Information Economics and Policy 25 (2013) 246–256

Contents lists available at ScienceDirect

Information Economics and Policy

journal homepage: www.elsevier.com/locate/iep

# Empirical validation of Metcalfe's law: How Internet usage patterns have changed over time

António Madureira [a,*], Frank den Hartog [b], Harry Bouwman [a,c], Nico Baken [a,d]

[a] Delft University of Technology, PO Box 5, 2600 AA Delft, The Netherlands
[b] TNO, PO Box 5050, 2600 GB Delft, The Netherlands
[c] Åbo Akademi University, Tuomiokirkontori 3, FI-20500 Turku, Finland
[d] Royal KPN, PO Box 30000, 2500 GA The Hague, The Netherlands

ARTICLE INFO

*Article history:*
Received 6 September 2011
Received in revised form 21 April 2013
Accepted 2 July 2013
Available online 25 July 2013

*JEL classification:*
B5
D2
D8
L9
O3

*Keywords:*
Telecommunications
Value of digital information networks
Eurostat data
Metcalfe's law

ABSTRACT

Few doubt that Digital Information Networks (DINs) such as the Internet constitute the basis of a new technology-driven economic era. A large body of literature tries to understand and quantify the value of DINs to help policy makers justify investments in new or improved infrastructures. The prevailing methodological approach is to depict DINs as an observable production input changing the uncertainty regarding the performance of an economic system. In such context, the value of DINs is typically measured with regression techniques between the penetration rate of DINs and economic growth. This approach provides too little insight on the actual causality between DINs and economic value. We recently developed a framework that identified 13 different ways ("capabilities") how users convert information into economic value. In this article, we show how a simple quadratic relation (Metcalfe's law) can be used to quantify how adequate these capabilities are in converting the ability to access information into economic value. To our knowledge, this is the first time that Metcalfe's law is empirically validated as such.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Since the 1980s, the telecommunication sector has been expanding rapidly (Shiu and Lam, 2008). This is mainly caused by the conversion of analoge communication networks designed for telephony or TV services into multi-functional Digital Information Networks (DINs). The exponential growth of services offered over DINs can be explained by many factors, including technological advancements, market liberalization and privatizations. The worldwide extraordinary level of interest in deploying information networks is due to the strong perception that information networks bring economic, social and environmental benefits (Firth and Mellor, 2005). Some authors speculated that DINs may have a similar impact on society as transportation networks had during the 20th century (OECD, 2001). In long wave theory, this information driven economic era is known as the 5th Kondratieff economic cycle (Perez, 2003). A Kondratieff cycle manifests itself by a sinusoidal-like long-term cycle from approximately 40 to 60 years in length with a semi-period of high productivity growth followed by a semi-period of relatively slow growth (Freeman and Louçã, 2001). The benefits of DINs can be observed directly. For example, construction of network infrastructures leads to direct increase in job employment. In addition, the benefits might also be more intangible, such as better quality of health care services,

* Corresponding author.
*E-mail addresses:* ajpsmadureira@gmail.com (A. Madureira), frank.denhartog@tno.nl (F. den Hartog), w.a.g.a.bouwman@tudelft.nl (H. Bouwman), nico.baken@kpn.com (N. Baken).

0167-6245/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.infoecopol.2013.07.002

improved education and organizational efficiency. The Organization for Economic Co-operation and Development (OECD) considered broadband DINs as key to enhancing competitiveness and sustaining economic growth (OECD, 2001). Many governments are increasingly committed to extending DINs to their citizens (Katz et al., 2009), particularly in the developing nations (Kagami et al., 2004). Consequently, the levels of interdependency between users and DINs' providers increased dramatically (Dijk and Mulder, 2005) and the DIN infrastructure became an essential facility for all economic sectors.

In order to justify policy support for further investments in DINs (e.g. in Fiber To The Home (FTTH)), it is necessary to learn from expenditures that have already been made and demonstrate their value. We recently developed a Holonic Framework (HF) which identified and defined so-called "capabilities" of users in DINs (Madureira et al., 2011). Capabilities are mechanisms that users apply to convert information into economic value. In this article, we show how a simple quadratic relation (Metcalfe's law) can be used to quantify how adequate these capabilities are in converting the ability to access information into economic value. To our knowledge, this is the first time that Metcalfe's law is empirically validated as such.

The next section describes the state of the art on studies aiming at understanding the value of DINs, including a brief overview of our HF. Section 3 provides the equations with which the behavior of the capabilities of the HF can be quantified. The characteristics of our data source, our conceptual operationalizations and our validation methodology are described in Section 4. The results of our analysis are presented in Section 5, whereas Section 6 discusses them, identifies potential implications, and describes the limitations of our work. The last section is reserved for our conclusions.

## 2. State of the art

We reviewed 24 studies on the value of DINs spanning a period from 1980 to 2010. These studies can be grouped into three classes: (1) macro-economic studies using general equilibrium theories and/or input–output tables (Katz et al., 2009; Greenstein and McDevitt, 2009; Correa, 2006; ACIL Tasman, 2004; CEBR, 2003; Röller and Waverman, 2001; Hardy, 1980); (2) econometric studies not addressing the issue of causality (Thompson and Garbacz, 2008; Thompson and Garbacz, 2007; Shideler et al., 2007; Duggal et al., 2007; Crandall et al., 2007; Lehr et al., 2006; Datta and Agarwal, 2004; Sridhar and Sridhar, 2004; Madden and Savage, 2000; Madden and Savage, 1998; Greenstein and Spiller, 1995; Leff, 1984); and (3) econometric studies addressing causality deterministically (Majumdar et al., 2010; Koutroumpis, 2009; Shiu and Lam, 2008; Ford and Koutsky, 2005; Cronin et al., 1991). The first class of studies provide a tool to policy analysts to study the effect of DINs across the interdependences and feedbacks of an economy (Borges, 1986). Empirical validation is not addressed due to the nature of the underlying assumptions, e.g. perfectly rational behavior and equilibrium solutions (Farmer and Foley, 2009). Hence, claims such as "the economic impact of broadband development over a ten year period in Germany amounts to 968,000 additional jobs" (Katz et al., 2009) tend to have a speculative character.

Madden and Savage (1998) found that the causality between DINs and economic growth works in both directions. Similar observations were made by Shiu and Lam (2008) who observed a "bidirectional relationship between telecommunications development and economic growth for European countries and those belonging to the high-income group". Thus, the direction of causality is a methodological challenge inherent in disentangling the value of DINs. The results of the class 2 studies, not addressing causality, should therefore be interpreted cautiously. Recently, some econometric studies (class 3) have addressed the issue of causality deterministically. In such context, the value of DINs is typically measured with regression techniques between the penetration rate of DINs and economic growth. However, this approach provides few insights on the actual causal mechanisms that explain how DINs generate value.

