

Intergenerational Effects in the Diffusion of New Technology: The Case of Mobile Phones

Jukka Liikanen^{*},

Paul Stoneman⁺

and

Otto Toivanen⁺⁺

Abstract

Many new technologies exhibit clear generational changes. The empirical literature on technology diffusion traditionally analyses the spread of new technologies generically. We use data from the mobile phone industry, where first-generation (1G) and second-generation technologies (2G) can be clearly identified, to analyze the role of generational effects in diffusion. The results from a generation specific approach differ significantly from those of a generic model. There are positive within-generation network effects. 1G (2G) has a positive (negative) effect on 2G (1G) diffusion. Both generations are substitutes for fixed phones. Effects of competition and payment schemes are analyzed.

^{*} Helsinki School of Economics, and Mobirex Ltd, ⁺ Warwick Business School, ⁺⁺ Helsinki School of Economics, and ETLA, the Research Institute of the Finnish Economy. Corresponding author: Otto Toivanen, Department of economics, Helsinki School of Economics, PO Box 1210, FIN 00101, Helsinki, FINLAND. Email: toivanen@hkkk.fi. Email Jukka Liikanen: jukka.liikanen@mobirex.com. Email Paul Stoneman: paul.stoneman@warwick.ac.uk. We would like to thank the editor (Denicoló), two anonymous referees, and seminar participants in Oulu for comments. The authors gratefully acknowledge financial support from the Yrjö Jahnesson foundation. This research is part of the wireless communication research program (brie-etla.org) of BRIE, the Berkeley Roundtable on the International Economy at the University of California at Berkeley, and ETLA, the Research Institute of the Finnish Economy. Financial support from Nokia and Tekes is gratefully acknowledged. All opinions expressed are those of the researchers. The usual caveat applies.

1. Introduction

Technical change can occur both as discrete steps and in a more continuous fashion. For some products or technologies, clear generations can be identified. Mobile phones are one example of such a technology: one generational shift, from analog to digital (1G to 2G) has taken place, and we are in the early stages of the next one. This paper analyzes mobile phone diffusion during a period when the generational shift was under way. The tradition in the diffusion literature has largely been to either ignore such generational changes or to take them into account only as represented by associated changes in the price or quality of the generic technology being diffused. Karshenas and Stoneman (1995, pp. 266),¹ suggest that there is no other reason than tradition for proceeding in this way. The outcome of this is that the (empirical) literature largely has not explicitly considered the effect that an old generation (s) of a technology (the stock of it) will have on the diffusion of new generation (s) of that technology and vice versa. Such inter-generational effects can be of crucial importance to entire industries: think for example of European telecommunications operators' huge investments into the third generation of mobile-phone technology. The presumption there seems to be that despite high penetration rates of second generation (2G) mobile phones, diffusion of third generation (3G) will be swift. There seems to be little basis in the literature to support such a presumption.

In particular, employing an international data-set covering 80 countries for up to seven years per country, in this paper we illustrate the importance of an explicit consideration of generational change by comparing the results from a diffusion model that uses the standard approach, neglecting technological generations, with estimates obtained by estimating generation specific diffusion models. In doing this we place emphasis upon one important characteristic of mobile phones: they are network

goods. Our specific interest is in the network externalities within and between generations, and between an obvious substitute that shares at least some of the network effects: fixed line telephony. The case of mobile telephony has some parallels to the famous QWERTY example (David, 1985, Liebowitz and Margolis, 1990), and the well documented standard battles in video recorders. The outcome however seems rather different for in this case the market has not got locked into an inferior technology.

Existing studies of mobile phone diffusion (see in particular Gruber and Verboven 2001a,b, henceforth GV; Dekimpe et al., 1998) utilize the standard epidemic model often found in the empirical literature on technology diffusion (surveyed recently by Geroski, 2000, and Stoneman, 2001). Other papers (Katz and Aspden, 1997) employ survey data. This paper follows GV in utilizing standard diffusion models, but we differ from the aforementioned studies in the following respects: first, we measure network effects directly (for a recent survey on network effects, see Gandai, 2002);² second, we make a distinction between 1G and 2G mobile phones; third, we use more information on technology than previous studies; and finally, our estimation methods differ significantly from those used in these other studies.

