapidly being closed, the digital divide in terms of bandwidth is rapidly evolving and here to stay. Bandwidth potential is closely linked to income levels, and therefore subject to a similar persistency as global income inequality.

This leads to an interesting challenge in terms of finding the right indicator. Bandwidth has certainly two main dimensions to it: the installed capacity (much in the sense measured here) and effective traffic (the fraction of bandwidth effectively used) [24]. The first is a general condition sine qua non, while the second allows to detect differences in terms of efficiency and effectiveness of infrastructure supply and demand.

There are other, additional metrics that can be used to complement these two fundamental indicators. For example, [12] have added network latency as part of their broadband quality score. Following the traditional argument of going beyond access and looking at usage patterns [34], other have started to complement broadband data with data about social media usage [35] and user skills [36]. Just like the first subscription-driven stage of the digital divide was filled with very different and sometimes contradicting proposals for adequate ways of measuring ICT access [37], the second bandwidth-driven stage of the digital divide will still have to find an adequate way of measurement. While the details are not yet clear, this article made clear that it is important to advance quickly with this task of taking the measure of the internet. Independent of the final motivation of measurement, be it for business models, systems design and management, governance and the public interest, the measurement of the size of the basic telecommunication infrastructure is an essential part of this effort.

Figure 7: Installed bandwidth per country
5. References

14. Following 3 letter ISO code: KOR, HK, DNK, JPN, SGP, MAC, NLD, SWF, FIN, CHE, NOR, LTU, USA, GBR, TWN, CAN, BEL, FRA, ESP, DEU, RUS, PRT, AUS, MDA, EST, NZL, CZE, BGR, SVN, HUN, POL, SVK, ISR, ITA, URY, UKR, GRC, CHN, KAZ, SAU, AZE, THA, CHL, TUR, BRA, PRI, MNG, MEX, MYS, ARG, ECU, JOR, DOM, PAN, JAM, MUS, CRI, COL, KGZ, VNM, GHA, ZAF, EGY, NAM, ZWE, IDN, TUN, SLV, NPL, PHL, GTM, PER, VEN, SEN, BOL, PRY, MLI, HND, NIC, IRN, NGA, DZA, UZB, TJK, CIV, IND, KEN, SWZ, PAK, YEM, RWA, CMR, BGD, TZA, MOZ, MWI, SLE, ETH, NER, ERI, ALB, AND, AGO, ATG, ARM, AUT, BHS, BHR, BRA, BLR, BLZ, BEN, BMU, BTN, BWA, BRN, BFA, BDI, CPV, CAF, TCD, COM, COG, COD, HRV, CYP, DJI, DMA, GNQ, FJI, GAB, GMB, GEO, GRL, GRD, GIN, GUY, ISL, IRL, KIR, KWT, LAO, LVA, LBN, LSO, LBR, LIE, LUX, MDG, MLT, MRT, MAR, OMN, PNG, ROU, KNA, LCA, WSM, SYC, LKA, VCT, SDN, SUR, SYR, TGO, TON, TTO, TKM, UGA, ARE, VUT, ZMB.
16. For example through different regional networks of the Fiber-to-the-House-Council:
   http://www.ftthcouncil.org;
   http://www.ftthcouncil.eu;
24. M. Hilbert and P. López, “How to Measure the World’s Technological Capacity to Communicate,
27 M. Hilbert, “How much of the global information and communication explosion is driven by more, and how much by better technology?,” J Assn Inf Sci Tec, vol. 65, no. 4, pp. 856–861, Apr. 2014.