In Madureira et al. (2011), we described a framework, labeled Holonic Framework (HF), that provides an overarching account for the intermediate processes between DINs and economic value. The HF provides deeper insights on the causality between DINs and economic value than the economic studies mentioned above. From extensive literature analysis, the HF identifies 13 different ways (so-called "capabilities") how users apply convert information into economic value. Here, they are simply postulated and defined in Table 1, in no particular order. To name a particular capability we mixed action/verb/process specifics, being aware that the result is not always in line with English grammar. For more information on the HF we refer to Madureira et al. (2011) and Madureira (2011).

## 3. Model for value generation by users in DINs

We can derive a number (value) for how effective a capability is in creating economic value from how it is used to generate income. For example, if a worker uses DINs for online education, then he uses adoptability to obtain a certain part of his income. The value ($y_c$) generated by a capability $c$ is dependent on the size $x$ of the DIN. With a larger network more value is extracted by a capability. $k_c$ is the coupling strength between the size of the network and the value generated by capability $c$, and is a measure for $c$'s effectiveness in creating value by accessing information. We assume that the size of the DIN and the coupling strength of each capability are independent.

Metcalfe's law states that the value of a network is proportional to the square of its size, relying on the observation that for a network with $n$ members, each can make $n - 1$ connections with the other members (Metcalfe, 1995). If all those connections are equally valuable, the total value of the network is proportional to $n(n - 1)$, thus roughly to $n^2$. For example, if a network has 5 members, there are 20 different possible connections that members can make to each other. If the network doubles its size to 10 members, then the number of connections does not simply double, but roughly quadruples to 90.

**Table 1**
Labels and definitions of the capabilities.

| Capability | Definition |
|---|---|
| Coordinatibility | Capability of a node/user in a network to manage dependencies between activities that are performed to achieve a goal, i.e. to coordinate itself with other nodes/users |
| Cooperatibility | Capability of a node/user in a network to enter in a relationship with other nodes/users for a common purpose |
| Selectibility | Capability of a node/user in a network to scan or search for the unknown or generate courses of action that improve on known alternatives |
| Biddability | Capability of a node/user in a network to influence other nodes/users by making proposals |
| Adoptability | Capability of a node/user in a network to acquire novel knowledge from other nodes/users and integrate it in its existing internal knowledge structures |
| Creatibility | Capability of a node/user in a network to deliberately and purposely collate knowledge to generate new or novel ways to understand a particular phenomenon |
| Brokerability | Capability of a node/user in a network to act as a broker between unconnected nodes/users |
| Normatibility | Capability of a node/user in a network to share with other nodes/users norms as rules with at least a certain degree of consensus that are enforceable by social sanctions |
| Culturability | Capability of a node/user in a network to share with other nodes/users general assumptions, values and patterns of behavior emerging in time from their interaction, i.e. to share and collectively shape a culture |
| Trustability | Capability of a node/user in a network to engage in a common effort with another node/user before knowing how that node/user will behave |
| Decisability | Capability of a node/user in a network to evaluate and decide among strategic alternatives |
| Modelability | Capability of a node/user in a network to understand the cause-effect structure of a system, thus facilitating causal reasoning, categorization and induction |
| Perceptability | Capability of a node/user in a network to pick information to establish and update internal representations of the environment, i.e. to create a perception of the node's/user's external context |

If we assume that the capabilities contribute independently to the total value of a DIN, then we may expect that the value created by each individual capability is proportional to the square of the size of the DIN. This is a simplification, because the capabilities are in fact interrelated. Thus, we get the following model:

$$y_c = k_{c,M} x^2. \tag{1}$$

The size of the network $x$, is usually given by $n$, the amount of members or users of the network. However, thanks to the proportionality of $y_c$ with $x^2$, $x$ may also be expressed in terms of the relative size of the network. Such a transformation only affects the value of the proportionality constant $k_{c,M}$. We use the latter, because our data set from Eurostat provides direct numbers for the fraction of potential members being connected to DINs, and to keep our measurement conditions constant (see Section 6). The maximum of $x$ therefore equals 1.

The limitations of Metcalfe's law have been described by various authors, such as Briscoe et al. (2006) and Briscoe et al. (2009). Metcalfe's law assumes that each user adds equal value to the network, and this is not the case in general. This effect is well described and quantified by Swann (2002). For example, a connection between people communicating with different languages has in principle smaller value than within a single language domain. For large $n$, Briscoe et al. (2006) provided an alternative to Metcalfe's law which states that the value of a network of size $n$ is proportional to $n$ ln ($n$). The term $n$ comes from the fact that there are $n$ members, each drawing ln ($n$) value for the capability. The term ln ($n$) comes from an empirical rule known as Zipf's law that is used to characterize a vast range of real-world phenomena (Zipf, 1949). Zipf's law states that if some large collection of items is ordered by value, then the $m$th ranked item contributes to the total value with about $1/m$ of the value of the first item. So, if an information network has $n$ members, the value for each member is in total proportional to $1 + 1/2 + 1/3 + \cdots + 1/(n-1)$ which approaches ln ($n$):

$$\lim_{n\to\inf} \sum_{m=2}^{n} \frac{1}{m-1} = \ln(n-1) + \gamma, \tag{2}$$

with $\gamma$ equal to the Euler-Mascheroni constant (0.577… ). To use Briscoe's law, we need to know the absolute size $n$ of the DIN, rather than its relative size $x$:

$$n = xI, \tag{3}$$

with $I$ being the potential maximum size of the DIN. We thus get an alternative model for large $n$:

$$y_c = k_{c,B} xI \ln(xI). \tag{4}$$

As $\ln(xI) = \ln(x) + \ln(I)$, we can see from this equation that $y_c$ is not proportional to $x \ln(x)$, which is why we cannot replace $n$ with $x$ in Briscoe's law as we did in Metcalfe's.

## 4. Methodology

### 4.1. Data collection

We proxied the value created by the capabilities of the HF individually and their dependence on the size of the DIN using data from Eurostat. Eurostat, the European Union's official organization to collect statistical data, provides one of the richest data sources about the usage of Information Technology (IT) in enterprises and households. We were allowed to use a significant part of their data set for our research. The data comes in two separate files with a total size of approximately 350 megabytes, which can be obtained at (Eurostat, 2010). By applying data mining techniques, we were able to relate many Eurostat variables more or less directly to our capabilities, and extract numbers representing the size of the relevant DIN.

Data was collected for every single year between 2002 and 2009, and for the following individual countries: Belgium (*BE*), Bulgaria (*BG*), Czech Republic (*CZ*), Denmark (*DK*), Germany (*DE*), Estonia (*EE*), Ireland (*IE*), Greece (*EL*), Spain (*ES*), France (*FR*), Italy (*IT*), Cyprus (*CY*), Latvia (*LV*), Lithuania (*LT*), Luxembourg (*LU*), Hungary (*HU*), Malta (*MT*), Netherland (*NL*), Austria (*AT*), Poland (*PL*), Portugal (*PT*), Romania (*RO*), Slovenia (*SI*), Slovak Republic (*SK*), Finland (*FI*), Sweden (*SE*), United Kingdom (*UK*), Turkey (*TR*), Iceland (*IS*), Norway (*NO*), MK (*MK*), Croatia (*HR*) and Serbian Republic (*RS*). Furthermore, the data points are collected individually for various economic sectors and geographic regions.