First and second mobile phone generations, while offering the same basic product, voice transfer, differ significantly from each other technologically and otherwise, as we will detail in the next section. They therefore are imperfect substitutes. Also, although one could argue that all phones, mobile and fixed, share the same direct externality of network size, indirect network effects such as lower prices due to economies of scale in production may be specific to technologies and/or generations. Similarly, mobile phone services vary from country to country even

though the base product is again the same. Governments make different institutional choices, e.g. the number of licenses, or whether the licenses are nationwide or local. Also, contractual terms offered to customers differ between countries. In some countries, for example, the so-called calling-party-pays system is in use. Such issues are discussed further in the following section. The data is presented in section 3, and section 4 is devoted to specifying the econometric model. Econometric results are presented in section 5, and section 6 provides conclusions.

2. Mobile Phone Generations

2. 1. Network Effects Within and Between Generations

When two generations of a technology co-exist, we may expect that they affect the diffusion of each other. In Figure 1, we plot the within-sample diffusion curves for 1G and 2G (see below on data sources and samples).³ The 1G curve approximates the traditional sigmoid curve documented in numerous diffusion studies. By the time 2G was introduced, diffusion of 1G had clearly levelled off. The 2G curve, on the other hand, has a very steep slope up to the end of our observation period. Figure 1 neatly illustrates how the relative diffusion of different generations can be very different at different points in time. As the steep increase in the number of digital phones coincides with the levelling off and decline in the diffusion of analog phones, it also suggests that the diffusions of the two different generations are likely to be interdependent.

[FIGURE 1 HERE]

Inter-generational diffusion effects may come either directly from the competition between goods in a market, or more indirectly from the impact of accumulated stocks. In the absence of stock effects, the decision of what type of phone to adopt is driven

by the standard economics of demand for differentiated goods (see e.g. Verboven, 1996). As digital phones in most aspects technically dominate analog phones (i.e. they are superior in a vertically differentiated sense), it is to be expected that for a given price and in the absence of network externalities, consumers would purchase digital rather than analog phones.

The traditional literature on technology diffusion has emphasised the role that accumulated stocks play in the diffusion of generic technologies. The earlier literature often incorporates information spreading effects arising from the growing familiarity with a technology that comes from use (e.g. Geroski, 2000, Stoneman, 2001). Thus, as more individuals use mobile phones, the number of people that know of their existence, performance characteristics, and costs of operation will increase. This would suggest a positive link between mobile phone generations, as a larger stock of 1G adopters means better informed consumers and increased demand for 2G phones.

More recent literature emphasises the role of network externalities (Katz and Shapiro, 1985, Farrell and Shapiro, 1985, Arthur, 1989, Choi, 1997, Choi and Thum, 1998). The essence of network externalities in this context is that the benefits to phone users increase as the network size increases. As one can use phones of any one standard, whether mobile or fixed, to call a phone using another standard, one is inclined to think that such network externalities will be shared between all phones. However, some features such as text (SMS) messages and picture transmission are generation specific. It may therefore be the case that although there are system wide network externalities that affect 1G, 2G and fixed line phone demand, there exist also network externalities that are specific to mobile phones, or to a particular mobile phone generation.⁴ The general positive effect from a larger installed base would

suggest a positive inter-generational effect. These positive externalities we subsequently label network effects.

On the negative side old-generation specific network externalities may weaken the relative appeal of the new generation due to a larger installed base. Also, the higher is the penetration rate of the old technology, the larger is the degree to which demand for the phones of the new generation is “replacement” demand, again meaning that the degree of diffusion of the old technology may slow down the diffusion of the new technology (in that the acquisition decision for the new technology will be based upon the additional benefits of changing from 1G to 2G rather than the absolute benefits of 2G versus no mobile phone). These negative externalities we subsequently label substitution effects.

If the positive effects dominate, a higher degree of 1G penetration should speed 2G diffusion; if the negative ones dominate, the effect will be reversed. In our empirical analysis, we concentrate on what effects dominate. This issue also has relevance for the next generational shift. The oncoming 2.5G and 3G mobile phones will be technologically superior to 2G,⁵ but will not enjoy generation-specific network externalities initially. For market forecasting purposes, it is thus necessary to have some indication as to the importance of such externalities in the take up of a new generation.

