

POLICY RESEARCH WORKING PAPER

4420

Diffusion of the Internet: A Cross-Country Analysis

Luis Andrés
David Cuberes
Mame Astou Diouf
Tomás Serebrisky

The World Bank
Latin America and the Caribbean Region
Sustainable Development Department
December 2007



Abstract

This paper analyzes the process of Internet diffusion across the world using a panel of 199 countries during 1990-2004. The authors group countries in two categories—low and high-income countries—and show that the Internet diffusion process is well characterized by an S-shape curve for both groups. Low-income countries display a steeper diffusion curve that is equivalent to a right shift of the diffusion curve for high-income countries. The estimated diffusion curves provide evidence of a “catching-up” process, although a very slow

one. The paper explores the determinants of Internet diffusion at the country level and across the same income groups. The most novel finding is that network effects seem to be crucial—the number of Internet users in a country in a given year is positively associated with the number of users in the previous year. The findings also show that the degree of competition in the provision of Internet service contributes positively to its diffusion, and there are significant positive language externalities.

This paper—a product of the Sustainable Development Department in the Latin America and the Caribbean Region—is part of a larger effort in the department to understand the determinants for performance in the infrastructure sectors. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at landres@worldbank.org and tsebrisky@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

*Diffusion of the Internet: A Cross-Country Analysis**

Luis Andrés
The World Bank

David Cuberes
Clemson University

Mame Astou Diouf
International Monetary Fund

Tomás Serebrisky
The World Bank

JEL classification: O14, O33, O57

Keywords: Technological diffusion, Internet, S-shape curve, Network externalities, Digital divide

* We would like to thank Georgeta Dragoiu, Antonio Estache, J. Luis Guasch, Charles Kenny, Aitor Lacuesta-Gabarain, and John Nash for useful comments. Authors' email addresses are: landres@worldbank.org, cuberes@clemson.edu, mdiouf@imf.org, tserebrisky@worldbank.org

1. Introduction

The process of technological adoption and diffusion has been extensively studied in the literature¹. A casual look at data on the diffusion of different technologies reveals that, at a given point in time, there are important differences in the degree of diffusion or adoption across countries².

In this paper we focus on the diffusion of the Internet. Understanding the process of Internet adoption and diffusion as well as the main determinants of cross-country differences in this process seems to be of particular interest in the case of Internet adoption since, as it has long been acknowledged, the Internet is a key tool of economic development (Röller, H., and L. Waverman, 2001; Sanchez-Roble, B 1998; Kenny, C. 2003). Figure 1 illustrates the significant disparity in Internet diffusion for eight countries. Although the percentage of Internet users was quite low in the United States in 1990, the use of this technology increased to 22 percent in 1997 and 63 percent in 2004. In China, the percentages were 0.03 percent in 1997 and 7.2 percent in 2004, whereas in Tanzania it was virtually 0 percent in 1997 and 0.9 percent in 2004. The observed differences in the levels of Internet adoption across countries raise important policy questions. Of particular interest to policymakers in developing countries is the need to understand the process of diffusion to anticipate if their countries will eventually catch up and close the digital divide and, if not, what policies should be adopted to increase the speed of Internet adoption.

In this paper we make use of a unique dataset to study the process of Internet diffusion in a large set of countries for the period 1990-2004. Our analysis includes both developed and developing countries and provides separate results for different groups of countries, based on their level of income. The paper has two different parts. In the first part, we show that the process of Internet diffusion across the world is well described by an S-shape pattern. An important advantage of our dataset that differentiates it from previous studies aimed at estimating the adoption pattern followed by a new technological product is that our study includes data for the first years in which Internet was adopted and hence it facilitates the estimation of a complete S-shape curve. We estimate S-shape curves for different groups of countries and find that low income countries have a much steeper adoption profile and their curve lies to the right of that of high income countries. This finding provides support to the hypothesis that Internet adoption follows a leader-follower model whereby low income countries, as followers, have lower adoption costs³.

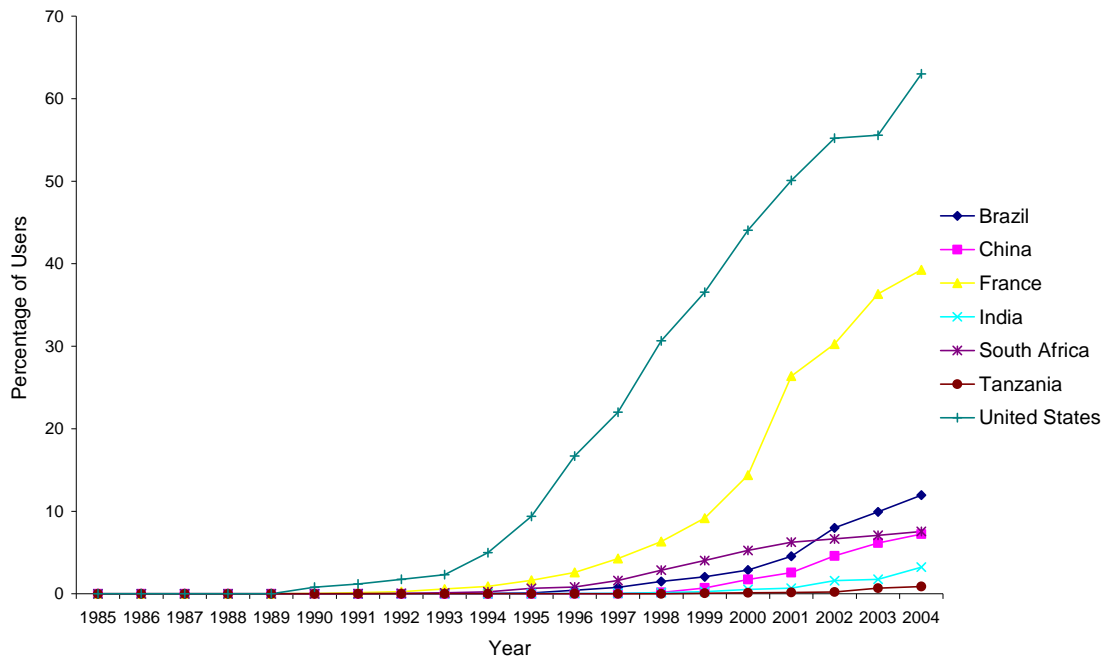
¹ See for instance Keller (2001), Comín and Hobijn (2003), Caselli and Coleman (2001), Comín *et al.* (2006), Barro and Sala-i-Martin (1997), and Jovanovic and Lach (1989).

² Chinn and Fairlie (2007) show that, in the year 1993 many developing countries had computer and Internet penetration rates that were 1/100th of the rates found in North American and Europe

³ This hypothesis was developed by Chong and Micco, 2003 but these authors did not test their hypothesis.

Figure 1:

Internet Diffusion in Different Countries



Source: ITU dataset

In the second part of the paper we identify candidates to explain the different cross-country patterns in Internet diffusion. Our most important finding is the presence of network effects in Internet diffusion: the number of Internet users in the previous year is a powerful determinant of the number of Internet users in the current year. The presence of network effects as a determinant of Internet adoption has been largely ignored in existing literature⁴, a striking fact considering that the utility derived from Internet consumption is affected by the number of people using it (Shy, 2001). In addition, we study two determinants that have not been accounted for in the literature: the competition in the market for the provision of Internet services and the role of language. Controlling for different relevant variables, we find that in countries where there is more competition in the distribution of Internet, the number of users increases more rapidly. Sharing a common language has a positive impact on the spread of Internet use.

The paper is organized as follows. Section 2 briefly summarizes the existing theoretical and empirical literature on technology diffusion and, in particular, on Internet diffusion. Section 3 describes the main dataset used throughout the paper. The empirical estimation of Internet diffusion curves is presented in Section 4. In Section 5 we explore the determinants of Internet diffusion across countries. Finally, Section 6 concludes the paper.

⁴ Estache *et al.* (2002) is an exception as these authors included a lagged variable of Internet users but found it not significant as determinant of Internet adoption.

2. Related Literature

There is a vast literature exploring the process of technology diffusion across different countries. Since the emphasis of our paper is empirical, the following summary will omit most of the theoretical analysis⁵.

Most of the empirical papers on technological diffusion focus on identifying variables that can explain some features of the diffusion process of different technologies. For instance, Gort and Klepper (1982) trace the history of diffusion for 46 new products and correlate it with different economic indicators. Caselli and Coleman (2001) focus on personal computer adoption and provide a comprehensive cross-country analysis that tries to disentangle its main determinants. Finally, Pohjola (2003) studies observed investment in information and communication technology in 49 countries during the period 1993-2000.

Other papers have focused on Internet diffusion; Chong and Micco (2003) study the spread of Internet in Latin America and argue that although latecomers, Latin American countries have the advantage of lower costs of adoption and could easily catch up with technological leaders. They also find that a country's capacity to innovate helps explain the extent to which Internet is adopted. Other empirical studies of the determinants of Internet usage are Chinn and Fairlie (2007), Estache *et al.* (2002), Canning (1999), Klobas and Clyde (1998), Kiiski and Pohjola (2002), and Quibria *et al.* (2002).