### *4.2. Operationalization*

Obviously, Eurostat did not obtain its data with the HF in mind. Therefore, the data does not provide enough empirical variables to cover fully and perfectly each capability of the framework. The empirical variables chosen in this article are limited by what is being measured in the Eurostat surveys, and some can better be considered to be proxies to the HF capabilities than others. Consequently, we have some capabilities that are relatively well operationalized (particularly coordinatibility, selectibility, adoptability, creatibility, normatibility and trustability), and others that are far from optimal (particularly cooperatibility, biddability, decisability, modelability and perceptability). Table 2 provides a summary of the operationalization. Below we explain how we related capabilities to the Eurostat parameters.

Within organizational management, coordination mechanisms are used in various applications, namely in systems of production, logistics and service operations. Therefore, we operationalize the value $y_c$ created by coordinatibility with the fraction of enterprises using systems for managing production, logistics or service operations (Eurostat reference *e_lnkpls*). Selectibility is operationalized with the fraction of enterprises using Internet information search engines (*e_iif*). Adoptability is operationalized with the fraction of individuals that have used the Internet for training and education (*i_iedut*). Creatibility is operationalized with the fraction of enterprises that consider the Internet significant for the development of new products and services (*e_beictnps*). Normatibility is operationalized with the fraction of enterprises that use agreed proprietary standards for automated data exchange (*e_adefpro*). Trustability is operationalized with the fraction of enterprises that regard the improving of the company image as an important reason why they are selling via the Internet (Eurostat references *e_benimv* and *e_benims*).

The capabilities in the previous paragraph are the ones that could be relatively easily related to the Eurostat variables. For the ones below, we feel that the relationship is not so straightforward. The deployment of DINs fueled the rise of electronic commerce, matching the goals of buyers and sellers to cooperate in a supply and demand relationship (Weiss, 2009). Therefore, we operationalize cooperatibility with the fraction of enterprises that have ordered products or services via the Internet (Eurostat reference *e_ibuy*). DINs have lowered costs of organizing bidding auctions, leading to an increasing number of transactions (Lucking-Reiley, 2000). Milgrom (1989) stated that Internet transactions reduce the state space of the negotiation to the bid alone and has the ''additional advantage of being an institution [Internet] where the conduct can be delegated to an unsupervised agent''. We operationalize biddability with the fraction of individuals that have used the Internet for selling goods (e.g. via auctions) (Eurostat reference *e_iusell*).

Modelability, decisability, and perceptability have an intricate and dependent nature. Modeling endows organisms to learn contingencies among events and actions, and therefore it is a vital capability for making decisions in dynamic environments (Newell and Broder, 2008). Moreover, both decisability and modelability are limited by the fact that biological organisms have limitations on how much information can be perceived (Miller, 1956). Thus, approaches have been proposed in the literature to integrate decisability, modelability and perceptability (Hommel et al., 2001). Nevertheless, the HF values their conceptual separation, based upon empirical evidence such as direct parameter specification (Neumann, 1989). Unfortunately, capturing each of these capabilities individually is not possible with the data provided by the Eurostat surveys. Therefore, we operationalize these three capabilities together with the fraction of enterprises that regard Internet sales as very important or of some importance in improving the quality of their services (Eurostat references *e_benquv* and *e_benqus*). Unfortunately, culturability and brokerability do not map at all with any variable from the Eurostat surveys.

Two out of nine operationalizations have been carried out using Eurostat proxies referring to individual Internet usage rather than enterprises. Of course, we would have liked to use usage statistics of either individuals or enterprises, but they were not available for all our capabilities. Besides, the distinction between individuals and enterprises is blurring, as consumers become also producers, employees work more and more from home, ever more companies have only one employee, and work and private activities are often mixed during the day. We therefore deem the distinction between individuals and enterprises not relevant for our study.

It is important to realize that we do not operationalize the full causal chain (DINs → capabilities → economic value), but only the causal chain DINs → capabilities. That is, the numbers we use for $y_c$ do not represent real economic value (e.g. €s), but are proxies for this value. But we took care that our operationalizations of capabilities are on one hand clearly resulting from availability of DINs and on the other hand are closely associated with economic value creation.

To represent DINs and especially their relative size $x$ in terms of users who are interconnected as a fraction of the total population, the Eurostat data provides two empirical variables: the fraction of enterprises that have access to the Internet (Eurostat reference *e_iacc*) and the fraction of households with access to the Internet (Eurostat reference *h_iacc*). Again, we consider both as good proxies for $x$, and when analyzing our data we indeed observed only very

**Table 2**
Operationalization of the capabilities.

| Capability | Operationalization | Eurostat reference |
|---|---|---|
| Coordinatibility | Fraction of enterprises using systems for managing production, logistics or service operations | *e_lnkpls* |
| Cooperatibility | Fraction of enterprises that have ordered products or services via the Internet | *e_ibuy* |
| Selectibility | Fraction of enterprises using Internet information search engines | *e_iif* |
| Biddability | Fraction of individuals that have used Internet for selling goods (e.g. via auctions) | *i_iusell* |
| Adoptability | Fraction of individuals that have used Internet for training and education | *i_iedut* |
| Creatibility | Fraction of enterprises that consider the Internet significant for the development of new products and services | *e_beictnps* |
| Normatibility | Fraction of enterprises that use agreed proprietary standards for automated data exchange | *e_adefpro* |
| Trustability | Fraction of enterprises that regard the improving of the company image as an important reason why they are selling via the Internet | *e_benimv* *e_benims* |
| Decisability, modelability and perceptability | Fraction of enterprises that regard Internet sales as very important or of some importance in improving the quality of their services | *e_benquv* *e_benqus* |

small, negligible difference between those numbers for a given year, country, sector, etc.

Furthermore, the empirical variables use different sample sizes: samples per year, per country, and per economic sector or geographical region. We assume that the resulting fractions are representative for all the domains observed. For example, if 10% of the enterprises in the construction sector, in Portugal, in 2004 used cooperatibility, and 50% of all enterprises in Portugal in 2004 was connected to the Internet, then we assume that we have found a generic data point $(x,y) = (0.5, 0.1)$, irrespective of the fact that it was measured in 2004, in Portugal, in the construction sector, with enterprises. By treating our data as such, we thus basically assume that it does not really matter how big the sample size is (as long as it is big enough), or which country we are in, or which sector, or if we are looking at enterprises or individuals: the value a network node (i.e. an enterprise or individual) creates only depends on the size of the network. This means that we assume linear user utility (Swann, 2002) for the whole Internet. Said otherwise, if 50% of all enterprises in Portugal in 2004 was connected to the Internet, we assume that this indicates that worldwide 50% of all enterprises are connected to the Internet and that they on average have equal interest in communicating with each other. This assumption is plausible, because the results for large aggregation areas such as the whole EU did not differ significantly or structurally from the smaller samples (Eurostat also provides data for aggregated areas such as the European Union – 27 countries (Eurostat reference *EU27*), European Union – 25 countries (*EU25*), European Union – 15 countries (*EU15*), EuroZone – 15 countries (*EA*), and EuroZone + SK – 16 countries (*EA16*)).

