Discussion Paper No. 2001/11

Cross-country Diffusion of the Internet

Sampsa Kiiski[1] and Matti Pohjola[2]

June 2001

Abstract

This paper investigates the factors which determine the diffusion of the Internet across countries. The Gompertz model of technology diffusion is estimated using data on Internet hosts per capita for the years 1995-2000. For a sample of the OECD countries, the basic finding is that GDP per capita and Internet access cost explain best the observed growth in computer hosts per capita. Competition in telecommunications markets does not seem to exert any independent influence on Internet penetration. Neither is investment in education a statistically significant predictor of its diffusion. For a larger sample of both industrial and developing countries, the results change in such a way that also education becomes significant.

Keywords: diffusion of technology, Internet, Internet access cost, Gompertz model, technology adoption, technology diffusion

JEL classification: L5, L9, O1, O3

[1] Nokia Corporation; [2] UNU/WIDER

This study has been prepared within the UNU/WIDER project on Production, Employment and Income Distribution in the Global Digital Economy, which is directed by Professor Matti Pohjola.

UNU/WIDER gratefully acknowledges the financial contribution to the project by the Ministry for Foreign Affairs of Finland.

Acknowledgements

The authors are grateful for helpful comments to Sumit Majumdar and the other participants at the UNU/WIDER project meeting on The New Economy in a Global Perspective.

1 Introduction

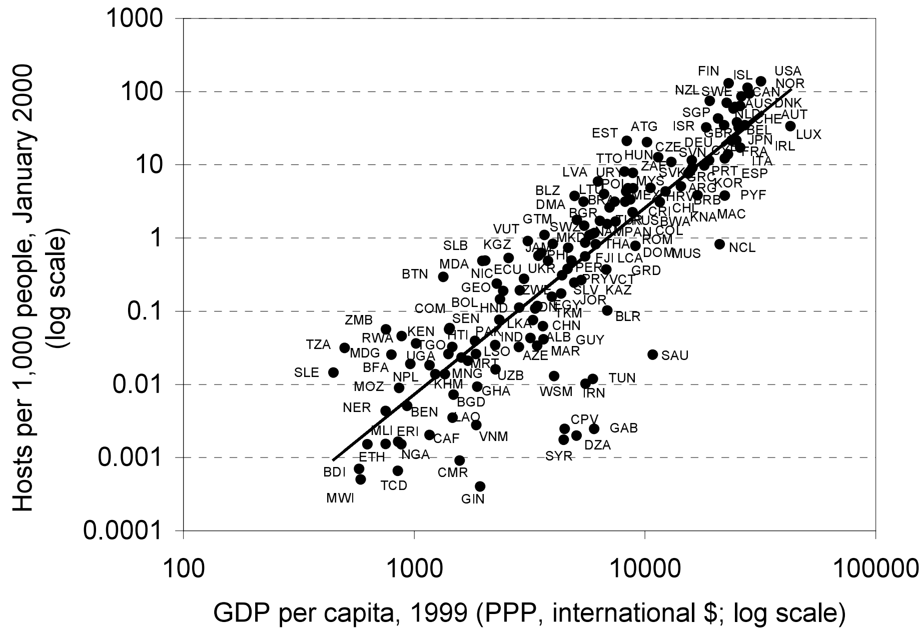
The computer and the Internet are at the vanguard of the modern information technology revolution. The Internet, being a network of computer networks, allows everyone with a PC and a modem to communicate with all the other computers connected to it without regard for geographic location. It is at once a world-wide broadcasting capability, a mechanism for information dissemination, a medium for interaction between individuals, and a marketplace for goods and services. A recent survey of the Internet domain name system indicates that in January 2001 there were 110 million computer hosts on the Internet (Internet Software Consortium 2001). A host is a computer that has users who access network services through it. The growth in the number of Internet hosts has been spectacular as in January 1991 their number was only 376,000. Internet traffic is estimated to double every 100 days (US Department of Commerce 1998).

However, stark differences exist between countries in the accessibility to the Internet. Rich industrial countries lead the rankings and, as Figure 1 displays, there exists a strong correlation across countries between GDP per capita and Internet connectivity. But differential access to the Internet technology is even more pronounced than national income inequalities across the world. In 1997, the fifth of the world's people living in the highest income countries had 86 per cent of world GDP but 93 per cent of Internet users, whereas the bottom fifth had 1 per cent of GDP and 0.2 per cent of the people 'online' (UNDP 1999). Given that information technology is believed to be a new source of economic growth, there is indeed concern that the Internet is becoming a factor which contributes to the widening of income differentials between countries.