2.2. *Standards and Intergenerational Differences*

Different mobile phone generations are easily identifiable by the technology they use. Similarly, all mobile phone standards can be allocated into a technological generation. The first public mobile communication systems appeared in the US in the late 40's and in Scandinavia in the early 70's. These relied on manually switched calls

and used large and heavy radio transmitters in vehicles. Therefore, they are usually viewed as predecessors to modern mobile communication technologies rather than a part of them. They were followed by analog technologies that featured automatic dialing to and from external networks. Various standards for these technologies were designed simultaneously in different countries, with the US and Japan being the first countries to adopt an analog (1G) standard in 1978 (Advanced Mobile Phone System, AMPS, in the US) and 1979 (Nippon Telegraph and Telephone, NTT, in Japan), respectively. While AMPS developed into a dominant standard in the Americas, the Japanese NTT (NTT's proprietary standard) was never adopted in any other country. Scandinavian countries all adopted a common (Nordic Mobile Telephones, NMT) standard in the early 80's. This standard was to be adopted by several other countries over the next fifteen years. In total, eight different 1G standards were adopted in at least one country (see Table 1). Whilst the analog standards made wide-spread use of mobile phones possible, technical features set constraints on their applicability. The early analog phones were large and heavy and were better suited to communication from a base such as a car rather than for mobile communication as we currently understand the word. In addition, analog technologies use radio spectrum inefficiently, and as the technology spread, capacity constraints were met in many countries. In practice this meant that customers could not necessarily make calls at the time they wanted.

Even early in the 1980's, it was realized that the next (2G) generation of mobile phones would be based on digital technology. The main benefits of digital over analog technology are a more efficient use of radio spectrum, and clarity of voice. European countries coordinated their standards choice on GSM in the 80's, several years before the first GSM network was opened (in Finland in 1991). Table 1 illustrates the

diffusion of different 1G and 2G standards over countries, and the extent of their adoption in terms of numbers of subscribers. One can see that of the analog standards, the AMPS standard (originating in the US) was most widely in use by late 90's both in terms of number of countries and number of subscribers but with 2G the GSM standard was most widely used. What is also apparent from this data is that analog mobile phones are being replaced by digital phones: for example, the number of countries that have an NMT network is declining, as is the number of NMT subscribers.

[TABLE 1 HERE]

2G introduction was also often accompanied by other measures. Most notably, several countries who did not introduce competition in 1G, but only issued licenses for the existing (government owned) telecom-monopolies, licensed competitive suppliers with 2G (see e.g. GV). Whereas in 1991, there were 56 monopoly markets out of a total of 69, in 1998 there were 90 monopoly markets out of 177 markets (EMC figures). Similarly, contracting terms offered to customers evolved over time. The two most critical ones were the so-called "calling-party-pays" practice, and the introduction of prepaid cards. The latter have proved especially important in countries that lack credit institutions. In our empirical analysis we take account of both competition and contracting features.

3. Data Sources and Diffusion Patterns

Our data come mainly from two sources: the data on mobile phones comes from EMC.⁶ Our demographic and economic data comes from World Development Indicators (WDI) of the World Bank. These are currently only available until 1998, which therefore determines the observation period for us. As we are interested in the

generational shift that has taken place, we condition our data on the existence of both generations. In other words, our sample consists of those country-year observations for which we know that both 1G and 2G were available for consumers. From a potential sample of 80 countries each with a maximal seven-year observation period we are left with 316 observations, after deleting observations for which we do not have data on explanatory variables. Table 2 presents sample descriptive statistics on all the variables used in the econometric analysis below. In particular we observe the number of adopters of analog, digital, and fixed line telephones at the end of each calendar year. In addition, we know which mobile phone standards are in use in each of the countries (although we only use information on whether the most widely used standard is in use in a particular country or not).