Regarding model (4), we estimate $n$ by multiplying $x$ with the number of Internet Protocol (IP) addresses advertised in 2010 ($I \approx 2.2 * 10^9$ (Potaroo, 2010)). Again, we have then assumed that observations for smaller (but still large enough) samples, such as the example of the portuguese construction sector in 2004, are portable to larger aggregation areas. To make our results independent of sample size (see also Section 6.1) we therefore multiply with the largest realistically achievable network size, knowing that most users and enterprises have obtained just 1 public IP address. Furthermore, we assume that the size of the population (i.e. the theoretical maximum size of the network) is the same for every year in the Eurostat data source (2004–2009) and that each IP address acts as a node in the network and adds the same value as all the other addresses.

Note that this study is done on the conceptual level of digital information networks, and therefore it does not distinguish different types of DINs (e.g. POTS, DSL, fiber, mobile, etc.). As such, for our study, it is enough that some form of DIN, even if basic (e.g. Internet over POTS), was already available worldwide in the first year of our study (i.e. 2004). Moreover, note that the previous assumption refers to the availability of the infrastructure, not access to the infrastructure, which is an exogenous variable in our study and varies considerably.

### 4.3. Analysis method

For model (1), the coupling strength $k_{c,M}$ of each capability is estimated by minimizing the sum of squared residuals:

$$\tilde{k}_{c,M} = \frac{\sum_i y_{c,i} x_i^2}{\sum_i x_i^4}, \tag{5}$$

in which $y_{c,i}$ is the operationalization for the value created by a certain capability and $x_i$ is the operationalization for the relative size of the DIN. For model (4), the coupling strength $k_{c,B}$ of each capability follows from:

$$\tilde{k}_{c,B} = \frac{\sum_i y_{c,i} x_i I \ln(x_i I)}{\sum_i x_i^2 I^2 \ln^2(x_i I)}. \tag{6}$$

Our results are presented in a set of graphs such as shown in Fig. 1(b) in which the horizontal axis represents $x$ or $n$, for model (1) and (4) respectively, and the vertical axis represents $y_c$, the normalized proxy for the value created by a capability. The regression line is shown by the thick curve. For optimal representation of the results, a binning process was used due to the large number of available samples and their relatively large spread. For example, we have more than 3000 samples available for coordinatibility and their relative standard deviation is 16% of the expectation value. The bin size $\Delta x$ that we

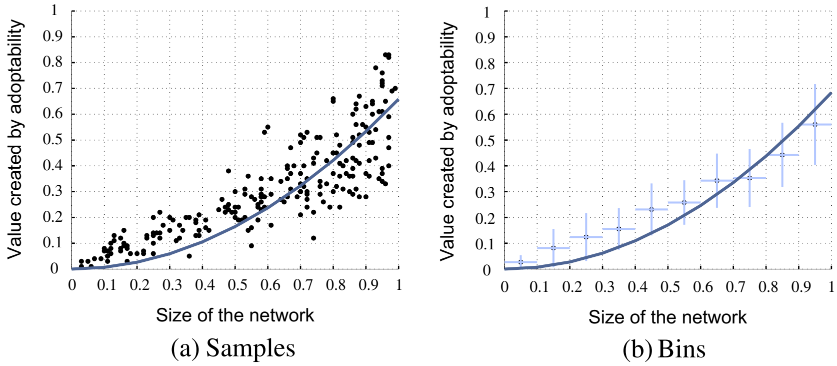


**Fig. 1.** Binning process for adoptability.

used is 0.05 for the regressions with Metcalfe's law, which corresponds to the horizontal error bar in Fig. 1(b). For the fits with Briscoe's law, the horizontal error bar is $\Delta n = 10^8$. The vertical error bar corresponds to the standard deviation of the samples in each bin. Fig. 1 provides an illustration of the binning process taking adoptability as an example. The individual samples are represented with black dots.

## 5. Results

### 5.1. Metcalfe's law

Fig. 2 shows the results obtained with model (1). All curves fit well within the limits provided by the error bars. The exception is selectibility, which behaves linearly with a slope of approximately 1, meaning that roughly every