To be able to bridge this 'digital divide' or to eradicate 'information poverty', it is important for policy-makers to know what factors explain the differences across countries in Internet connectivity. Providing universal access may not be enough if people cannot afford the services or are not capable of utilizing them. Surprisingly few studies have been conducted on these issues. The only econometric studies we are aware of are Hargittai's (1999) analysis of Internet connectivity in 18 OECD countries and Norris' (2000) investigation of Internet penetration in a larger group of 179 countries. Hargittai regresses across OECD countries the number of Internet hosts per capita in 1998 on the average values in 1994-96 of a number of explanatory variables measuring income and educational levels as well as telecommunications policy and infrastructure. Her basic conclusion is that it is the regulatory environment and its influence on competition that have the largest impact on Internet penetration: competition between telecom service providers increases Internet connectivity. GDP per capita turned out be the other significant explanatory factor.

For a cross-section of 179 nations, Norris (2000) regresses the number of people online in spring 2000 on a number of variables measuring economic, social and political development. Besides regional dummies, only GDP per capita and the share of R&D spending in GNP turned out to be statistically significant. Neither adult literacy, secondary education nor the level of democratization had any explanatory power. Consequently, economic factors outweighed all others in predicting cross-country differences in access to the Internet. Regional dummies for the Nordic countries, North America and Western Europe were also significant which Norris takes to suggest that there are social and cultural factors that are not being picked up in the model.

Figure 1
Internet host density and GDP per capita in 145 countries, 1999-2000



Data source: Matrix.Net, Inc. for Internet hosts; World Bank for GDP

In this paper we follow a similar approach but extend the previous analyses in an important way. We study the diffusion of the Internet by the aid of an explicit model of technology diffusion. Besides for the OECD countries, we also carry out the estimation for larger samples of countries. Our aim is to find out what the global pattern of Internet diffusion looks like and what factors determine technological diffusion across countries.

2. The model

The diffusion of the Internet will be analysed by the aid of a model which is similar to the ones that have been applied in modelling the spread of the computer (Stoneman 1983: Ch 10). Let H_{it} denote the number of Internet hosts per capita in country *i* in year *t* and let H_i^* be its post-diffusion or equilibrium level. Most models of technology adoption assume that over time H_{it} tends to H_i^* along an S-shaped path. The Gompertz model of diffusion specifies the rate of change in the number of Internet hosts as

$$(1) \qquad \ln H_{it} - \ln H_{it-1} = \alpha_i (\ln H_i^* - \ln H_{it-1})$$

in which α_i is the speed of adjustment taken to be constant in our analysis.

The equilibrium level of Internet hosts per capita will be a function of at least the basic demand-side variables, i.e. the level of income Y_{it} and the cost of Internet access P_{it} in country *i*. Given that income and prices change over time, we may argue that H_i^* is time-dependent, and express it as

(2) $$\ln H_{it}^{*} = \beta_{i0} + \beta_{i1} \ln Y_{it} + \beta_{i2} \ln P_{it} + \boldsymbol{\gamma}' \mathbf{Z}_{it}$$

where $\mathbf{Z}_{it}$ is the vector of other possible variables describing the demand or supply conditions or the technological infrastructure in country *i*. The estimable equation is obtained by inserting (2) into (1):

(3) $$\ln H_{it} - \ln H_{it-1} = \alpha_i \beta_{i0} + \alpha_i \beta_{i1} \ln Y_{it} + \alpha_i \beta_{i2} \ln P_{it} + \alpha_i \boldsymbol{\gamma}' \mathbf{Z}_{it} - \alpha_i \ln H_{it-1} .$$

The computer host data which we will use in our analysis are based on an Internet host domain survey produced every six months by Network Wizards (www.nw.com). These data have been postprocessed by Matrix.Net (www.mids.org) to take into account the physical location of all the com, net, edu and org domains. The dependent variable H_{it} is the number of Internet host per capita in January 2000. Its initial value H_{it-1} is the corresponding per capita host count in January 1995. All other data and their sources will be explained below.