Several authors have studied the well-known fact that technology diffusion follows an S-shaped pattern. This empirical recurrence is documented by Griliches (1957), Davies (1979), and Gort and Klepper (1982), and Mansfield (1961), and is modeled in Jovanovic and Lach (1989) among others. More recently, Comín *et al.* (2003, 2006) study the diffusion processes of several technologies in different countries over the last 200 years. They find that, once the intensive margin of technological diffusion is taken into account, the evolution of the level of technology in a country does not typically follow an S-shaped pattern.

Finally, another strand of the literature has analyzed the positive effect of technology adoption (and, in particular the adoption of Internet) on the growth performance of a country and on the Digital Divide across countries—the gap in access to information technologies between developed and developing countries. Some interesting studies along these lines are Röller and Waverman (2001), Dewan *et al.* (2004), Gramlich (1994), World Bank (1994), Sanchez-Robles (1998), Guillén and Suárez (2005), Norris (2000), OECD (2001), and Kiiski and Pohjola (2002).

⁵ A more complete summary of the theoretical literature can be found in Keller (2001).

3. The Data

Technological diffusion is defined in Gort and Klepper (1982) as "the spread in the number of producers engaged in manufacturing a new product". Given the nature of the Internet, we have to adapt this definition, whereby diffusion refers to the number of consumers of Internet. Although one could think of many indicators of Internet diffusion⁶, the two most widely used are: the number (and percentage) of Internet users and subscribers⁷. Tables 1A-2A in the Appendix present descriptive statistics for Internet users. Conceptually, both variables are different since users include intrahousehold access to Internet as well as access in public places (universities, libraries). The results using both measures of Internet diffusion are qualitatively similar. Hence, in this paper, we focus only on the results for Internet users.⁸

The main dataset that we use in this paper is from the International Telecommunication Union Data (2006). This dataset contains information on a set of 199 countries for the period 1990-2005.⁹ By including both developed and developing countries in our sample we take into account the importance of problems of sample selection raised by De Long (1988) in the context of the literature on growth convergence.

One important thing to note is that, for most technologies, the relevant measure of their diffusion is the ratio of actual users to *potential* users. Measuring potential users is problematic since it requires access to micro data, which is unavailable for most of the countries in our sample. However, as noted by Dasgupta *et al.* (2001), in the case of Internet, human capital requirements to use its basic applications (electronic mail and information search) are relatively low. Therefore, it seems reasonable to assume, as we do, that the entire population is a potential user.

Finally, we use the *World Bank Country Classification* to study potential different diffusion patterns across income levels¹⁰. The World Bank classifies countries in the year 2005 according to their GNI per capita in five different groups. These groups are *low-income* economies (54), *lower-middle-income* economies (58), *upper-middle-income* economies (40), *high-income* economies (56), and *high-income OECD members* (24). All the countries included in the last group are also present in the previous one, so we disregard OECD members throughout the analysis. The list of countries included in each

⁶ Press (2000) provides a long list of such indicators: connectivity, host count, number of web sites, language distribution, compound indices of pervasiveness, geographic dispersion, sectoral absorption, connectivity infrastructure, organizational infrastructure, and sophistication of use.

⁷ Ideally, we would also like to have measures of the quality and the intensive use of the Internet but this information is not available for a large enough group of countries. A developing nation might have a very slow connection to the Internet. For example, in the spring of 1999, Cuba's total international bandwidth was 832 kb/s, which is less than a home with high-speed DSL service or cable modem. Furthermore, connectivity was concentrated in Havana (though less concentrated than in many developing nations) and limited to relatively few people, almost exclusively through their work (See Martinez, 1999). The picture is much worse in Africa (see Jensen, 2005).

⁸ Results for Internet subscribers are available from the authors upon request.

⁹ See the Appendix 2 for a list of the countries considered in this study.

¹⁰ The income thresholds are: low income, \$875 or less; lower middle income, \$876-\$3,465; upper middle income, \$3,466-\$10,725; and high income, \$10,726 or more.

group is shown in Appendix 2. Throughout the paper we follow the World Bank classification but opted to group countries in two categories: Low-income and lower-middle-income countries and upper-middle-income and high-income countries. This grouping provides more degrees of freedom to the panel estimations.

4. An Estimation of the Diffusion Process of the Internet

As noted in Jovanovic and Lach (1989), there exists strong empirical evidence to support the view that the diffusion path of both new processes and product innovations follow an S-shaped or logistic pattern. Kotler (1986) interpreted this fact as evidence in favor of the existence of four phases for technology adoption: introduction, growth, maturity, and decline.

This paper presents two contributions to this stylized fact. First, as acknowledged by Comín and Hobijn (2003), most of the previous empirical studies on the diffusion of technology do not have access to data for the initial years in which the innovation (or new product) was adopted. This translates into a lack of precision in the estimation of diffusion and, in particular, to S-shape curves that grow “too fast” during the early introductory phase. This lack of data is due to the fact that, in most cases, information on the use of a given new technology has only been collected after it has been widely adopted.¹¹

Our first contribution is to make use of a complete dataset that allows us to study the diffusion process of Internet since the first years in which it started to spread in the leading country- the U.S. This clearly overcomes the selection effect present in most of the existing studies.

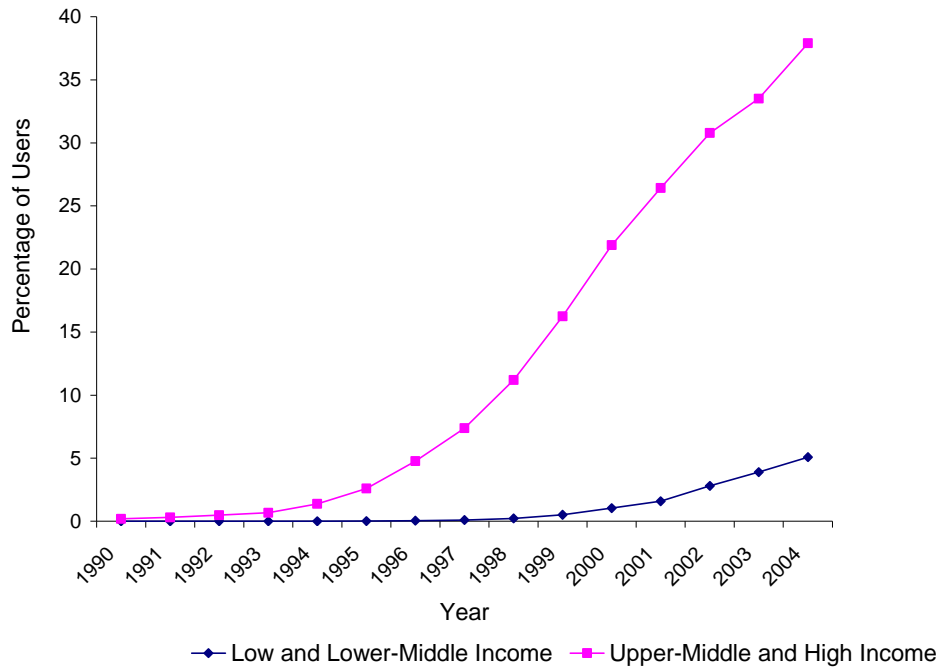
The second contribution is to estimate the hypothesis that Internet diffusion follows an S-shape curve for low and high income groups of countries. There are theoretically sound reasons to believe that the diffusion process of a given technology should be significantly different for poor and rich countries (see for instance Barro and Sala-i-Martin 1997, 2004).

Figure 2 plots the actual percentage of Internet users across income groups. It is clear from this graph that, at any point in time, the degree of Internet adoption is much lower in poorer economies than in more developed ones. The use of Internet in low and lower-middle income countries did not start until 1994, whereas in several high income countries it started already in 1990. In 2004, in high income countries, about 40 percent of the population enjoyed Internet services, while the percentage was less than 3 percent in low income countries. In other words, the diffusion curve of low income countries can be described as a right shift of the one displayed by high income countries. This stylized fact seems to support the leader-follower model of Chong and Micco (2003).

¹¹ Comín and Hobijn (2003) mention the example of the telephone, which was invented by Alexander Graham Bell in 1876 but most countries did not publish official statistics on its diffusion until the early years of 1900s.

Figure 2:

Internet Users (actual data) in low and high income Countries



To estimate the diffusion process of Internet, we use a logistic function. This functional form has often been used to approximate the S-shaped diffusion process due to its relative simplicity¹². Equation (1) presents the expression of the logistic function used for the estimations:

$$Y_{it} = \frac{\delta_0}{1 + e^{-(\delta_1 + \delta_2 t)}} + \varepsilon_{it} \quad (1)$$

where Y_{it} represents the percentage of Internet users in country i at period t , and ε_{it} is a white noise. The parameter δ_0 is the long-run outcome, i.e. the limit of Y_{it} when t goes to infinity, δ_1 is a constant of integration that positions the curve on the time scale, and δ_2 reflects the speed of adoption. We estimate (1) using a nonlinear least squares method that fits an arbitrary nonlinear regression function by least squares.