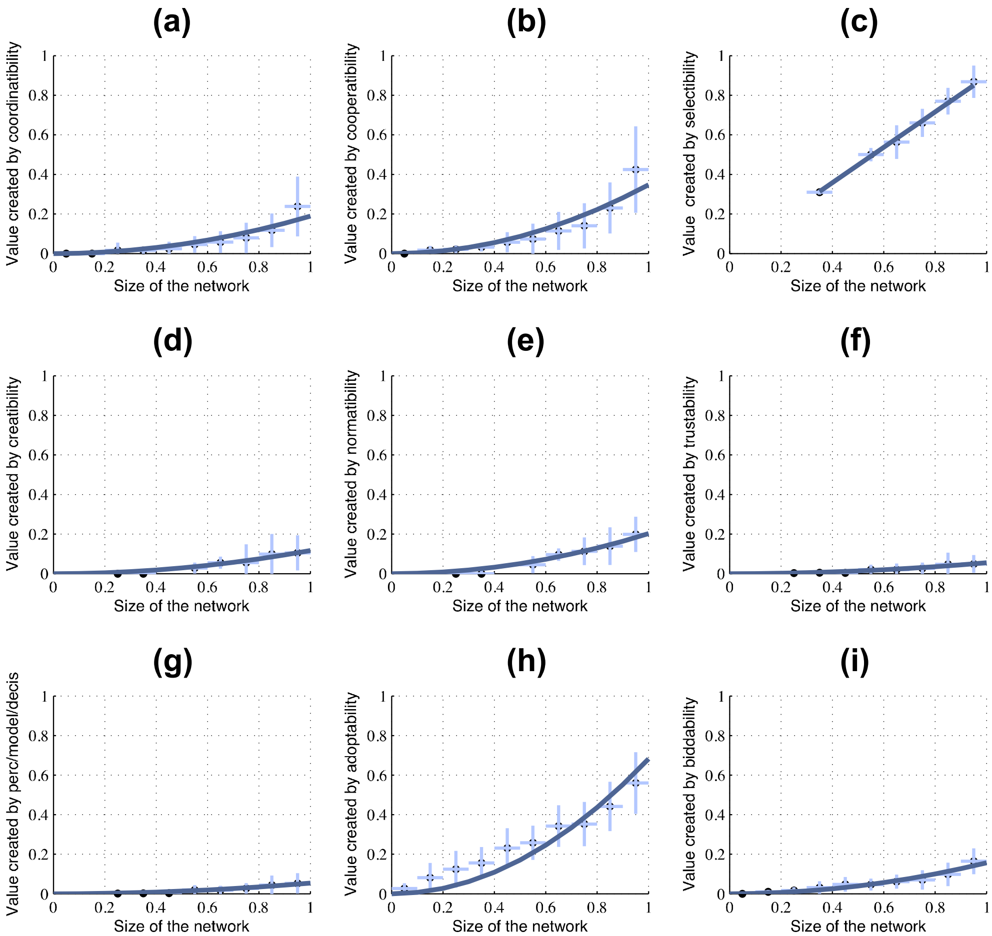


**Fig. 2.** Applying Metcalfe's law to the Eurostat data (___ regression, ___ precision, • bin points).

**Table 3**
Coupling strength ranking of the capabilities with Metcalfe's law. **n**: number of samples, $k_{c,M}$: coupling strength, $\mathbf{RSD}_{k_c,M}$: relative standard deviation of the coupling strength, $R^2$: coefficient of determination of Metcalfe's law.

| | Capability | **n** | $k_{c,M}$ | $\mathbf{RSD}_{k_c,M}$ (%) | $R^2$ |
|---|---|---|---|---|---|
| 1 | Selectibility (fraction of enterprises using Internet information search engines) | – | – | – | – |
| 2 | Adoptability (fraction of individuals that have used Internet for training and education) | 220 | 0.68 ± 0.05 | 7 | 0.85 |
| 3 | Cooperatibility (fraction of enterprises that have ordered products or services via the Internet) | 3635 | 0.35 ± 0.05 | 14 | 0.88 |
| 4 | Normatibility (fraction of enterprises that use agreed proprietary standards for automated data exchange) | 887 | 0.20 ± 0.01 | 5 | 0.96 |
| 5 | Coordinatibility (fraction of enterprises using systems for managing production, logistics or service operations) | 3347 | 0.19 ± 0.03 | 16 | 0.86 |
| 6 | Biddability (fraction of individuals that have used Internet for selling goods (e.g. via auctions)) | 191 | 0.17 ± 0.01 | 6 | 0.93 |
| 7 | Creatibility (fraction of enterprises that consider the Internet significant for the development of new products and services) | 805 | 0.117 ± 0.008 | 7 | 0.94 |
| 8 | Trustability (fraction of enterprises that regard the improving of the company image as an important reason why they are selling via the Internet) | 839 | 0.055 ± 0.004 | 7 | 0.93 |
| 9 | Perc/model/decis (fraction of enterprises that regard Internet sales as very important or of some importance in improving the quality of their services) | 836 | 0.054 ± 0.004 | 7 | 0.94 |

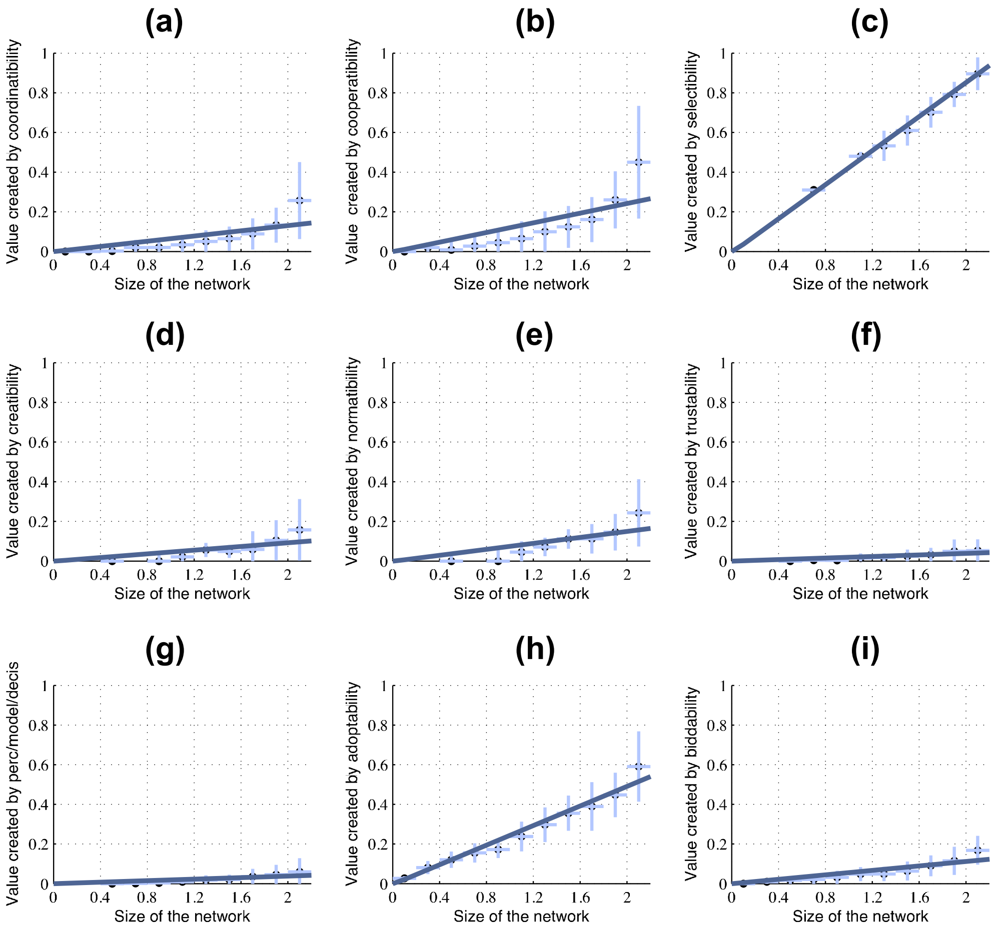


**Fig. 3.** Applying Briscoe's law to the Eurostat data (___ regression, ___ precision, • bin points).

additional node will use selectibility. This can be theoretically expected. When a quadratic curve following model (1) gets close to the line $y = x$, it means that the fraction of enterprises or individuals using a capability is equal to the relative size of the network. This strongly indicates that every enterprise or individual that is connected to the network uses the capability. From there on the curve follows $y = x$. This is in line with our model (4) and the literature (Briscoe et al., 2006) where for large networks, the increase in the value is expected to tend more towards a linear

**Table 4**
Coupling strength ranking of the capabilities with Briscoe's law **n**: number of samples, $k_{c,M}$: coupling strength, $\mathbf{RSD}_{k_c,M}$: relative standard deviation of the coupling strength, $R^2$: coefficient of determination of Briscoe's law.