[TABLE 2 HERE]

The data indicates that although mobile phones as a generic technology have diffused fast during the 90's, their penetration ratio still lags far behind those of fixed line phones. This is illustrated in Figure 2, which shows histograms of the penetration ratios of all three phone technologies in our sample countries in 1998 (1G, 2G, mainline: the bin width is five percentage points). At that point in time, 1G penetration ratios were already declining in a number of countries, 2G penetration rates were increasing fast, and main line penetration ratios were much more stable in comparison, but at a higher level. In the vast majority of countries, 1G penetration rates are below five per cent, with a maximum of 19%. 2G penetration ratios are much more widely distributed, with a maximum of 50%. This statement holds even more strongly for fixed line telephones, where the maximum penetration ratio is as high as 67.5% (penetration ratios are calculated in relation to population). We have calculated the sample correlations of the 1998 penetration ratios. Analog and digital

penetration ratios are positively correlated, but the coefficient is only 0.23. The correlation between analog and mainline penetration ratios is also positive, at 0.44, and that between digital and mainline is 0.78. These values suggest the presence of positive network externalities between mobile phone generations and mobile and fixed line telephone systems. In contrast, GV found a negative relationship between mainline (fixed) network size and diffusion speed of mobile phones.

[FIGURE 2 HERE]

The most widely used analog standard (in terms of number of country-year observations) is NMT, used in 40.8% of our observations, and 28 out of 80 countries in 1998 use NMT.⁷ GSM is the most widely used digital standard. It was used in 84.5% of our observations – and 63 out of 80 countries in 1998.

4. Econometric Models

We concentrate on estimating nonlinear diffusion models (see e.g. Zettelmeyer and Stoneman, 1995). Some recent papers on international diffusion of technology (e.g. Caselli and Coleman, 2001) have estimated a simple linear model, where the dependent variable is the number of adopters of technology j in country i at the end of period t . As a specification check, we estimate variants of this approach, and comment on the results briefly.⁸ In line with GV, we take as our starting point the standard epidemic diffusion model (see e.g. Geroski, 2000). In that model, one posits that the number of adopters of technology j in country i at the end of period t is affected by three factors: the number of potential adopters N_t^{ji} ; the number of actual adopters in the previous period, and a multiplier. The multiplier tells us the proportion of those who could potentially adopt, but have not by beginning of period t , who will adopt during that period. This leads to the following equation:

$$(1) \quad S_t^{ji} = \lfloor g(\cdot)(N_t^{ji} - S_{t-1}^{ji}) + S_{t-1}^{ji} \rfloor \Delta t.$$

In (1), N_t^{ji} is the number of potential adopters of technology j in country i at (end of) year t ; $g(\cdot)$ is the learning coefficient that measures the strength of the epidemic effect (taking on values in the unit interval; for a detailed discussion, see Geroski, 2000).

The literature on mobile phone diffusion (GV) has used a version of (1) where Δt is allowed to go to zero. We assume that $\Delta t = 1$, where one refers to a calendar year.⁹

Rearranging (1), and assuming that $N_t^{ji} = POP_t^i f(X_t^{ji}, \beta^j)$ and $g = g(t^{ji}, \gamma^j)$ we arrive at our estimating equation:

$$(2) \quad \Delta S_t^{ji} = g(t^{ji}, \gamma^j) [POP_t^i f(X_t^{ji}, \beta^j) - S_{t-1}^{ji}] + \varepsilon_t^{ji}.$$

In (2), POP_t^i refers to the size of the population in country i in year t ; X_t^{ji} are exogenous or predetermined explanatory variables; t is a (country and possibly technology specific) vector of time-based variables; ε_t^{ji} is an i.i.d., zero-autocorrelation error term; and β and γ are (technology specific) parameter vectors to be estimated. The linear and nonlinear models can be used to estimate either generic diffusion, i.e., not accounting for generational differences within a technology, or generation specific diffusion. We will do both.

In (2), we thus postulate that the number of potential adopters is determined, in the case of mobile phones, by population. As mobile phones are used on an individual basis, and the number of people with more than one phone is small especially within our observation period, it seems sensible to assume that population determines an upper bound on N_t^{ji} . We therefore assume that $f(\cdot)$ is a function that takes on values in the unit interval. Similarly, $g(\cdot)$ is a function that takes on values in the unit interval as it measures the proportion of potential adopters who adopt.