¹² Other S-shaped functions used include the cumulative normal and the Gompertz model.

The results of the estimation are displayed in Table 1. The first column presents the estimates for the entire world. Columns 2 and 3 show the estimates for the two income categories of countries. The parameter of the speed of adoption (δ_2) is higher for low and lower-middle income countries than for upper-middle and high income countries. This difference suggests that low income countries have a higher adoption speed, which is consistent with the hypothesis that low income countries are ‘catching up’, and also consistent with the hypothesis that followers tend to adopt technologies faster than leaders.

Table 1: Estimates of the S-shape Function for Internet Users

	Entire World	Low and Lower-Middle Income Countries	Upper-Middle and High Income Countries
	[1]	[2]	[3]
δ_0	16.31*** (0.5)	6.99 (0.28)	40.94*** (0.87)
δ_1	-5.83*** (0.14)	-9.57*** (0.26)	-5.76*** (0.16)
δ_2	0.49*** (0.02)	0.7*** (0.02)	0.53*** (0.02)
Observations	15	15	15
R^2	0.99	0.99	0.99

Notes: Standard errors in brackets. *, **, ***: significant at the 10, 5, and 1 percent level, respectively.

Figures 3, 4, and 5 display the actual and predicted percentages of Internet users in the world during the years 1990-2004 and for each of the country income groups for the same period of time. The first thing to notice is that the fit of the data is very accurate as the R^2 's of the regressions in Table 1 indicate. Moreover, in particular for upper-middle and high income countries one can appreciate a very clear S-shape pattern—while the diffusion of Internet is very slow in the initial years, it speeds up in the middle years, and then finally slows down in the last periods. As mentioned above, the fact that our dataset contains information from the very first years of Internet diffusion helps explain why we obtain a complete S-shape curve while most of the related literature does not.

One interesting thing to notice in Figures 4 and 5 is that in low and middle income countries diffusion of Internet accelerates in the last years, whereas in upper middle and high income countries the process significantly flattens out. This graphical evidence also suggests that low income countries are catching up with upper-high income countries in the diffusion of Internet.

Figure 3:

Internet Use in the World

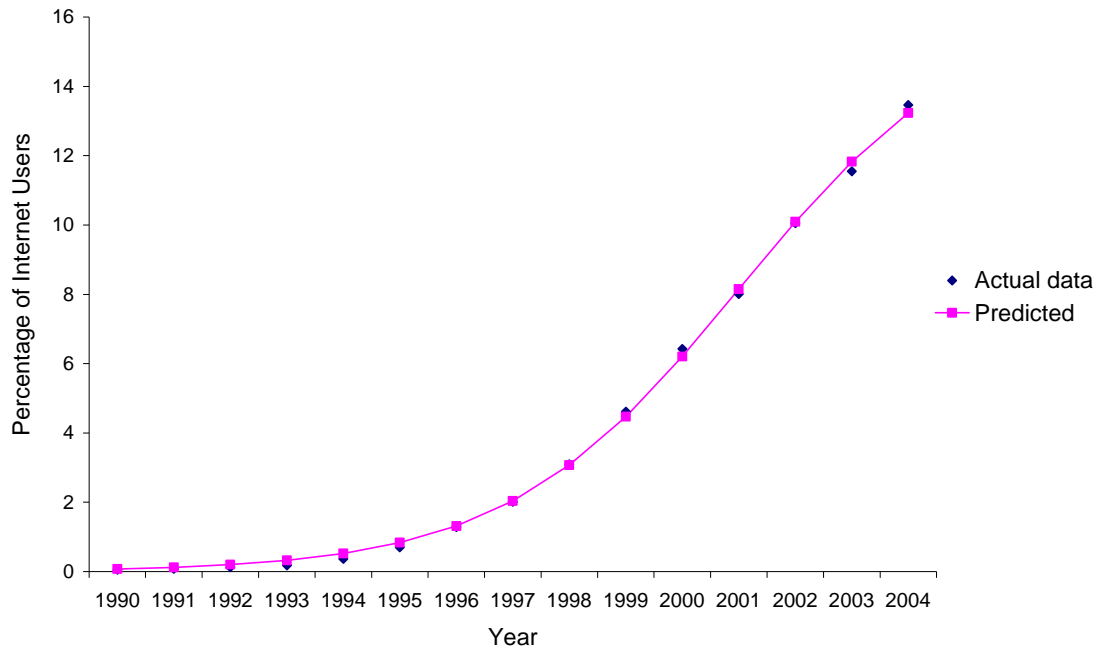


Figure 4:

Internet Use in Low and Lower-Middle Income Countries

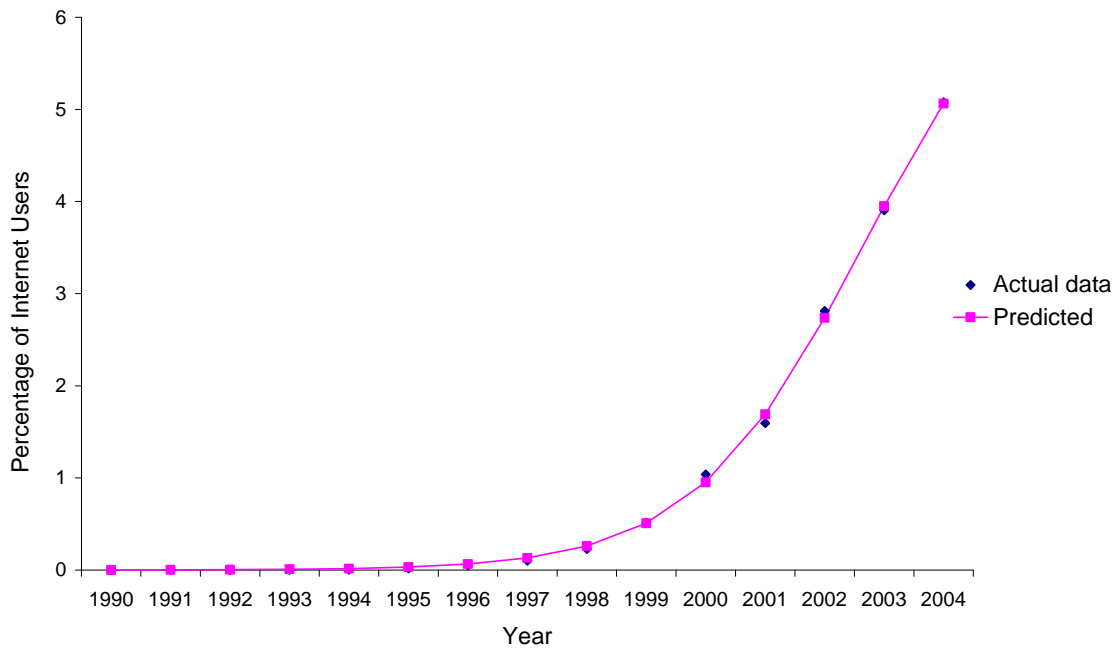
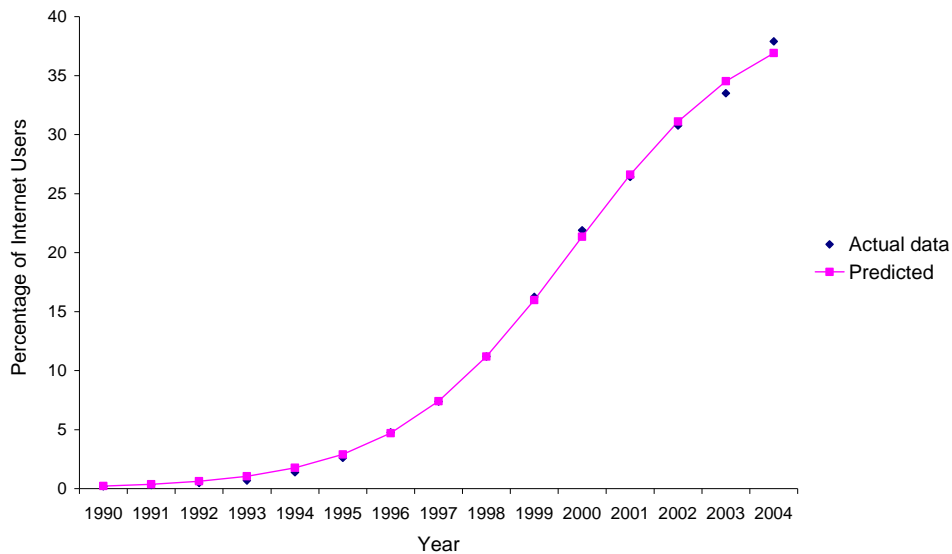