| | Capability | **n** | $k_{c,M}$ **(E-12)** | $\mathbf{RSD}_{k_c,M}$ (%) | $R^2$ |
|---|---|---|---|---|---|
| 1 | Selectibility (fraction of enterprises using Internet information search engines) | 248 | 19.7 ± 0.2 | 1 | 0.99 |
| 2 | Adoptability (fraction of individuals that have used Internet for training and education) | 220 | 12.4 ± 0.5 | 4 | 0.97 |
| 3 | Cooperatibility (fraction of enterprises that have ordered products or services via the Internet) | 3634 | 6 ± 1 | 17 | 0.67 |
| 4 | Normatibility (fraction of enterprises that use agreed proprietary standards for automated data exchange) | 887 | 3.5 ± 0.5 | 14 | 0.68 |
| 5 | Coordinatibility (fraction of enterprises using systems for managing production, logistics or service operations) | 3346 | 3.0 ± 0.7 | 23 | 0.64 |
| 6 | Biddability (fraction of individuals that have used Internet for selling goods (e.g. via auctions)) | 191 | 2.6 ± 0.3 | 12 | 0.83 |
| 7 | Creatibility (fraction of enterprises that consider the Internet significant for the development of new products and services) | 805 | 2.1 ± 0.4 | 19 | 0.62 |
| 8 | Trustability (fraction of enterprises that regard the improving of the company image as an important reason why they are selling via the Internet) | 839 | 0.9 ± 0.1 | 9 | 0.75 |
| 9 | Perc/model/decis (fraction of enterprises that regard Internet sales as very important or of some importance in improving the quality of their services) | 836 | 0.9 ± 0.1 | 9 | 0.70 |

behavior. Unfortunately, Eurostat does not provide data for the behavior of selectibility for small network sizes. However, we can theoretically expect that the use of selectibility also behaves quadratically with $x$ for small network sizes: if there is only 1 node in the network, there is nothing to search for, and with 2 or 3 nodes only a very little. $dy_c/dx$ therefore equals 0, and the first beginning of the curve can only be modeled with a second-order (or higher) polynomial. We therefore assume that the selectibility curve behaves quadratically until the first few available bin points or earlier, and from thereon linearly. This would lead to an estimation for the coupling strength of selectibility of $k_{c,M} \geqslant 2$, which suggests that selectibility is the most relevant capability, as one would expect. The statistics for model (1) are presented in Table 3.

### 5.2. Briscoe's law

Fig. 3 shows the results obtained with model (4). For each sub-graph, the horizontal axis represents $n$ in billions ($10^9$), and the vertical axis represents $y_c$, the normalized proxy for the value created by a capability. The regression curve fits all the data quite well, including selectibility (see the statistics in Table 4). If we compare the relative standard deviations in $k_{c,B}$ with those of $k_{c,M}$ (Table 3), we observe that Briscoe's law fits the strongly coupled capabilities selectibility and adoptability better than Metcalfe's law. For less strongly coupled capabilities, Metcalfe's law fits better. Briscoe's law also explains better the variance of selectibility and adoptability. For the other capabilities, Metcalfe's law has higher coefficients of determination. These results are in concordance with observations about the validity interval of Metcalfe's law (Briscoe et al., 2006).

## 6. Discussion

### 6.1. Analysis of the models and the results

Both our regression models result in the same ranking of coupling strengths with selectibility on top and perceptability/modelability/decisability at the bottom. Selectibility is followed by adoptability and cooperatibility. Within the error bars, normatibility, coordinatibility, and biddability have the same coupling strength and so does trustability and perceptability/modelability/decisability.

Selectibility and, to a lesser extent, adoptability support the use of Briscoe's law rather than Metcalfe's law. The selectibility curve basically states that everyone who has a DIN connection uses it to select information. This is not surprising given the popularity of Internet search engines. Also the adoptability curve is best fitted with model (4) rather than (1). This is somewhat remarkable since the curve of Fig. 3h has not yet approached $y = n$, but seems to follow $y \approx 0.5n$. Apparently, there is a group of users which do not require adoptability at all, independently of the size of the network. The remaining capabilities seem to be better fitted with Metcalfe's law. Overall, we can safely state that the capabilities have either a quadratic or a linear dependency with the size of the DIN infrastructure.

An important aspect in our methodology is the use of normalized values in model (1) instead of absolute values. This has the advantage of keeping the measurement conditions constant. If instead the model would have been of the form:

$$y_c * P = q_{c,M}(x * P)^2, \qquad (7)$$

with $P$ being the sample size, then expression (5) for $q_{c,M}$ would be dependent on the sample size and thus be meaningless. In model (4), we must use absolute values due to the ln function present. Instead of sample size we chose for $I$, the size of the global Internet, a number which is equal for every sample and which varies negligibly over the years. Given our general assumption of linear user utility (Swann, 2002) for the whole Internet, this is a reasonable approach. We could have used such absolute values for model (5) also, but that would only lead to a correction of all $q_{c,M}$ with a factor $I$.

### 6.2. Implications for theory and practice

In the mainstream literature, Metcalfe's law has been used more as a heuristic or metaphor than an iron-clad empirical rule. To our knowledge, empirical work validating and employing Metcalfe's law was nonexistent up to now. Therefore, this article is most likely the first empirical study which supports the implications of Metcalfe's law and its extension mentioned in Briscoe et al. (2006) concerning large networks.

Our work opens the possibility of using mediation analysis techniques for the study of large scale economic impacts of DINs. A mediation model is one that seeks to identify and explicate the mechanism that underlies an observed relationship between an independent variable and a dependent variable via the inclusion of a third explanatory variable, known as a mediator variable (Baron and Kenny, 1986). Extensions of our work could rely on the capabilities as mediator variables and their behavior as investigated in this article. The current literature on mediation analysis only applies to specific and small-scale impacts of general IT (e.g. Grover et al., 1998).

Our study also helps relating small-scale studies on specific impacts of DINs with the macro-level studies reviewed in Section 2. The capabilities of the HF, the coupling strengths $k$, and the functional forms provided by Metcalfe's and Briscoe's law can be used to extrapolate results from the micro- to the macro-level. For example, one may apply a $k$ obtained from this article to a specific country for a nation-wide study. Additionally, they can be used to validate impact changes at the macro-level, because these changes necessarily need to be preceded by changes in the use of capabilities.

The macro-economic studies reviewed in Section 2 rely on Cobb-Douglas production functions to model an economic system as a black box and investigate the relationship between DINs as a production input and economic value as an output (e.g. see Majumdar et al., 2010 and Koutroumpis, 2009). Generally speaking, Cobb-Douglas functions take the form:

$$Y = AL^{\alpha}K^{\beta}, \qquad (8)$$

in which $Y$ is the total production, $A$ is the total factor productivity, $L$ is the labor input, $K$ is the capital input, and $\alpha$ and $\beta$ are the output elasticities of labor and capital respectively. $\alpha$ and $\beta$ are assumed to be constants determined by the available production technology, such as DINs. However, our results show that the value created by capabilities are linearly or quadratically dependent on the size of the DIN infrastructure, including a transition region between linear and quadratic behaviors. Such complex behaviors might not be well captured with constant output elasticities. Thus, studies on the economic impact of DINs need to introduce more complex production functions, such as functions with variable returns to scale (Kim, 1992).