We specify our empirical model as follows:

$$(3) \quad g(.) = 1 - \exp(-(\gamma_o + \gamma_1 TIME_t^{ji} + \gamma_2 (TIME_t^{ji})^2)) ,$$

(the exponential distribution) where $TIME_t^{ji}$ is the number of years that technology j has been available in country i , and $TIME_t^{ji2}$ is its square). The idea is that information diffusion to a large extent is population (hence country) specific, and that different generations may necessitate different processes of information diffusion. For the generic model, we specify $TIME_t^{ji}$ as time from first introduction of mobile phones.

We specify that the proportion of population that is a potential adopter of mobile phones is given by

$$(4) \quad f(.) = \frac{1}{1 + \exp(-h(.))}$$

where $h(.)$ is given by (we omit the j superscript from the β coefficients)

$$(5) \quad h(.) = \beta_0 + \beta_1^{j'} NETWORK_t^{ji} + \beta_2^{j'} DEMECON_t^{ji} + \beta_3^{j'} TECHCOST_t^{ji} .$$

In other words, $f(.)$ takes the form of a logistic distribution. The functional forms for $g(.)$ and $f(.)$ have been chosen partly for their analytic simplicity: given our limited sample size, explicit functional forms yield added efficiency. We multiply $h(.)$ by minus one (see equation (4)) in order to have positive coefficients increase diffusion speed. We arrive at our nonlinear specification by inserting (3), (4) and (5) into (2).

The explanatory variables in (5) can be divided into three categories, each consisting of a vector of variables. First, there are variables measuring network effects within and between technologies (NETWORK). These are: (i) the penetration rates of mobiles (pen_M for generic mobile, pen_A for analog phones, and pen_D for digital phones where Pen_M = pen_A+pen_D) (ii) the penetration of the other mobile technology (either pen_A or pen_D) and (iii) penetration of fixed line telephones

(pen_fixed). These are all defined as the number of phones of a given technology, divided by population. We also include their squared values to capture any nonlinearities, and to mitigate the effects of our functional form assumptions.¹⁰ The interpretation of the coefficients of these variables is as follows. The coefficient of the penetration of the own technology measures network externalities within the given mobile phone technology. The coefficient of the other mobile technology's penetration rate measures both substitutability and network externalities between the technologies. We therefore expect the first of these to have a positive effect, and the second either a positive or a negative effect (depending on whether substitutability or network externalities dominate), on the number of potential adopters. For fixed lines, the coefficient measures network effects on the one hand, and substitutability on the other hand. A negative coefficient for fixed line penetration, for example, would imply that network effects are outweighed by the effect of substitutability. An argument can be made for including penetration rates in the $g(\cdot)$ function in addition to having them in $f(\cdot)$: if diffusion of information is not only affected by time, but also by penetration effects, then one should include them into $g(\cdot)$. However, separately identifying these effects would pose great problems. We have therefore opted for including penetration rates into (4) alone.

Second, there are demographic and economic variables (DEMECON). We include the following: GDP per capita (gdpcap) and its square (gdpcap²), the percentage of the population that lives in cities (urb), the proportion of population aged 65 or more (pop65), the age-dependency ratio (agedep), and the geographical area of country i (surface). All but surface are time-varying. The reasons for inclusion of these variables are probably obvious, but a few remarks are in order. The geographic size of a country may have an effect for at least two reasons. First, the

larger the country, the longer the distances, and the more likely it is that people are not reachable by fixed line phones as easily. Second, the larger the country, the more expensive it is to build and maintain a fixed line network.¹¹ The proportion of over-65 year olds and the proportion of urban population are included because older people are generally regarded as less likely to adopt new technologies whereas people in urban areas often are better informed of new technologies, but on the other hand have a lower opportunity cost of using fixed line phones (due to shorter distances).