Figure 5:

Internet Use in Upper-Middle and High Income Countries



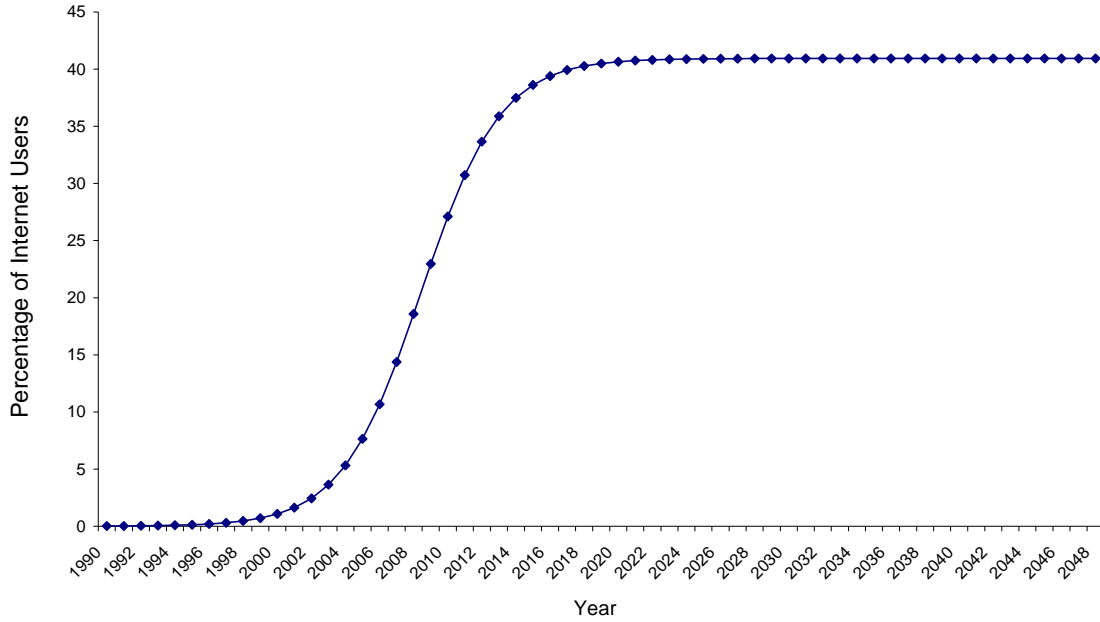
A striking result of Table 1 is the estimate of δ_0 . According to the results from Table 1, there is no long term convergence in Internet diffusion between low and high income countries. Upper-Middle and high income countries converge to an adoption rate of 40.9 percent, whereas low and lower-income ones reach a much lower percentage in the long run (16.3 percent). This result is explained by the short time period of available data – fifteen years, and the significant current difference in percentage of users in low and high income countries. Provided the estimation indicates low income countries are catching up as the speed of adoption is higher, a relevant question would be: how many years would it take low and lower-middle income countries to reach the long term adoption rate of upper-middle and high income countries? To answer this question we estimate equation (1) imposing the restriction that the δ_0 coefficient of low and lower-middle income countries equals 40.9, the estimated long run adoption rate of upper-middle and high income countries. The result of the estimation is shown in Figure 6. Using the actual available diffusion data, it is estimated that lower-middle income countries would take 8 years, i.e. in the year 2015, to catch up to the 90 percent long term level of upper-middle-income countries.

In this section we showed that Internet adoption follows an S-shape pattern, but this pattern is different for low and high income countries. Internet diffusion in the former group started with a lag¹³, but is enjoying a higher adoption speed. However, the digital divide, in absolute terms, is impressive. A simplified simulation indicates it would take low income countries more than 50 years to close the digital divide. These results motivate the need to further study the determinants of Internet adoption, a task we take in the following section.

¹³ While upper-middle income countries reach a 2 percent usage of Internet in 1995, low and lower-middle income countries reached the same usage rate in 2002.

Figure 6:

Predicted Catch-Up in Internet Adoption between Poor and Rich Countries



5. The Determinants of Internet Adoption

To identify the main determinants of Internet adoption we follow Estache *et al.* (2002) and Caselli and Coleman (2001) and estimate a reduced form model presented in equation (2).

$$\ln IU_{it} = \alpha + \beta_1 \ln Y_{it} + \beta_2 \ln P_{it} + \beta_3 \ln L_{it} + \beta_4 \ln C_{it} + \eta_i + \varepsilon_{it} \quad (2)$$

where IU_{it} represents the number of Internet users in country i and period t , Y_{it} is the real GDP per capita, P_{it} represents the real cost of a local phone call, and L_{it} and C_{it} are the number of phone lines and computers per capita, respectively. The last two variables are intended to capture the level of telecommunication infrastructure and the availability of infrastructure facilities needed to access the Internet of a country at a given point in time, respectively. Finally, η_i is a country fixed effect and ε_{it} is a standard error term.

One would expect β_1 to be positive since a higher level is naturally associated with better technological infrastructure, better organized labor markets and a higher purchasing power of goods and services associated with the Internet. The coefficient of the cost of a local phone call (β_2) is expected to be negative and its magnitude would depend on the price elasticity of the demand for Internet usage. Both the effects of the number of phone lines (β_3) and the number of computers (β_4) should be positive, since they both are

necessary inputs to use the Internet. Moreover, one would expect a strong positive complementarity between computer and Internet use¹⁴.

Other studies¹⁵ added additional explanatory variables to equation (2), including a country's level of human capital—proxied by the number of years of education, its degree of trade openness, the percentage of urban population and the extent of property rights protection. We opted not to incorporate these variables in the analysis since our data suggests that, in most cases, they display very little variation in the time interval covered by this paper. Anyhow, the inclusion of a country fixed effect in our estimation should be able to capture cross-country differences explained by these variables¹⁶.

The results of estimating equation (2) using panel data country-specific fixed effects are presented in Table 2.¹⁷ All the coefficients are highly significant and have the expected signs. The levels of income and telecommunication infrastructure (lines per capita) have significant positive effects on Internet adoption and seem to have a similar magnitude. A 10 percent increase in per capita GDP is associated with a 22.6 percent increase in the number of Internet users. Similarly, increases of 10 percent in the number of lines and computers per capita drive up the number of Internet users by 21.6 and 22.9 percent respectively. On the contrary, the higher the cost of a local call, the lower is the probability that Internet will be adopted.

¹⁴ A possible criticism to the specification of the model is the inclusion of the cost of local phone calls given the increase in alternative technologies to access the Internet (for instance broadband access). Up to 2004, the participation of alternative technologies was very low, in particular in low income countries.

¹⁵ For a sample of these studies see Chinn and Fairlie (2007), Wallsten (2005), Kiiski and Pohjola (2002), and Chong and Micco (2003).

¹⁶ Interestingly, coefficients associated to some of these variables have often been found non significant or controversial in internet adoption models. This is the case of education in Kiiski and Pohjola (2002) and Chinn and Fairlie (2007), and openness and property rights protection in Caselli and Coleman (2001).

¹⁷ Henceforth we use robust standard errors in our regressions to account for potential heteroskedasticity of the unbalanced panel.

Table 2: Benchmark model

	[1]
Log real GDP per capita	2.255*** [0.559]
log real cost	-0.210*** [0.077]
Log lines per capita	2.155*** [0.266]
Log computers per capita	2.292*** [0.132]
Constant	-12.727** [5.242]
Method of estimation	GLS
Number Observations	955
R ²	0.91

Notes: The dependent variable is the log of Internet users per capita. Robust standard errors in brackets. *, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively. The regression uses countries with GDP per capita growth in the interval [-4.9, 9.2 percent] and controls for country fixed effects.

5.1. The diffusion model

Specification (2) has often been criticized because it does not account for the process of diffusion in Internet adoption. Following Estache and al. (2002), we include the lag of the number of Internet users (in logs) as a right-hand-side variable:

$$\ln IU_{it} = \alpha + \beta_1 \ln Y_{it} + \beta_2 \ln P_{it} + \beta_3 \ln L_{it} + \beta_4 \ln C_{it} + \beta_5 \ln IU_{it-1} + \eta_i + \varepsilon_{it} \quad (3)$$

Equation (3) is the reduced form of a Gompertz model of technology diffusion with a constant speed of adjustment. In such a model, the change in the number of users (from the current period to the next one) is expressed as a fraction—the speed of adjustment—of the gap between the number of users at the equilibrium and the number of current users. Hence, the number of new users who adopt a certain good or service in a given period depends on both the number of existing and potential users—which is itself determined by demand-side variables (income, costs, etc.), and other variables describing the demand or supply conditions or the technological infrastructure in each country i (see Stoneman, 1983; Kiiski and Pohjola, 2002; and Estache *et al.* 2002 for more details).