3 Internet diffusion in the OECD countries

Since data on Internet access costs are only available for the OECD countries, we first report the findings for this group of industrial economies, excluding those whose population is less than one million (Iceland and Luxembourg) and those for which access cost data are not available for the period of our analysis (Czech Republic, Hungary, South Korea and Poland). Table 1 presents the results of linear least squares estimation of the Gompertz equation (3) across the remaining 23 OECD countries. The underlying assumption here is that the diffusion process is the same in all countries, i.e. that the parameter values of equation (3) take the same value for all *i*.

The first column displays the estimation results for the basic model in which the explanatory variables are GDP per capita and Internet access cost, both measured in purchasing power parities. OECD's (1996, 1999) estimates of the average cost to users to access the Internet, for 20 hours per month at peak rates, are used as the indicator of Internet prices. This measure includes both the charges of an Internet service provider as well as the fixed charge comprising the monthly rental and local call charges. GDP per capita is taken as the average over the period 1995-99 but, because of missing data, the Internet access cost is calculated as the average of the values for the years 1995 and 1998 only.

Both explanatory variables are statistically significant and have the expected signs. But to make GDP per capita significant, a dummy variable for the low-income countries of Mexico and Turkey had to be included in the regression. The estimates imply that the elasticity of the number of computer hosts per capita with respect to the Internet access cost is -0.98 in the equilibrium, i.e. when the saturation level has been reached. Consequently, a 50 per cent reduction in Internet access cost would raise the equilibrium number of computer hosts per capita by 49 per cent. Over a five-year period, their number would increase by about 25 per cent. This relationship is illustrated in Figure 2 where the growth in the number of hosts per capita, shown on the vertical axis, is corrected for the effect of all the other explanatory variables.

Table 1
Cross-section estimation results for OECD countries

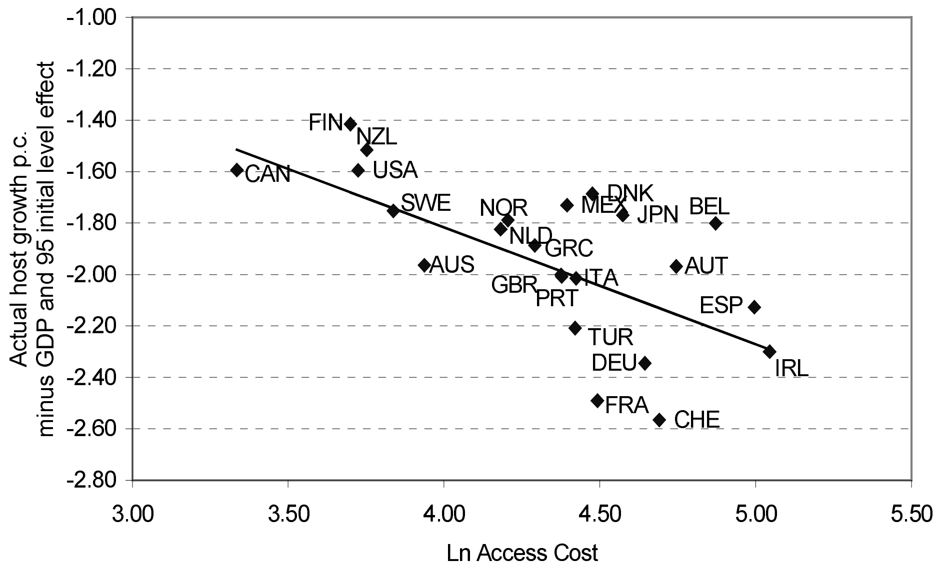
Dependent variable: $\ln H_{00} - \ln H_{95}$					
	(1)	(2)	(3)	(4)	(5)
Speed of diffusion α	0.468*** (0.079)	0.472*** (0.080)	0.502*** (0.091)	0.250** (0.090)	0.667*** (0.092)
Constant $\alpha\beta_0$	-5.023 (3.729)	-5.312 (3.793)	-4.800 (3.782)	3.091 (4.384)	-1.045 (2.470)
GDP per capita $\alpha\beta_1$	0.997** (0.397)	0.982** (0.403)	0.884* (0.428)	0.586 (0.348)	0.683** (0.282)
Internet access cost $\alpha\beta_2$	-0.459** (-0.160)	-0.368* (0.201)	-0.430** (0.166)	-0.249* (0.135)	-0.587*** (0.145)
Telecom competition		0.126 (0.166)			
Average years of schooling			0.366 (0.450)		
English proficiency				-1.139* (0.560)	
Nordic country dummy					0.278* (0.142)
'Southern' country dummy					-0.487** (0.175)
Dummy for Mexico and Turkey	0.790* (0.391)	0.799* (0.396)	0.679 (0.421)	N/A	
F-test	26.18***	20.57***	20.60***	8.74**	27.70***
Adjusted R^2	0.821	0.816	0.817	0.659	0.859
Number of observations	23	23	23	17	23