Finally, the third category (TECHCOST) of variables aims to capture the (relative) cost of adopting and using a mobile phone (of a particular generation). Unfortunately, as so often with data on new technology, we lack data on prices.¹² This is a further reason for estimating reduced form models of this kind. In the absence of prices the cost of using and adopting is measured by the number of standards in use, both within own generation and in the other generation (tech_M, tech_A and tech_D). These are defined as the number of mobile, analog and digital standards, respectively, in use in country i in year t . More standards most likely means that some standard specific network effects (such as economies of scale in production of services) are lost, leading to higher usage costs. For the generic model, we pool the number of analog and digital standards. The variable “split” refers to a situation where digital licenses are not country-wide, but defined over smaller regions. Again, this can lead to lost (indirect) network effects, and thereby to higher marginal costs of production of services of 2G phones. We measure market structure through the 3-firm Herfindahl index, and the number of licenses. The latter corresponds to the number of firms in the market, and therefore the hhi3 coefficient measures how concentration affects prices, conditional on the number of firms. Finally, we include a technology dummy in both 1G and 2G equations ($TECH_t^{ji}$ and $TECH_t^{k \neq ji}$ in equation (6)). For both, we

allow those countries with the most widely adopted standard (in terms of the number of countries using that standard: NMT for 1G, GSM for 2G) to have a different constant term in $f(\cdot)$. This allows for differences in prices and quality of the service that different mobile phone standards yield customers. The variable YEAR (and its square) is included to capture world-wide economies of scale and learning. It is the calendar year minus 1991 divided by seven (to normalize the variable onto the unit interval).

5. Results

In Table 3, we present the results from our nonlinear estimations: column (1) reports results for the generic model, column (2) for 1G, and column (3) for 2G. Recall that now the interpretation of most variables (those in $f(\cdot)$) is as an effect on the proportion of population who are potential adopters. Below, we will discuss the results in two stages. In the first subsection, we discuss all the other variables but those measuring network effects. These are discussed in a separate subsection. In both subsections, we discuss all three models. In the final subsection, we discuss robustness issues.

5.1. Results on Learning, Demographics, and Market and Technology

For the generic model, we find that time in the market has no effect on the learning coefficient. In contrast, YEAR has a positive impact on the share of population of potential adopters, but YEAR² (squared years) no impact. We find that gdp per capita also has a nonlinear effect: the linear term has a positive, the squared term a negative effect on the proportion of population that are potential adopters. Within our sample, the linear effect dominates. Urban population decreases the share of population that

are potential adopters; the share of over 65-year olds, the age-dependency ratio, and the geographical size of a country have no effect.

NMT (1st generation) mobile phones do not affect the share of potential analog phone adopters, nor does the GSM (2nd generation) standard affect the number of digital mobile phone adopters. The number of technologies in use has a negative effect on the share of potential adopters. Calling party pays-contracts and the ability to buy prepaid phones have no effect. We find that the number of licenses has no effect on the proportion of potential adopters, but more highly concentrated markets have a smaller proportion of population belonging to the set of potential adopters. Having split licenses has no effect.

[TABLE 3 HERE]

For 1G (Column (2)), we find that TIME has a positive, but decreasing impact on the proportion of non-adopters that adopt, i.e., on the learning effect. The value of $g(.)$ goes from 63% in the year of digital phone introduction to 100% in the 7th year. We find that the effect of time (YEAR) on the proportion of potential adopters is nonlinear. Gdp per capita has a positive impact, and its square no impact. Urban population and the proportion of pensioners affect the share of potential adopters positively, whereas the age-dependency-ratio has no effect. Countries with a large geographical area experience slower analog diffusion.

Turning to the technology variables, we find that the number of digital and analog standards play no role in the diffusion of analog phones, nor does the number of licenses or having (geographically) split licenses. Offering prepaid phones slows down diffusion, as does having a calling-party-pays system. Both of these findings are surprising, and could reflect the fact that these are more often in use in digital networks. The negative coefficients would then reflect substitution effects between

generations. Somewhat surprisingly, diffusion is faster in more concentrated markets. This could be due to operator-specific (in)direct network effects outweighing any negative effects of concentration on competition. NMT analog phones' diffusion speed does not differ from that of other standards.