The coefficient β_5 measures the importance of network externalities in the diffusion of Internet. In the absence of diffusion, β_5 is not significant. When β_5 is positive and smaller than 1, the diffusion model is accepted. Thus the number of users in the current period explains (in a percentage given by β_5) the number of Internet users the subsequent year.

As it is well known, including the lagged dependent variable in the right hand side of equation (3) creates an endogeneity problem. By construction, the unobserved panel-level effects are correlated with the lagged dependent variables, making standard estimators inconsistent. Arellano and Bond (1991) derived a consistent generalized method-of-moments (GMM) estimator for this model. To correct this problem, we estimate Equation (3) using the instrumental variables (IV) procedure proposed by Arellano and Bond (1991) where the lagged values of the dependent variable are used as instruments and country fixed effects are accounted for. Table 3 presents the results of the GLS and IV estimations (specifications (2) and (3) respectively).

First of all, the diffusion coefficient (the lag of the dependent variable) is positive and highly significant with very similar values in both regressions, indicating that the diffusion model cannot be rejected. A 10 percent increase in the number of Internet users in the current year leads to an increase of about 6.8 percent in the number of Internet users the next year. In the naïf regression (GLS estimation presented in column [2]) income and the number of computers per capita seem to have a positive effect on diffusion, although both variables become insignificant when we use IV.¹⁸

The importance of the lagged dependent variable is in line with the results of several papers including Goolsbee and Klenow (2002) and Kiiski and Pohjola (2002). It however contradicts the results of Estache *et al.* (2002) where the diffusion hypothesis is rejected. The fact that in the IV specification the lagged dependent variable is the only significant explanatory variable suggests that network externalities drive Internet diffusion and might indeed be its most important determinant.

Tables 2 and 3 provide results on what determines Internet diffusion aggregated at the world level. However, in Section 4 we presented conclusive evidence that the process of Internet diffusion is far from being uniform across countries. Thus, the question that needs to be answered is if the variables that explain the process of Internet diffusion differ between low and high income countries. Only by identifying differences in the explanatory power of variables that influence the diffusion process, will it be possible to adopt policies aimed at reducing the digital divide.

¹⁸ Using the lagged values of the dependent variable considerably reduces the number of available observations. We have estimated (2) by GLS using the same sample as in the IV specification to make sure that the different results does not result from this loss of information. The results show that this is not the case.

Table 3: Naïve and instrumental variable estimations of the diffusion model

	[1]	[2]
Log real GDP per capita	0.782*** [0.261]	0.585 [0.550]
log real cost	0.016 [0.024]	0.009 [0.051]
Log lines per capita	0.034 [0.135]	-0.109 [0.226]
Log computers per capita	0.540*** [0.083]	0.096 [0.172]
Lagged dependent variable	0.675*** [0.022]	0.676*** [0.049]
Constant	-5.916** [2.448]	0.100*** [0.038]
Method of estimation	GLS	IV
Number Observations	886	764
R ²	0.97	-

Notes: The dependent variable is the log of Internet users per capita. The robust standard errors are in brackets.

*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.

The regression uses countries with GDP per capita growth in the interval [-4.9 , 9.2 percent] and controls for country fixed effects.

5.2. Does the magnitude of Internet diffusion explanatory variables vary with the level of income?

The benchmark model showed that the level of income is positively and highly correlated with Internet adoption. However, although income has the expected sign in the IV estimation, it is no longer significant. This result does not provide information about the existent varying processes of Internet adoption across countries, each with very different levels of income. Section 4 provided sound evidence that the shape of the Internet diffusion process is different for low and high income countries. To assess if there are significant differences in the explanatory power of the variables that are more likely to account for the Internet diffusion process, we estimate equation (3) dividing the sample in two groups: low income countries (defined as an aggregate of low and lower middle income countries) and high income countries (an aggregate of upper middle and high income countries).

When the number of users in the previous period is included in the regression (columns [2] and [4] of Table 4¹⁹), it becomes the only significant variable for both low and high income countries. This result would indicate that network effects are the drivers of Internet diffusion. The literature that studies digital divide and the factors that account have, except for few exceptions like Estache *et al.* (2002), ignored the importance of network effects.

Our estimates indicate that network effects are higher in high income countries. A 10 percent increase in the number of Internet users in high income countries in one year leads to an increase of 8 percent in the number of users the following year, 2 percent more than in low income countries. One possible explanation for this finding may be related to the structure of high income countries economies. The relative importance of services (usually high intensive in Internet usage) in these economies is higher than in less developed ones. Accordingly high income countries may have more Internet-demanding and network-demanding jobs.

Table 4. Impact of income categories on the diffusion process

	Low and Lower Middle Income		Upper Middle and High Income	
	[1]	[2]	[3]	[4]
Log real GDP per capita	2.853** [1.103]	0.576 [0.881]	1.083 [0.661]	0.447 [0.612]
Log real cost	-0.143 [0.114]	0.028 [0.065]	-0.472*** [0.120]	-0.03 [0.079]
Log lines per capita	2.721*** [0.358]	0.066 [0.293]	1.130*** [0.385]	-0.273 [0.302]
Log computers per capita	1.952*** [0.194]	0.236 [0.213]	2.800*** [0.146]	-0.162 [0.257]
Lag Internet users		0.576*** [0.057]		0.801*** [0.088]
Constant	-9.903 [8.967]	0.142** [0.067]	-9.225 [6.822]	0.072* [0.042]
Method of estimation	GLS	IV	GLS	IV
Fixed effects	Yes	Yes	Yes	Yes
Number Observations	464	358	491	406
R ²	0.87	-	0.9	-

Notes: The dependent variable is the log of Internet users per capita. The robust standard errors are in brackets.

*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.

The regression uses countries with GDP per capita growth in the interval [-4.9 , 9.2 percent] and controls for country fixed effects.

Besides studying the determinants of Internet adoption dividing countries by income level, another important question to understand the process of Internet adoption is whether estimates are constant through time. In order to do so, we divide the sample in two sub-periods: 1985-1998 and 1999-2004 (the results are presented in Table 5). First,

¹⁹ For the remaining of the paper we rely on the IV estimation. We display the results of the benchmark model when we consider the comparison is relevant.

the model of diffusion is accepted in both sub-periods and all income categories. Second, for all income categories, the importance of network effects increased with time. An increase of 10 percent in the number of Internet users in low income countries in one year advanced the number of users the next year by 3.9 percent between 1985 and 1998 and augmented it by 6.11 during the 1999 to 2004 period. A similar result holds for high income countries where the diffusion coefficient increases from 0.617 in 1985-1998 to 0.859 in 1999-2004 and globally from 0.518 to 0.736. Interestingly, we note that these estimations are consistent with the S-shape curve we estimated for each income category in Section 4. The value of the lagged number of Internet users captures the average contribution of network effects to the speed of adoption, controlling for other variables, and thus shows that the speed of diffusion increases between the two periods considered for this exercise, implying that we are identifying the first phase of the Internet adoption process. It also shows how the low and lower middle income countries are following a similar adoption path of upper middle and high income countries. As Table 5 shows, the diffusion coefficient for the last period for low income countries resulted similar to the one we obtained for the period 1985-98 for the leading group of countries.

Table 5: World Diffusion and Diffusion by Income Category: 1985-1998 and 1999-2004

	Low and Lower Middle Income		Upper Middle and High Income		World	
	1985-1998	1999-2004	1985-1998	1999-2004	1985-1998	1999-2004
	[1]	[2]	[3]	[4]	[5]	[6]
Log real GDP per capita	-0.862	1.607*	2.476***	0.217	0.019	0.79
	[2.311]	[0.826]	[0.707]	[0.728]	[0.925]	[0.557]
Log real cost	-0.099	0.077	0.055	0.002	-0.031	0.053
	[0.158]	[0.072]	[0.089]	[0.089]	[0.112]	[0.063]
Log lines per capita	-0.049	0.16	-0.997***	0.305	0.053	0.327
	[0.552]	[0.384]	[0.315]	[0.558]	[0.417]	[0.304]
Log computers per capita	0.052	0.173	-0.428	-0.114	-0.001	0.191
	[0.445]	[0.186]	[0.337]	[0.227]	[0.308]	[0.152]
Lag Internet users	0.395***	0.611***	0.617***	0.859***	0.518***	0.736***
	[0.109]	[0.063]	[0.092]	[0.078]	[0.084]	[0.052]
Constant	0.502***	0.083	0.252***	-0.032	0.282***	-0.016
	[0.188]	[0.064]	[0.074]	[0.035]	[0.080]	[0.032]
Method of estimation	IV	IV	IV	IV	IV	IV
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number Observations	101	257	215	191	316	448

Notes: The dependent variable is the log of Internet users per capita. The robust standard errors are in brackets.

*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.

The regression uses countries with GDP per capita growth in the interval [-4.9 , 9.2 percent] and controls for country fixed effects.