Note: Standard errors are in parentheses. * = significant at 10 per cent, ** = at 5 per cent, and *** = at 1 per cent level.

For the average values of the explanatory variables in the sample, the saturation level of computer hosts H^* is estimated at 480 hosts per 1,000 people. The speed of adjustment $\alpha = 0.468$ suggests that the number of hosts will increase from $H / H^* = 0.1$ to $H / H^* = 0.9$ in about 33 years.

Regressions (2)-(4) in Table 1 test the robustness of our basic finding that GDP per capita and Internet access cost explain best the observed growth in computer hosts per capita. It is shown in column (2) that the statistical fit is not improved by including in the model a dummy variable describing the existence of some form of competition (i.e. the lack of monopoly) in telecommunications markets. The data for this variable come from OECD (1996). This conclusion is in sharp contrast with Hargittai's (1999) cross-country analysis. She found that the existence of a monopoly *per se* in the telecommunications sector has a considerable negative impact on Internet connectivity in OECD countries whereas access cost is not a statistically significant predictor of computer hosts per capita. However, our result seems to be better in line with both economic theory and intuition: why would competition matter irrespective of its impacts through market prices?

Figure 2
The scatter diagram for computer host growth and Internet access cost



Columns (3) and (4) of Table 1 extend the basic model by including in it variables measuring educational levels. The first measure is the average number of years of schooling for the population over age 15. The observations, obtained from Barro and Lee (2000), are for the year 1995, and the variable enters the model in the logarithmic form. The level of education may have an impact on the diffusion of the Internet for at least two reasons. First, education contributes directly towards basic literacy, and reading and writing skills are essential in using this technology. More educated people are also likely to be faster to adopt new innovations than people with less education. Second, given that the early users of the Internet were people working in higher education and research, academic institutions may play an important role in spreading the Internet. They are often among the first institutions in a nation to be wired.

However, column (3) reveals that education is not a statistically significant predictor of the diffusion of the Internet among the OECD countries. It can be argued that average schooling years are not the best measure of educational attainment for these countries, but our experiments with other variables, such as the percentage of higher education attained in the population, were not any more successful.

English proficiency is the second component of education which is often thought to influence the diffusion of the Internet (see, e.g., the discussion in Norris (2000)). The United States is the birthplace of this technology, and even today over one-half of the world's 110 million computers connected to the Internet are located there. Also about one-half of all the people using this media regularly reside in Northern America. A recent survey conducted by Inktomi Corp. (2000) found that 87 per cent of all the one billion Web pages available on the Internet are in English. Therefore, it is somewhat surprising to see from column (4) of Table 1 that English proficiency enters the regression with the wrong sign. English skills are here measured by the percentage of pupils in secondary education learning English (European Commission 2000). The problem here is that the lack of data which reduces the number of observations to 17. However, our conclusion that education does not seem to have any predictive power in explaining Internet diffusion is consistent with the findings of both Hargittai (1999) and Norris (2000).

The statistical fit of the basic model displayed in column (1) can be improved somewhat by the aid of regional dummy variables. For example, as shown in column (5), by introducing a dummy variable for the Nordic countries (Denmark, Finland, Norway and Sweden) and another dummy for the 'southern' countries (France, Greece, Italy, Mexico, Portugal, Spain and Turkey) the adjusted R^2 improves from 0.82 to 0.86 and the parameters are more precisely estimated. This suggests that there are cultural or technology policy related differences between the countries in the adoption of the Internet which are hard to capture in the model. But the problem with the cross-section analysis is that it assumes the diffusion process to be the same in all countries.