Interoperability refers to the ability of two or more systems or components to exchange information and to use the information that has been exchanged (IEEE, 1991). Interoperability is related to how information is conceptualized. Most of the existing research discusses interoperability from a purely technical perspective, assuming that information can be directly observed as a production input (e.g. data such as an ontology) (Legner and Lebreton, 2007). We hypothesize that each capability corresponds to a fundamental dimension of interoperability. For example, interoperability on the dimension of coordinatibility may correspond to interoperability between IT systems for supply-chain management such as choreography and synchronization of activities and milestones between business partners (McAfee, 2005). Second, the empirical results presented in this article may provide propositions to test why investments in IT integration did not result in the expected pay-offs. For example, the returns on investment of an IT system to enable interoperability of a certain capability did not occur because the coupling strength of the capability was small.

### 6.3. Limitations and future study

Causality in social and, more generically, holarchical systems is problematic to establish. Nowhere in this article, we claim to have established causality, but only to provide further insights on the co-occurrence of DINs and creation of economic value. The state-of-the-art literature does not advance the mechanisms from DINs to economic value, i.e. the capabilities. In Madureira et al. (2011), we have introduced the HF together with its set of capabilities. Regarding causality, this article simply evidences that the use of the capabilities grows (correlates) quadratically with the size of the DIN infrastructure. Taking selectibility as an example, we assume that enterprises are using Internet search engines, because they have direct returns on value from that use. One could also think of a reverse-causality story where value creation leads to higher use of Internet search engines. However, the first hypothetical causal logic seems to us stronger than the reverse, which would require two conditions: (1) the correlation observed is mainly a result of highly valued enterprises that decide to use Internet information search engines; (2) the use of Internet information search engines does not bring significant value for the enterprises. Nevertheless, both causal logics play a role in the correlations observed. The questions therefore is: which one is stronger and how much relatively to the other? Future work relying on the existing statistical data, but especially in experimental setups, could attempt to clarify this question. Given the added value and complexity of such task, it is a research topic *per se*, and is therefore left for future work.

In this article, we did not operationalize the full causal chain (DINs → capabilities → economic value), but only the causal chain DINs→ capabilities. That is, the numbers we use for $y_c$ do not represent real economic value in €s, but are proxies for this value. But we took care that our operationalizations of capabilities are on one hand clearly resulting from availability of DINs and on the other hand are closely associated with economic value creation. For future work, we suggest to operationalize the full causal chain using mediation analysis. Following this approach, the capabilities could serve as intermediate variables between DINs and one single notion of value (e.g. GDP). Note that with the current Eurostat data, this task is not possible, because we do not possess operationalizations for all the capabilities, and a few operationalizations are limited. Nevertheless, we expect that, independently of the notion of value taken, Metcalfe's law and its derivative would still hold.

We realize that some operationalizations are far-stretched. But most are acceptable (e.g. coordinatibility, selectibility, adoptability and creatibility). The main contribution of our article is to show how Metcalfe's law can be used to quantify the relation between the size of the DIN infrastructure with the use given to the capabilities. Having this goal in mind, we feel that the quality of our operationalization can be improved, but is sufficient for now. For the more ambitious goal of operationalizing all the causal chain (DINs → capabilities → economic value) for predictive purposes or for empirically testing the completeness of the HF, the operationalization should indeed be improved. Looking into other data sources is an obvious way to improve our empirical results (e.g. the United Nations (UN) Statistical Commission). Even better would be the understanding and construction of a targeted measuring and data-gathering campaign to further validate and quantify the importance and completeness of the capabilities identified by the HF. In this sense, our work can also help statistical agencies to redefine the data to be collected.

Mapping the Eurostat data with the capabilities of the HF was challenging. We assumed that this was due to the measurement limitations of the Eurostat data. Nevertheless, we should not exclude the possibility that the definitions of the capabilities need better formalizing, and the HF as a whole requires further conceptual development.

In this article, network aspects are only accounted for through the metric *size of the network*, because it is the only network metric available in the Eurostat data. For example, in Fig. 1, we basically present in the $x$-axis the size of the DIN infrastructure, and in the $y$-axis the size of the adoptability network. However, other network metrics may capture better the effect of the capabilities. For example, future work could investigate if topological network metrics such as betweenness, centrality or assortativity do a better job on capturing the effects of capabilities on value creation.

Finally, our study was done on the conceptual level of digital information networks, and therefore it does not distinguish different types of DINs (e.g. POTS, DSL, fiber, mobile, etc.). Future work could also attempt to distinguish the quality of the DIN infrastructure, even if roughly (e.g. narrowband and broadband).

## 7. Conclusions

To justify further investments in Digital Information Networks (DINs)' infrastructures (e.g. in FTTH), it is necessary to analyze expenditures that have already been made and demonstrate their value. Madureira et al. (2011) presented a Holonic Framework (HF) which identifies these mechanisms as *capabilities* and specified 13 of these capabilities. Building upon the HF and Eurostat data, this article demonstrates that the value that these capabilities create by using information shows either a quadratic or a linear dependency with the size of the digital information network infrastructure. The quadratic dependency can be explained by Metcalfe's law. The linear dependency is explained by an extension of Metcalfe's law as described in Briscoe et al. (2006). We were able to quantify the economic coupling strength of the capabilities and showed that the results are qualitatively the same irrespectively of using Metcalfe's law or Briscoe's adaptation of it. To our knowledge, this is the first time that Metcalfe's law is empirically validated in a scientific way.

## Acknowledgments

We thank the Royal Dutch KPN, TNO and the Delft University of Technology for funding this research. We thank Eurostat for the data used in this article. Finally, we thank the editors and reviewers of Information Economics and Policy for constructive discussions during the submission process.

## References

ACIL Tasman, 2004. Economic Impacts of Broadband Adoption in Victoria. Technical report, ACIL Tasman.

Baron, R., Kenny, D., 1986. The moderatormediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology 51 (6), 1173–1182.

Borges, A., 1986. Applied general equilibrium models: an assessment of their usefulness for policy analysis. Technical report 7, Organisation for Economic Co-operation and Development.

Briscoe, B., Odlyzko, A., Tilly, B., 2006. Metcalfe's Law is wrong. IEEE Spectrum, 26–31.

Briscoe, B., Odlyzko, A., Tilly, B., 2009. When a law isn't a law at all. International Commerce Review 8 (2–4), 146–149.

CEBR, 2003. The Economic Impact of a Competitive Market for Broadband. Technical report, The Broadband Industry Group.

Correa, L., 2006. The economic impact of telecommunications diffusion on UK productivity growth. Information Economics and Policy 18 (4), 385–404.