5.3. Impact of the level of competition on Internet diffusion

As mentioned in the introduction, there is plenty of evidence that the adoption of information and communication technologies (ICT) in general and Internet in particular contribute to economic growth and development (for a summary of this literature see Grace *et al.*, 2004 and Zhen-Wei Qiang *et al.*, 2004). Thus, a key policy question is what can low income countries do to accelerate Internet diffusion? One possible answer is to liberalize the telecommunication markets, hoping that more competition will lower prices and facilitate the diffusion of Internet. To assess the validity of this argument, we analyze the impact of the number of Internet Service Providers (ISP) operating in a country on the speed of diffusion. We decompose the sample in two groups: countries with low level of competition—with a number of ISPs less than 4, and countries with high level of competition—with a number of ISPs equal or higher than 5.²⁰ Then, we estimate equations (2) and (3) for each group, also distinguishing between low and high income countries. Provided the lag variable of Internet users is, *ceteris paribus*, a proxy for the average speed of diffusion, we are interested in determining if a more competitive ISP market structure leads to a higher estimate of the lag Internet user variable.

Results are displayed in Table 6. First, once again, one cannot reject the model of diffusion. Network effects are the main determinant of Internet adoption for all income categories and degrees of competition (see specifications 2, 4, 7, 9, and 10). Second, competition has a larger impact on diffusion in high income countries. For low and lower income countries, increasing the number of ISPs from 4 or less to more than 5 increases the diffusion coefficient from 0.494 to 0.582 while for upper middle and high income countries, the jump is from 0.300 to 0.806.²¹ Finally, regressions 5 and 10 show that when we use the interaction between high competition and lagged Internet users a high degree of competition significantly increases the average speed of diffusion in high income countries while the effect is insignificant in low income countries. We believe this approach is more informative than the one used in Estache *et al.* (2002) and Wallsten (2005) who simply added a dummy variable on the right hand side of the regression to account for the existence of competition and/or regulation. In fact, this latter work only allows the intercept to adjust for the conditions on the telecommunication market while our approach allows all coefficients to adjust for the degree of competition.

²⁰ The median of the number of ISPs for the world is 5. We adopted this threshold to define countries with low and high competition.

²¹ When considering low competition countries, the diffusion coefficient of high income countries becomes lower than those of low and middle income countries (regressions 2 and 7). This result contradicts our former conclusions but its validity is questionable because of the low number of observations in the rich country regression (41 observations, 9 countries).

Table 6: Impact of the number of Internet Service Providers on the Internet diffusion process

	Low and Lower Middle Income Countries					Upper Middle and High Income Countries				
	1 ≤ # of ISPs ≤ 4		# of ISPs ≥ 5		All	1 ≤ # of ISPs ≤ 4		# of ISPs ≥ 5		All
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Log real GDP per capita	3.837** [1.728]	3.453** [1.393]	2.162 [1.469]	-0.036 [1.120]	1.166 [0.868]	2.34 [2.671]	-0.234 [1.182]	1.139* [0.678]	0.496 [0.689]	0.457 [0.635]
Log real cost	0.074 [0.216]	0.032 [0.155]	-0.175 [0.117]	0 [0.069]	0.027 [0.061]	0.224 [0.451]	0.006 [0.335]	-0.506*** [0.124]	-0.062 [0.083]	-0.058 [0.078]
Log lines per capita	3.002*** [0.573]	-0.033 [0.448]	2.585*** [0.433]	0.036 [0.401]	0.081 [0.304]	-0.3 [1.067]	0.711 [0.511]	1.298*** [0.412]	-0.175 [0.368]	-0.158 [0.329]
Log computers per capita	1.842*** [0.357]	0.334 [0.244]	2.040*** [0.225]	-0.013 [0.314]	0.169 [0.202]	2.762*** [0.409]	-0.018 [0.293]	2.766*** [0.155]	-0.17 [0.318]	-0.092 [0.258]
Lag Internet users		0.494*** [0.098]		0.582*** [0.067]	0.549*** [0.058]		0.300** [0.152]		0.806*** [0.094]	0.656*** [0.096]
Lag Internet users * I _{ISP ≥ 5}					-0.006 [0.047]					0.124* [0.064]
Constant	-13.174 [14.199]	0.156 [0.098]	-6.777 [12.082]	0.196** [0.082]	0.166*** [0.063]	-16.976 [28.375]	0.281*** [0.103]	-10.125 [7.014]	0.070* [0.042]	0.074* [0.038]
Method of estimation	GLS	IV	GLS	IV	IV	GLS	IV	GLS	IV	IV
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number Observations	204	153	257	203	356	57	41	434	365	406
R ²	0.87	-	0.87	-	-	0.89	-	0.91	-	-

Notes: The dependent variable is the log of Internet users per capita. Robust standard errors in brackets.

*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively. The regression uses countries with GDP per capita growth in the interval [-4.9, 9.2 percent] and controls for country fixed effects.

5.4. Impact of language externalities on diffusion of the Internet

Given the fact that Internet diffusion is heavily determined by network effects, it is important to understand if language plays a role in the size of the network effects. Is the number of Internet users in Argentina (Spanish speaking country) likely to foster the number of Internet users in China? It is intuitive to think that sharing the same spoken language as other users would also contribute to network effects and thus to Internet diffusion²²? The importance of language as an explanatory variable of Internet diffusion has been ignored in the literature.

To test the role of language on Internet diffusion, we classify countries by official language using the ten most used languages in the Internet. These languages are English (with 28.9 percent of the users), Chinese (14.7 percent), Spanish (8.9 percent), Japanese (7.6 percent), German (5.2 percent), French (5.0 percent), Portuguese (3.6 percent), Korean (3.0 percent), Italian (2.7 percent), and Arabic (2.5 percent).²³ All other languages with a share of less than 2.5 were grouped in the category “others”.

Table 7 presents the estimation results of equation (3) where the diffusion effect related to a network with the same language is identified separately from the general diffusion effect within a given country. For the estimation we define a variable equivalent to the lagged variable of the sum of users that speaks the same language around the world excluding the users of a given country. The first element that this table shows is that the results and estimates are consistent with the results presented in tables 3 and 4. It also demonstrates that language, as a proxy of network externalities of users speaking and sharing contents in a common language matters. Controlling by the diffusion coefficient within the country i , a 10 percent increase in the number of users sharing the same language in other countries will increase 2.4 percent the number of users in country i . Another result is that the language externality effect seems to be more relevant for upper middle income countries. Again, this may suggest that the two groups of countries under study are facing different stages of the diffusion process. It also may suggest that when users basically connect to Internet for social and personal reasons or for local work use, one can expect the (internal) national network to be the only significantly important network. However, when the use is more professionally oriented, with internationally opened and globalized firms, the international network is likely to play a preponderant role in the incentives of being an Internet user. It is very likely that this is the case of new users in the upper middle income countries. In the case of low income countries, it is probable that users that access the Internet in the first stage are not constrained by language. Assuming that the main group of users for a given country is the one with higher income and human capital, it is very likely that this group is also able to read English, making less relevant the need of available contents in their own language.

²² Continuing with the example of Argentina, the number of users in this country should increase when the number of users in the rest of Spanish speaking Latin American countries and Spain increase.

²³ Source: www.internetworldstat.com as of June 2, 2007. We group the rest of languages together. The official language of each country has been defined according to Wikipedia. Whenever a country has several official languages, we set English as the primary official language if among the choices, otherwise we choose French or the next most spoken language worldwide

However, it is probable that as soon as the low income countries reach higher Internet penetration, language will start to matter as indicated by the results for upper middle income.

Table 7: Impact of sharing the same language on the diffusion process

	Low and Lower Middle Income	Upper Middle and High Income	All
	[1]	[2]	[3]
Log real GDP per capita	0.904 [0.827]	0.671 [0.685]	0.282 [0.539]
Log real cost	-0.065 [0.084]	-0.02 [0.074]	0.005 [0.048]
Log lines per capita	0.258 [0.293]	-0.303 [0.307]	0.166 [0.245]
Log computers per capita	0.228 [0.210]	-0.204 [0.234]	0.107 [0.162]
Lag Internet users	0.539*** [0.049]	0.719*** [0.082]	0.622*** [0.047]
Lagged language network	0.08 [0.161]	0.239*** [0.075]	0.237*** [0.056]
Constant	0.11 [0.073]	-0.019 [0.050]	-0.004 [0.045]
Method of estimation	IV	IV	IV
Fixed effects	Yes	Yes	Yes
Number Observations	358	380	738

Notes: The dependent variable is the log of Internet users per capita. For each country i , the lagged language network is the logarithm of the total users in other countries (other than i) with the same official language, per capita. The robust standard errors are in brackets.

*, **, and *** indicate significance at the 10, 5, and 1 percent levels respectively.

The regression uses countries with GDP per capita growth in the interval [-4.9 , 9.2 percent] and controls for country fixed effects.