A panel data estimation would allow for more flexibility in this regard, but OECD's Internet access cost data are not suitable for this work. The methodology of estimating the costs has changed over the years but the time series has not been harmonized. We tried panel estimation by replacing Internet access costs with telephone tariffs obtained from ITU (2000). The results are not, however, worth reporting here in detail because telephone tariffs were not statistically significant. Telephone access costs do not seem to explain the differences across the OECD countries in Internet connectivity.

A non-linear least squares estimation of equation (3) slightly improves the precision of the estimates displayed in Table 1. But since the speed of adjustment is already quite accurately identified in the linear estimation, this will not change the qualitative conclusions. To retain comparability with the panel estimations to be reported on in the next section, we have here presented the results of the linear estimation only.

4 Internet diffusion across the world

Next the cross-country analysis will be extended to cover non-OECD countries as well. The sample is limited to countries where the size of the population exceeds one million and where the number of Internet hosts was larger than 50 in 1995. The latter restriction is imposed to avoid problems related to the modelling of the initial adoption of technology. It can be argued that this decision is influenced by factors other than the basic demand and supply conditions only. Since no information is available on the cost of Internet access for this larger group of countries, we have to be contented with telephone tariffs. From the data provided by ITU (2000) we have constructed a variable which measures telephone access costs, in purchasing power parities, by adding together monthly telephone subscription rates and the costs of making 30 three-minute local calls.

Table 2 displays the linear least squares estimation results for the basic model of equation (3) for two sets of data. Columns (1) and (2) contain the results of the cross-section estimations. The dependent variable is the logarithmic difference of Internet hosts per capita over the period 1995-2000 and the explanatory variables assume their average values in this period. Columns (3) and (4) present the results of the panel estimations in which annual values are used for both the dependent and explanatory variables.

The lesson from both the cross-section and panel estimation is that GDP per capita and the average years of schooling are statistically significant explanatory factors of the growth of Internet hosts per capita but that telephone tariffs are not. It is interesting that investment in education now seems to matter for the diffusion of the Internet. This finding is different from the conclusion obtained for the OECD countries, but it is quite intuitive and can be explained by the larger variability of the schooling variable in this sample of countries. It is also consistent with Caselli and Goldman's (2001) results for the diffusion of computer investment across countries.

It can, of course, be argued that education and telephone access costs are not the only infrastructure variables which are important for Internet connectivity. After all, the Internet is not a single innovation but might be better viewed as a cluster of related technologies. The use of the Internet is contingent on the adoption of the computer or some other terminal for access. Another necessary requirement is a means of connection, most commonly the existing network of telephone cables. Data on more specific infrastructure indicators, such as the share of fibre-optic cable in the telephone network or modem density, are unfortunately not available for a representative sample of countries.

Table 2
Cross-section and panel estimation results for countries with more than 50 Internet hosts

	Cross-section estimation		Panel estimation	
Dependent variable:	$\ln H_{00} - \ln H_{95}$		$\ln H_{t+1} - \ln H_t$	
	(1)	(2)	(3)	(4)
Speed of diffusion α	0.507***	0.527***	0.118***	0.147***
	(0.062)	(0.084)	(0.018)	(0.020)
Constant $\alpha\beta$	-7.578**	-6.055**	-1.570***	-1.936***
	(2.314)	(2.839)	(0.482)	(0.560)
GDP per capita $\alpha\beta_1$	1.085***	0.545*	0.229***	0.210**
	(0.254)	(0.325)	(0.055)	(0.073)
Telephone access cost $\alpha\beta_2$	-0.035	0.211	0.037	0.045
	(0.142)	(0.142)	(0.028)	(0.032)
Average years of schooling		1.450**		0.260**
		(0.613)		(0.126)
F-test	33.05***	14.21***	14.69***	14.00***
Adjusted R^2	0.565	0.464	0.096	0.142
Number of observations	75	62	389	315

Note: Standard errors are in parentheses. * = significant at 10 per cent, ** = at 5 per cent and *** = at 1 per cent level.