On digital diffusion (column (3)), we find that the time that a population has had access to digital technology has no effect on the share of potential adopters who adopt. Of the determinants of the population share of potential adopters, we find that the time that digital technology has been in use worldwide (YEAR) has a positive but nonlinear impact, as does *gdpcap*. The linear *gdpcap* term carries a positive, the squared a negative coefficient. This is in contrast to the analog results, but in line with the generic results. Contrary to what we found for analog phones, the more urban the population, the smaller the share of potential adopters. The share of over 65-year olds and the age-dependency ratio have no effect on the population share of potential adopters; the geographical surface increases it.

As to contracting variables, prepaid decreases the population share of potential adopters, but calling-party-pays increases it. The former is in line with analog results, the latter in contrast with them. In the linear estimation (see Table 4), prepaid carried a negative but insignificant, *cpp* a positive and significant coefficient. An increase in the number of digital or analog technologies has no impact on the population share of potential adopters. Geographical splitting of digital licenses has again no effect on digital diffusion. More licenses leads to a higher population share of adopters, whereas (conditional on the number of firms), concentration has no effect.

5.2. Results on Network Effects

In the generic model, we find evidence for network externalities within mobile phones. In contrast, any network externalities between fixed line phones and mobile phones is outweighed by substitution effects (corroborating GV's result): the positive second order term outweighs the negative linear effect only very close to the sample maximum for fixed line penetration. Generally, the (unreported) linear and nonlinear estimation results are in line with each other (regarding significant coefficients): the one notable exception is the coefficient of the squared gdp per capita variable (significant and negative in nonlinear estimation, insignificant and negative in linear regression).

In the analog model, we find that the squared penetration rate of digital mobile phones has a significant negative effect on the share of population that are potential adopters of analog phones. Thus, substitution effects dominate any network effect. Within analog technology, we find evidence of network effects, as the analog penetration rate and its square have positive effects on the proportion of potential adopters. Notice however that the absolute value of the analog penetration coefficients are much smaller than that of the squared digital penetration rate. Substitution effects dominate between analog and fixed line telephony once the fixed line penetration rate exceeds 0.11 (one third of the sample mean).

We find positive network externalities within digital technology. The coefficients on analog penetration rate and its square suggest a nonlinear relationship. The higher the analog penetration rate, the larger the absolute value of the effect on digital diffusion. Positive between generation-network effects exist once the analog penetration rate exceeds 0.086 (55% of sample mean for analog penetration rate). For land lines, substitution effects dominate.

Our results show that the estimated network externalities are very different when one allows for generational differences. Both in the linear and the nonlinear estimations, the estimated network effect is significantly smaller than that estimated for the digital mobile phones within generation. Some of the variables obtain coefficients of opposite sign for the two generations (e.g. *surface*); for others, the absolute values are very different (e.g. *pen_fixed* and its square).

[TABLE 4 HERE]

5.3. Robustness

In this subsection, we report on various robustness tests. The first issue concerns functional form specification. As discussed above, we use tightly parameterized functional forms due to our limited sample size. The linear model that we have estimated provides a useful check on the results. In unreported regressions, we find that our nonlinear results are confirmed.

The second issue has to do with unobservables or possible omitted variables, and endogeneity of some of our explanatory variables. Despite the fact that we have several control variables, it is still possible that there are country-specific unobservables that are correlated with phone usage, thereby affecting our network externality results. Also, it may be possible that such country-specific unobservables are correlated with some of our control variables, such as *hhi3* and the calling-party pays dummy. We believe it is much less likely that our control variables are correlated with the period-specific deviations from country-level means of the error term. To verify that our results are not affected by correlation between some of the explanatory variables and the error term, we re-ran the nonlinear model excluding the prime suspects for endogeneity (*licence*, *hhi3*, *ccp*, *split*, *tech_A*, *tech_D*) and found

that our results are essentially unchanged. To check the effects of unobservables, we re-ran the linear model using a within estimator (i.e., including country fixed-effects). Again, the results are in line with those we obtain without the control for country-specific unobservables. Unfortunately, there is no simple way to estimate the nonlinear model with country fixed effects. Given the large number of countries (80), we would severely lose on degrees of freedom if we inserted country dummies into either $g(\cdot)$ or $f(\cdot)$; furthermore, it is not clear a priori in which one we should insert the country dummies.