6. Conclusions

This paper provides a detailed empirical study of the process of adoption and diffusion of the Internet in a large sample of countries for the period 1990-2004. In the first part of the paper we show that this process is well characterized by an S-shape pattern. This is true if we use the world aggregate or if we consider different groups of countries based on their income. The slope of this pattern is positively related to the level of development of each group of countries although, due to the nature of the curve, with time less developed countries adopt Internet faster than more developed ones.

The second part of the paper explores the main determinants of Internet diffusion and finds that national network effects (measured as the lag of the number of users in a given country) are a crucial determinant of the spread of Internet. This network effect is stronger in more developed countries than in developing ones, implying that both groups are facing different stages of Internet adoption. We explored potential differences in this process across time and we found that low and lower middle income countries are following the path made by the upper middle and high income countries.

A positive reading of our results is that less developed countries are, although very slowly, converging with more developed ones, in terms of Internet usage. The most clear policy implication of our findings is that, in order to help narrow the Digital Divide, policymakers should emphasize policies that liberalize the telecommunications markets in less developed countries. Our results suggest that increasing the number of Internet providers has a substantial positive effect on the spread of the Internet.

References

- Arellano, M., and S. Bond (1991), "Some tests of specification for panel data," *Review of Economic Studies*, 58.
- Barro, R. J., and X. Sala-i-Martin (1997), "Technological Diffusion, Convergence, and Growth," *Journal of Economic Growth*, 2: pp.1-27.
- Canning D. (1999), "Internet use and telecommunications infrastructure." CAER II Discussion Paper 54, December.
- Caselli, F., and W. J. Coleman II (2001), "Cross-country Technology Diffusion: The Case of Computers." *American Economic Review*, vol. 91, No. 2, Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association, May, pp. 328-335.
- Chinn, M. D., and R. W. Fairlie (2007), "The Determinants of the Global Digital Divide: A Cross-Country Analysis of Computer and Internet Penetration." *Oxford Economic Papers* 59, pp. 16-44.
- Chong, A., and A. Micco (2003), "The Internet and the ability to innovate in Latin America." *Emerging Markets Review* 4, 53-72.
- Comín, D., and B. Hobijn (2003), "Cross-Country Technology Adoption: Making the Theories Face the Facts." paper prepared for the Carnegie Rochester Conference on Public Policy.
- Comín, D., Hobijn, B., and E. Rovito (2006), "Five facts you need to know about technology diffusion." NBER working paper 11928.
- Dasgupta, S., Lall, S., and D. Wheeler (2001), "Policy Reform, Economic Growth, and the Digital Divide: An Econometric Analysis." World Bank Working Paper, No. 2567, March.
- Davies, S. (1979), *The Diffusion of Process Innovations*, Cambridge, Cambridge University Press.

- De Long, J. B. (1988), "Productivity Growth, Convergence, and Welfare: Comment." *American Economic Review* 78: 5, December, pp. 1138.
- Dewan S., D. Ganley, and K.L. Kraemer (2004), "Cross the Digital Divide: A Cross-Country Analysis of the Determinants of IT Penetration", Personal Computing Industry Center, Graduate School of Management, University of California, Irvine.
- Estache, A., M. Manacorda, and T. M. Valletti (2002), "Telecommunications, Reform, Access Regulation, and Internet Adoption in Latin America," *Economía*, Volume 2, Spring.
- Goolsbee, A. and P. J. Klenow (2002), "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *Journal of Law and Economics*, vol. XLV, October.
- Gort, M. and S. Klepper (1982), "Time Paths in the Diffusion of Product Innovations," *The Economic Journal*, 92, September, 630-653.
- Grace, J., C. Kenny, and C. Zhen-Wei Qiang (2004), "Information and Communication Technologies and Broad-Based Development. A Partial Review of the Evidence," World Bank Working Paper No 12, Washington D.C.
- Gramlich, E. M. (1994), "Infrastructure Investment: A Review Essay," *Journal of Economic Literature*, Vol. XXXII, pp. 1176-1196
- Griliches, Z. (1957), "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, October, 25, pp. 501-522.
- Guillén, M. F. and Suárez, S. (2005), "Explaining the Global Digital Divide: Economic, Political and Sociological Drivers of Cross-National Internet Use," *Social Forces*, Volume 84, Number 2, pp. 681-708.
- International Telecommunication Union (2006), "World Telecommunication Indicators/ICT indicators database," Geneva, Switzerland.
- Jensen, M. (2005), <http://www3.sn.apc.org/africa/afrmain.htm>
- Jovanovic, B., and S. Lach (1989), "Entry, Exit, and Diffusion with Learning by Doing," *American Economic Review*, vol. 79, No. 4, September, pp. 690-699.
- Keller, W. (2001), "International Technology Diffusion," NBER working paper 8573.
- Kenny, C. (2003), "The Internet and Economic Growth in Less-developed Countries: A Case of Managing Expectations?," *Oxford Development Studies*, Vol. 31, No. 1, 99-113.

- Kiiski, S., and M. Pohjola (2002), "Cross-country diffusion of the Internet," *Information Economics and Policy*, 14, pp. 297-310
- Klobas, J. E. and L.A. Clyde (1998), "Learning to use the internet in a developing country: validation of a user model," *Libri*, Vol. 48, pp. 163-175
- Kotler, P. (1986), *Principles of Marketing*, 3rd Edition, Englewood Cliffs, NJ: Prentice Hall
- Mansfield, E. (1961), "Technical Change and the Rate of Imitation," *Econometrica*, Vol. 29, No. 4, October, pp. 741-766.
- Martinez, J. (1999), "The Net in Cuba," *Matrix News*, Vol. 1, No. 1, January
- Norris, P. (2000), "The Global Divide: Information Poverty and Internet Access Worldwide," paper prepared for the Internet Conference at the International Political Science World Congress in Quebec City.
- OECD (2001), *Understanding the Digital Divide*, Organization for Economic Cooperation and Development, Paris.
- Pohjola, M. (2003), "The Adoption and Diffusion of ICT across Countries: Patterns and Determinants," *The New Economy Handbook* (Academic Press)
- Press, L. (2000), "The State of the Internet: Growth and Gaps," http://www.isoc.org/inet2000/cdproceedings/8e/8e_4.htm#_ftn4
- Quibria, M. G., S.N. Ahmed, T. Tschang, and M. L. Reyes-Macasaquit (2002), "Digital Divide: Determinants and Policies with Special Reference to Asia," Economics and Research Department Working Paper No. 27, Asia Development Bank, Manila.
- Röller, H., and L. Waverman (2001), "Telecommunications Infrastructure and Economic Development: A Simultaneous Approach," *American Economic Review*, 91 (4), 909-23.
- Sanchez-Robles, B. (1998), "Infrastructure Investment and Growth: Some Empirical Evidence," *Contemporary Economic Policy*, Vol.XVI, pp. 98-108.
- Shy, O. (2001), *The Economics of Network Industries*, Cambridge University Press.
- Stoneman, P. (1983), *The Economic Analysis of technological Change*, Oxford, Oxford University Press.

Wallsten, S. (2005), "Regulation and Internet Use in Developing Countries. Economic," *Development and Cultural Change*. 501-523.

World Bank (1994), *World Development Report, Infrastructure for Development*, Washington, D.C.

Zhen-Wei Qiang, C., and A. Pitt (2004), "Contribution of information and Communication Technologies to Growth," World Bank Working Paper No 24, Washington D.C.

Appendix 1:

Year	Observations	Mean	Standard Dev.	Min	Max
1990	214	12337	137204	0	2000000
1991	214	20185	206316	0	3000000
1992	214	32483	310060	0	4500000
1993	214	47021	414435	0	6000000
1994	214	96232	894788	0	13000000
1995	214	184903	1720460	0	25000000
1996	214	345366	3107373	0	45000000
1997	214	547334	4209711	0	60000000
1998	214	853528	5968722	0	84600000
1999	214	1287533	7453445	0	102000000
2000	213	1832099	9376954	0	124000000
2001	204	2405333	11300000	0	143000000
2002	201	3082273	13200000	1000	159000000
2003	200	3602316	14100000	1400	162000000
2004	201	4249475	16200000	1600	185000000
2005	161	4859031	13400000	2000	111000000

Year	Observations	Mean	Standard Dev.	Min	Max
1990	210	0.024	0.11	0	0.804
1991	211	0.052	0.217	0	1.404
1992	213	0.095	0.345	0	2.21
1993	213	0.144	0.463	0	2.775
1994	213	0.288	0.893	0	6.747
1995	213	0.624	1.75	0	13.714
1996	213	1.203	2.923	0	18.211
1997	213	2.118	4.546	0	27.487
1998	213	3.451	6.597	0	36.329
1999	213	5.451	9.591	0	53.818
2000	213	8.078	12.164	0	59.786
2001	204	10.546	14.424	0	59.93
2002	200	13.107	16.355	0.004	64.792
2003	200	15.275	17.612	0.053	67.474
2004	200	17.667	19.598	0.079	77
2005	152	20.668	21.356	0.1	87.755

Appendix 2: World Bank income classification

Source: The World Bank website as of April 2007. Link: <http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20420458~menuPK:64133156~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html>

Low-income countries: Afghanistan, Bangladesh, Benin, Bhutan, Burkina Faso, Burundi, Cambodia, Central African Republic, Chad, Comoros, Congo (Dem. Rep), Cote d'Ivoire, Eritrea, Ethiopia, Gambia, Ghana, Guinea, Guinea-Bissau, Haiti, India, Kenya, Korea (Dem Rep.), Kyrgyz Republic, Lao PDR, Liberia, Madagascar, Malawi, Mali, Mauritania, Mongolia, Mozambique, Myanmar, Nepal, Niger, Nigeria, Pakistan, Papua New Guinea, Rwanda, Sao Tome and Principe, Senegal, Sierra Leone, Solomon Islands, Somalia, Sudan, Tajikistan, Tanzania, Togo, Uganda, Uzbekistan, Vietnam, Yemen, Zambia, Zimbabwe.