The telephone access cost depends on factors determining the prevailing supply and demand conditions of telephone communication. The access cost of an enhanced form of communication, the Internet, can be expected to depend on the same factors. It is necessary to specify such factors in a general manner, which is done in the next set of estimations reported in Table 3. Investment in telecommunications is a natural choice for a variable to be included in such a specification. In addition to its role in satisfying demand in general, investment introduces new technology, which makes it possible to respond more efficiently to the current demand and which may also create new demand. Investment creates revenue, which is paid through subscription and calling costs. We take revenues from telecommunication activities to represent the monetary flows originating from the demand side with respect to the service level provided. The investment-revenue relation can be expected to depend on the general level of demand and technical development, controlled by the GDP per capita variable for retaining comparability of estimates across countries.

Telecommunication investment and revenue levels and their relative magnitude represent the major aggregate factors determining the price level in telephony. They are also major determinants of the initial conditions for the infrastructure which the Internet is built on. Throughout the diffusion process, telecommunications investment and revenue become increasingly directly determined by the adoption of new technologies and the Internet. The share of investment aimed at merging the Internet with traditional telephony has increased dramatically during the past decade, most recently with the 120 billion dollars paid by telecommunications consortiums for the third generation mobile communication network licences.

A technical problem is that multicollinearity in the set of new variables prevents precise estimation of coefficients in a single equation. We need to expand the model with the aid of additional equations, one for technological infrastructure, and another for price determination. Access cost is the natural dependent variable in the price equation. Following Figure 1, we choose GDP per capita for the dependent variable in the infrastructure equation. Internet host level would be the other possible candidate, but GDP does not directly affect the diffusion parameter and so allows the growth curve to be positioned by the model as freely as possible. The simultaneous equation system then becomes

$$
\text{(4)}\quad
\begin{aligned}
\ln H_{it} - \ln H_{it-1} &= \alpha_i \beta_{i0} + \alpha_i \beta_{i1} \ln Y_{it} + \alpha_i \beta_{i2} \ln P_{it} + \alpha_i \boldsymbol{\gamma}_D{}' \mathbf{Z}_{it} - \alpha_i \ln H_{it-1} \\
\ln Y_{it} &= \delta_Y + \boldsymbol{\gamma}_Y{}' \mathbf{Z}_{it} \\
\ln P_{it} &= \delta_P + \boldsymbol{\gamma}_P{}' \mathbf{Z}_{it}
\end{aligned}
$$

where the additional variables $\mathbf{Z}_{it}$ are now divided among the equations by their equation-specific coefficient vectors.

The three endogenous variables in the model are the difference and level of Internet hosts, and the access cost. GDP is exogenous along with other variables, so the system is identified and estimable with three-stage least squares. This methodology uses instrumenting to account for the effects of the exogenous variables on all endogenous variables. The endogenous variables are correlated with all equation disturbances, which increase in particular the precision of the access cost estimates. Unstructured specification of cross-equation correlations best utilizes cross-equation multicollinearity and adapts to autocorrelation as well.

Table 3

Simultaneous equations estimation results for countries with more than 50 Internet hosts

Diffusion equation	$\ln H_{t+1} - \ln H_t$			
	(1)	(2)	(3)	(4)
Speed of diffusion α	0.186*** (0.033)	0.191*** (0.033)	0.194*** (0.033)	0.193*** (0.031)
Constant $\alpha\beta_0$	-5.493*** (1.040)	-5.525*** (1.037)	-5.352*** (1.050)	-4.710*** (1.037)
GDP per capita $\alpha\beta_1$	0.695*** (0.118)	0.697*** (0.118)	0.659*** (0.131)	0.565*** (0.122)
Telephone access cost $\alpha\beta_2$	-0.164** (0.056)	-0.152** (0.055)	-0.135** (0.058)	-0.123** (0.053)
Average years of schooling			0.073 (0.154)	
Attendance in university education				0.149** (0.064)
Chi square -test	42.53***	43.29***	42.69***	47.80***
Pseudo R^2	0.255	0.263	0.275	0.309
Infrastructure equation	$\ln Y_t$			
Constant	6.979*** (0.027)	6.982*** (0.053)	6.982*** (0.053)	6.982*** (0.053)
Personal computers per 1000 people	0.332*** (0.027)	0.347*** (0.027)	0.347*** (0.027)	0.348*** (0.027)
Telephone main lines per 100 people	0.278*** (0.035)	0.257*** (0.036)	0.257*** (0.036)	0.257*** (0.036)
Chi square -test	2011.55***	2003.65***	2003.65***	2003.15***
Pseudo R^2	0.933	0.934	0.934	0.934
Price equation	$\ln P_t$			
Constant	-6.539*** (1.333)	-12.357*** (2.241)	-12.362*** (2.244)	-12.423*** (2.246)
Share of telecom investment in GDP	-0.470** (0.166)	-0.299* (0.170)	-0.301* (0.171)	-0.308* (0.171)
Share of telecom revenue in GDP	0.761** (0.255)	0.461* (0.259)	0.466* (0.260)	0.473* (0.260)
GDP per capita	0.946*** (0.104)	1.170*** (0.119)	1.170*** (0.119)	1.171*** (0.120)
Gini coefficient of net income inequality		0.986** (0.350)	0.989** (0.350)	1.004** (0.351)
Chi square -test	133.86***	157.92***	158.05***	158.54***
Pseudo R^2	0.503	0.545	0.545	0.545
Number of observations	141	141	141	141