Crandall, R., Lehr, W., Litan, R., 2007. The Effects of Broadband Deployment on Ontput and Employment: A Cross-sectional Analysis of US Data. Brookings Institution Press.

Cronin, F., Parker, E., Colleran, E., Gold, M., 1991. Telecommunications infrastructure and economic growth: an analysis of causality. Telecommunications Policy 15 (6), 529–535.

Datta, A., Agarwal, S., 2004. Telecommunications and economic growth: a panel data approach. Applied Economics 36 (15), 1649–1654.

Dijk, M.V., Mulder, M., 2005. Regulation of telecommunication and deployment of broadband. CPB Memoranda 131, CPB Netherlands Bureau for Economic Policy Analysis.

Duggal, V., Saltzman, C., Klein, L., 2007. Infrastructure and productivity: an extension to private infrastructure and it productivity. Journal of Econometrics 140 (2), 485–502.

Eurostat, 2010. ICT usage and e-Commerce. <http://epp.eurostat.ec.europa.eu/portal/page/portal/information_society/introduction> (accessed 22.11.10).

Farmer, J., Foley, D., 2009. The economy needs agent-based modelling. Nature 460 (6), 685–686.

Firth, L., Mellor, D., 2005. Broadband: benefits and problems. Telecommunications Policy 29 (2–3), 223–236.

Ford, G., Koutsky, T., 2005. Broadband and economic development: a municipal case study from Florida. Applied Economic Studies 17 (3), 216–229.

Freeman, C., Louçã, F., 2001. As Time Goes by: From the Industrial Revolutions to the Information Revolution. Oxford University Press.

Greenstein, S., McDevitt, R., 2009. The Broadband Bonus: Accounting for Broadband Internet's Impact on US GDP. NBER Working Papers 14758, National Bureau of Economic Research, Inc.

Greenstein, S., Spiller, P., 1995. Modern telecommunications infrastructure and economic activity: an empirical investigation. Industrial and Corporate Change 4 (4), 647–665.

Grover, V., Teng, J., Segars, A., Fiedler, K., 1998. The influence of information technology diffusion and business process change on

perceived productivity: the IS executive's perspective. Information Management 34 (3), 141–159.
Hardy, A., 1980. The role of the telephone in economic development. Telecommunications Policy 4 (4), 278–286.
Hommel, B., Musseler, J., Aschersleben, G., Prinz, W., 2001. The theory of event coding (TEC): a framework for perception and action planning. Behavioral and Brain Sciences 24 (5), 849–878.
IEEE, 1991. IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. IEEE Computer Society Press.
Kagami, M., Tsuji, M., Giovannetti, E., 2004. Information Technology Policy and the Digital Divide: Lessons for Developing Countries. Edward Elgar Publishing.
Katz, R., Vaterlaus, S., Zenhausern, P., Suter, S., Mahler, P., 2009. The Impact of Broadband on Jobs and the German Economy. Technical report, Deutsche Telekom AG.
Kim, Y., 1992. The translog production function and variable returns to scale. The Review of Economics and Statistics 74 (3), 546–552.
Koutroumpis, P., 2009. The economic impact of broadband on growth: a simultaneous approach. Telecommunications Policy 33 (9), 471–485.
Leff, N., 1984. Externalities, information costs, and social benefit-cost analysis for economic development: an example from telecommunications. Economic Development and Cultural Change 32 (2), 255.
Legner, C., Lebreton, B., 2007. Preface to the focus theme section: 'business interoperability' business interoperability research: present achievements and upcoming challenges. Electronic Markets 17 (3), 176–186.
Lehr, W., Gillett, S., Osorio, C., Sirbu, M., 2006. Measuring broadband's economic impact. Broadband Properties, 12–24.
Lucking-Reiley, D., 2000. Auctions on the Internet: what's being auctioned, and how? Journal of Industrial Economics 48 (3), 227–252.
Madden, G., Savage, S., 1998. CEE telecommunications investment and economic growth. Information Economics and Policy 10 (2), 173–195.
Madden, G., Savage, S., 2000. Telecommunications and economic growth. International Journal of Social Economics 27 (7/8/9/10), 893–906.
Madureira, A., 2011. Value of Digital Information Networks: A Holonic Framework. PhD thesis.
Madureira, A., Baken, N., Bouwman, H., 2011. Value of digital information networks: a holonic framework. Netnomics 12 (1), 1–30.
Majumdar, S., Carare, O., Chang, H., 2010. Broadband adoption and firm productivity: evaluating the benefits of general purpose technology. Industrial and Corporate Change 19 (3), 641–674.
McAfee, A., 2005. Will web services really transform collaboration? Sloan Management Review 46 (2), 78–84.
Metcalfe, R., 1995. Metcalfe's law: a network becomes more valuable as it reaches more users. Infoworld (17).
Milgrom, P., 1989. Auctions and bidding: a primer. Journal of Economic Perspectives 3 (3), 3–22.
Miller, G., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review 63, 81–97.
Neumann, O., 1989. Kognitive Vermittlung und direkte Parameter-spezifikation. Zum Problem mentaler Representation in der Wahrnehmung, Sprache und Kognition 8, 32–49.
Newell, B., Broder, A., 2008. Cognitive processes, models and metaphors in decision research. Judgment and Decision Making 3, 195–204.
OECD, 2001. The development of broadband access in OECD countries. Technical report, Organisation for Economic Co-operation and Development.
Perez, C., 2003. Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages. Edward Elgar Pub..
Potaroo, 2010. IPv4 Address Report. <http://www.potaroo.net/tools/ipv4/index.html> (accessed 22.11.10).
Röller, L., Waverman, L., 2001. Telecommunications infrastructure and economic development: a simultaneous approach. The American Economic Review 91 (4), 909–923.
Shideler, D., Badasyan, N., Taylor, L., 2007. The economic impact of broadband deployment in Kentucky. Regional Economic Development 3 (2), 88–118.
Shiu, A., Lam, P., 2008. Causal relationship between telecommunications and economic growth in China and its regions. Regional Studies 42 (5), 705–718.
Sridhar, K., Sridhar, V., 2004. Telecommunications infrastructure and economic growth: evidence from developing countries. Working Papers 04/14, National Institute of Public Finance and Policy.
Swann, P., 2002. The functional form of network effects. Information Economics and Policy 14, 417–429.
Thompson, H., Garbacz, C., 2007. Mobile, fixed line and Internet service effects on global productive efficiency. Information Economics and Policy 19 (2), 189–214.
Thompson, H., Garbacz, C., 2008. Broadband impacts on state GDP: direct and indirects impacts. In: International Telecommunications Society 17th Biennial Conference, Canada.
Weiss, A., 2009. Smart infrastructure matches supply and demand. netWorker 13 (3), 18–25.
Zipf, G., 1949. Human Behavior and the Principle of Least Effort. Addison-Wesley.