Lower-middle-income countries: Albania, Algeria, Angola, Armenia, Azerbaijan, Belarus, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Cameroon, Cape Verde, China, Colombia, Congo (Rep), Cuba, Djibouti, Dominican Republic, Ecuador, Egypt, El Salvador, Fiji, Georgia, Guatemala, Guyana, Honduras, Indonesia, Iran, Iraq, Jamaica, Jordan, Kazakhstan, Kiribati, Lesotho, Macedonia, Maldives, Marshall Islands, Micronesia, Moldova, Morocco, Namibia, Nicaragua, Paraguay, Peru, Philippines, Sri Lanka, Suriname, Swaziland, Syrian Arab Republic, Thailand, Tonga, Tunisia, Turkmenistan, Ukraine, Vanuatu, West Bank and Gaza.

Upper-middle-income countries: American Samoa, Argentina, Barbados, Belize, Botswana, Chile, Costa Rica, Croatia, Czech Republic, Dominica, Equatorial Guinea, Estonia, Gabon, Grenada, Hungary, Latvia, Lebanon, Libya, Lithuania, Malaysia, Mauritius, Mayotte, Mexico, Northern Mariana Islands, Oman, Palau, Panama, Poland, Romania, Russian Federation, Seychelles, Slovak Republic, South Africa, St. Vincent and the Grenadines, Trinidad and Tobago, Turkey, Uruguay, Venezuela.

High-income countries: Andorra, Antigua and Barbuda, Aruba, Australia, Austria, Bahamas, Bahrain, Belgium, Bermuda, Brunei Darussalam, Canada, Cayman Islands, Cyprus, Denmark, Finland, France, French Polynesia, Germany, Greece, Greenland, Guam, Hong Kong (China), Iceland, Ireland, Israel, Italy, Japan, Korea (Rep), Kuwait, Liechtenstein, Luxembourg, Macao (China), Malta, Monaco, Netherlands, Netherlands Antilles, New Caledonia, New Zealand, Norway, Portugal, Puerto Rico, Qatar, San Marino, Saudi Arabia, Singapore, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States, Virgin Islands (U.S.).

Appendix 3: Robustness Analysis

This subsection presents robustness checks to confirm the validity of the results presented in the paper.

First, we analyze the impact of outliers on the estimates of our models. Table A3 shows the results of regressions for the world and for different countries by income category where the condition that per capita GDP growth lies within the interval (-4.9, 9.2) has been relaxed. Second, we test whether changing the definition of “sufficiently high competition” (the cutoff value of 5 ISP) changes the impact of the degree of competition on Internet diffusion. Table A4 presents the test of the impact of competition using two alternative definitions: a number of ISPs larger than 4 (the median number of ISPs for low and lower middle income countries) and a number of ISPs larger than 17 (the median number of ISPs for higher middle and high income countries). Finally, we study the robustness of our results using the number of Internet subscribers instead of users as a dependent variable (Table A5).

As can be seen in the tables just described, the results of the key coefficients are extremely similar to the ones presented in the main section of the paper. Including outliers, changing the definition of competition, or using the number of subscribers instead of users to assess the diffusion of Internet does not change any of our results. In all the specifications it is still the case that network effects are identified as the main determinant of the spread of Internet.

Table A3: World diffusion and diffusion by income category including GDP growth outliers

	Low and Lower Middle Income		Upper Middle and High Income		All	
	[1]	[2]	[3]	[4]	[5]	[6]
Log real GDP per capita	1.634*	-0.146	0.629	0.607	1.761***	0.115
	[0.946]	[0.667]	[0.619]	[0.475]	[0.508]	[0.456]
Log real cost	-0.086	0.044	-0.468***	-0.032	-0.151*	0.021
	[0.111]	[0.051]	[0.120]	[0.076]	[0.079]	[0.047]
Log lines per capita	2.647***	0.087	1.412***	-0.264	2.125***	-0.082
	[0.355]	[0.276]	[0.357]	[0.302]	[0.266]	[0.223]
Log computers per capita	2.019***	0.226	2.823***	-0.114	2.333***	0.138
	[0.173]	[0.193]	[0.140]	[0.246]	[0.121]	[0.162]
Lag Internet users		0.584***		0.810***		0.673***
		[0.054]		[0.085]		[0.046]
Constant	-0.146	0.154**	-4.398	0.055	-7.892*	0.106***
	[7.614]	[0.062]	[6.363]	[0.039]	[4.735]	[0.037]
Method of estimation	GLS	IV	GLS	IV	GLS	IV
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number Observations	503	388	510	419	1013	807
R ²	0.86	-	0.91	-	0.91	-

Notes: The dependent variable is the log of Internet users per capita. The robust standard errors are in brackets.

*, ** and *** indicate significance at the 10 percent, 5 percent and 1 percent levels respectively.

The regression controls for country fixed effects.

Table A4: Impact of competition on Internet diffusion using alternative definitions of competition

	Low and Lower Middle Income		Upper Middle and High Income	
	[1]	[2]	[3]	[4]
Log real GDP per capita	1.096 [0.882]	0.294 [0.828]	0.457 [0.635]	0.969 [0.658]
Log real cost	0.022 [0.061]	0.004 [0.064]	-0.058 [0.078]	0.013 [0.078]
Log lines per capita	0.082 [0.304]	0.174 [0.261]	-0.158 [0.329]	0.009 [0.314]
Log computers per capita	0.179 [0.203]	0.196 [0.192]	-0.092 [0.258]	0.053 [0.234]
Lag Internet users	0.549*** [0.058]	0.549*** [0.052]	0.656*** [0.096]	0.685*** [0.073]
Lag Internet users * $I_{ISP \geq 4}$	-0.005 [0.048]		0.124* [0.064]	
Lag Internet users * $I_{ISP \geq 17}$		0.014 [0.038]		0.111*** [0.032]
Constant	0.164*** [0.063]	0.166*** [0.053]	0.074* [0.038]	0.042 [0.039]
Method of estimation	IV	IV	IV	IV
Fixed effects	Yes	Yes	Yes	Yes
Number Observations	356	356	406	406

Notes: The dependent variable is the log of Internet users per capita. The robust standard errors are in brackets.

*, ** and *** indicate significance at the 10 percent, 5 percent and 1 percent levels respectively.

The regression uses countries with GDP per capita growth in the interval [-4.9 , 9.2] and controls for country fixed effects.

Table A5: World diffusion and diffusion by income category using the number of subscribers

	Low and Lower Middle Income		Upper Middle and High Income		World	
	[1]	[2]	[3]	[4]	[5]	[6]
Log real GDP per capita	2.61 [1.605]	0.035 [1.137]	3.459** [1.717]	2.000 [1.703]	5.818*** [1.201]	0.612 [1.021]
Log real cost	0.128 [0.148]	-0.016 [0.047]	-0.168 [0.192]	-0.032 [0.112]	0.06 [0.130]	-0.002 [0.049]
Log lines per capita	2.725*** [0.459]	0.662* [0.346]	1.417 [0.993]	1.059* [0.569]	1.409*** [0.377]	0.877** [0.383]
Log computers per capita	0.792*** [0.260]	0.036 [0.332]	2.601*** [0.336]	0.118 [0.228]	1.476*** [0.285]	0.026 [0.277]
Lag Internet subscribers		0.605*** [0.108]		0.410** [0.167]		0.516*** [0.135]
Constant	-12.872 [13.080]	0.007 [0.052]	-31.469* [17.072]	0.081 [0.069]	-46.825*** [10.845]	0.05 [0.062]
Method of estimation	GLS	IV	GLS	IV	GLS	IV
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number Observations	287	168	250	153	537	321
R ²	0.91	-	0.91	-	0.95	-

Notes: The dependent variable is the log of Internet subscribers per capita. Robust standard errors in brackets.

*, ** and *** indicate significance at the 10 percent, 5 percent and 1 percent levels respectively.

The regression uses countries with GDP per capita growth in the interval [-4.9 , 9.2] and controls for country fixed effects.