Note: Asymptotic standard errors are in parentheses. * = significant at 10 per cent, ** = at 5 per cent and *** = at 1 per cent level.

We first estimate the system with personal computers per 1,000 people and telephone main lines per 100 people as the explanatory variables in the infrastructure equation, and with GDP shares of both telecommunications investment and telecommunications revenue as well as GDP per capita in the price equation. All the variables are in logarithmic form, and the data come from ITU (2000). The results are reported in column (1) of Table 3.

Because of limitations in the data on the covariates we use, the number of observations is here lower than in the panel estimations presented in Table 2. However, the sample still includes a fair selection of countries at various stages of development and Internet adoption. When comparing the results with those obtained from the single equation models of Table 2, it is interesting to see that telephone tariffs are now statistically significant in explaining the diffusion of the Internet. This can be explained by the fact that the independent variables of the price equation, being more directly related to the Internet hosts than the price variable by itself, exert their influence through the price variable. As expected, additional investment tends to lower the cost, whereas additional extracted revenue raises the cost. Higher demand levels associated with higher GDP per capita increase the price as well. In the infrastructure equation, the number of PC's and telephone lines has a direct, very significant relation to GDP, and can be expected to exert an indirect effect of a comparable magnitude in the diffusion equation.

Next, we turn to differential Internet access. It is partly determined by different levels of technology and development. However, when comparing countries at similar levels of development, notable differences in the number of Internet hosts per capita remain. Since Internet access and equipment costs still form a considerable sum of annual earnings per capita, especially in countries where the Internet is just starting to spread, the price level is an important factor.

Higher levels of diffusion within a country need a growing part of the population with resources to access the hosts. The more unequally income is distributed within a country, the fewer people with the money to invest in Internet access, and thus the fewer hosts needed for service provision. Benefits such as everyday commercial and educational services in the local language encourage investment but can be expected to grow substantially only with a sizeable share of regular users. When the share remains small, inequality may be the factor preventing growth. Any such effect should be visible in the price equation where GDP already controls for the average level of development.

Column (2) of Table 3 shows the estimation results obtained by including the logarithm of the Gini index of inequality in percentage points in the price equation. The impact of income inequality on telephone access cost is positive and quite large. Consequently, the adverse effects of high prices are further amplified by income inequality. The Gini index is a representative number from the past decade as documented in UNU/WIDER (2000).Varying definitions are roughly corrected to household income net of tax inequality with dummy variable adjustments estimated from the total of 5,000 observations in the inequality database.

Finally, columns (3) and (4) include education variables in the diffusion equation. In contrast to the single-equation estimates, the variable measuring the average years of education is not statistically significant in column (3). It seems that, in addition to the more precise estimation method, the additional variables capture the effects of this variable. However, in column (4) the share of population over 15 years old attending tertiary education in 1995, also from Barro and Lee (2000), is significant with a positive coefficient. Thus, given the conditions set by the infrastructure and price levels, a larger share of university students, and in turn a larger share of early adopters, has positively contributed to the growth of the Internet.

5 Conclusions

We have investigated the factors which determine the diffusion of the Internet across countries by estimating the Gompertz model of technology diffusion with data on Internet hosts per capita for the years 1995-2000. For a sample of the OECD countries, the basic finding is that GDP per capita and Internet access cost explain best the observed growth in computer hosts per capita. In contrast with some earlier findings (Hargittai 1999), competition in telecommunications markets does not seem to exert any independent influence on Internet penetration. Consequently, the deregulation of the telecommunications sector improves Internet connectivity only if it lowers the access cost. There is in fact no intuitive reason to expect competition to matter irrespective of its impact through market prices.

Somewhat unexpectedly, investment in education does not have any statistically significant explanatory power in our model. This may, however, rather reflect the low variability of the schooling variable in the OECD sample than the lack of impact of education on the diffusion of the Internet. Moreover, the variable measuring proficiency in English enters the regression with the wrong sign. Also, the fact that the regional dummy variables for the ‘northern’ and ‘southern’ OECD countries improve the precision of the parameter estimates and the statistical fit of the model suggests that there are cultural or technology policy related differences between the countries in the adoption of the Internet which have not been captured in our model.

For a larger sample of both industrial and developing countries, the results change in such a way that also education becomes significant. However, since data on Internet access costs are not available for non-OECD countries we have to use telephone access costs instead. This variable becomes statistically significant in the diffusion equation only when the original model is refined to control for the factors affecting telephone tariffs and the technological infrastructure of Internet adoption. This refinement also changes the role of human capital in the model: investment in university education seems to matter more for Internet connectivity than the average years of schooling.

There are at least two important avenues for future research. One avenue is to study in greater detail the diffusion of the cluster of Internet-related technologies. After all, the use of the Internet is contingent on the adoption of a terminal for access (like the computer or the mobile phone) as well as on the availability of the concomitant infrastructure. Another avenue is to analyse the factors affecting the growth of traffic on the Internet. This would be a better indicator of its economic importance than Internet connectivity itself.

References

Barro, R. J., and J.-W. Lee (2000). 'International Data on Educational Attainment Updates and Implications'. NBER Working Paper No. 7911. Cambridge, MA: NBER.

Caselli, F., and W. J. Coleman II (2001). 'Cross-country Technology Diffusion: The Case of Computers'. NBER Working Paper No. 8130. Cambridge, MA: NBER.

European Commission (2000). *Key Data on Education in Europe, 1999-2000*. Luxembourg: Eurostat Press Office.

Hargittai, E. (1999). 'Weaving the Western Web: Explaining Differences in Internet Connectivity among OECD Countries'. *Telecommunications Policy* 23: 701-718.

Inktomi Corp. (2000). 'Web Surpasses One Billion Documents'. Available at: http://www.inktomi.com/new/press/billion.html.

Internet Software Consortium (2001). 'Internet Domain Survey, January 2001' Available at: http://www.isc.org/ds/WWW-200101/index.html.

ITU (2000). *International Telecommunication Indicators Database.* Available at: http://www.itu.int/ti/publications/world/world.htm.

Norris, P. (2000). 'The Global Divide: Information Poverty and Internet Access Worldwide'. Available at: http://www.ksg.harvard.edu people/pnorris/acrobat/IPSA 2000.pdf.

OECD (1996). *Information Infrastructure Convergence and Pricing: The Internet.* Committee for Information, Computer and Communications Policy. Paris: OECD. Available at: http://www.oecd.org/dsti/sti/it/cm/prod/e_96-73.htm.

OECD (1999). *Communication Outlook*. Paris: OECD.

Stoneman, P. (1983). *The Economic Analysis of Technological Change*. Oxford: Oxford University Press.

UNDP (1999). *Human Development Report 1999*. United Nations Development Programme (UNDP). New York: Oxford University Press.

UNU/WIDER (2000). UNU/WIDER-UNDP World Income Inequality Database 15 September 2000. Available at: http://www.wider.unu.edu.

US Department of Commerce (1998). 'The Emerging Digital Economy'. Available at: http://www.ecommerce.gov/emerging.htm.

UNU World Institute for Development Economics Research (UNU/WIDER) was established by the United Nations University as its first research and training centre and started work in Helsinki, Finland in 1985. The purpose of the Institute is to undertake applied research and policy analysis on structural changes affecting the developing and transitional economies, to provide a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and to promote capacity strengthening and training in the field of economic and social policy making. Its work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.

UNU World Institute for Development Economics Research (UNU/WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Liisa Roponen at UNU/WIDER
Printed at UNU/WIDER, Helsinki

ISSN 1609-5774
ISBN 952-455-146-2 (printed publication)
ISBN 952-455-147-0 (internet publication)