6. Conclusions

The objective of this paper was to shed light on how technology diffusion takes place under a period when two consecutive generations of a technology are on offer. In the industry under study, mobile phones, this is a timely question as the industry is about to enter the next generational shift, having invested heavily in the future (third) generation to come. We find strong evidence that the new generation impedes the diffusion of the old generation, whereas the (penetration rate of the) old generation has a positive effect on the diffusion of the new one, at least once the old generation penetration rate is high enough. We thus find that an “intermediate” technology such as 1G (it was known by early 80’s that 1G was going to be replaced by digital 2G) can have a positive impact on the diffusion speed of a superior (new) technology, and thus substitution effects are dominated by network effects. Thus lock-in into an inferior technology seems not to have been a major problem regarding mobile phones. It clearly emerges that the network effects are largest within technological generations. With regard to fixed line telephony, substitution effects dominate any network effects (as found by GV). Understanding the effects of within- and between-generation

network externalities and substitutability will be of great importance also in the ongoing generational shift from 2G to 2.5G and 3G.

We find that both economic variables and demographics strongly influence the diffusion process, and that even the geographic structure of a country plays a role. Further, contracting features, market structure, and governments' decisions as to licensing (through the number of firms) all influence the diffusion speed of mobile phones. Calling-party-pays system and use of the prepaid-technology enhance diffusion of digital phones, as does an increase in the number of licenses.

References

- Arthur, W. B., 1989, Competing technologies, increasing returns and lock-in by historical events, *Economic Journal*, 99, 116 -31
- Brynjolfsson, E., and Kemerer, C., 1998, Network externalities in microcomputer software: an economic analysis of the spreadsheet market., MIT CISR wp. No. 265.
- Caselli, F., and Coleman, J., 2001, Cross-country technology diffusion: the case of computers, *American Economic Review, Papers and Proceedings*, 91, 2, 328-335.
- Choi, J.-P., 1997, Herd behavior, the penguin effect and the suppression of informational diffusion: an analysis of informational externalities and payoff interdependency, *Rand Joirnal of Economics*, 28(3), 407 -25
- Choi, J.-P. and Thum, M.,1998, Market structure and the timing of technology adopton with network externalities, *European Economic Review*, 42(2), 225 -44
- David, P., 1985, Clio and the economics of QWERTY., *American Economic Review Papers and Proceedings* ,75, 332 - 7
- Dekimpe, M. G., Parket, P. M., and Sarvary, M., 1998, Staged estimation of international diffusion models: An application to global telephone adoption, *Technology Forecasting and Social Change*, 57, 105-132.
- Farrell, J., and Saloner, G., 1985, Standardisation, compatibility and innovation, *RAND Journal of Economics*, 16, 70-83.
- Gandal, N., 1994, Hedonic price indexes for spreadsheets and an empirical test of the network externalities hypotheses, *RAND Journal of Economics*, 25, 1, 160-170.
- Gandal, N., 2000, Compatibility, Standardization, & Network Effects: Some Policy Implications, *Oxford Review of Economic Policy*, 18, 80-91.
- Geroski, P., 2000, Models of Technology Diffusion, *Research Policy*, 29, 4-5 , 603-625.
- Gruber, H., and F. Verboven, 2001a, The evolution of markets under entry and standards regulation: the case of global mobile telecommunications, *International Journal of Industrial Organization*, 19, 7, 1189-1212.
- Gruber, H., and F. Verboven, 2001b, The diffusion of mobile telecommunications services in the European Union, *European Economic Review*, 45, 3, 577-588.
- Karshenas, M., and P. Stoneman, 1995, *Technological Diffusion*, Handbook of the Economics of Technological Change, P. Stoneman (ed.), Edwar Elgar.

- Katz, M., and Shapiro, C., 1985, Network externalities, competition and compatibility, *American Economic Review*, 75, 424-440.
- Katz, J. E., and Aspden, P., 1998, Theories, data, and potential impacts of mobile communications: A longitudinal analysis of US national surveys, *Technology Forecasting and Social Change*, 57, 133-156.
- Liebowitz, S., and Margolis, S., 1990, The fable of the keys, *Journal of Law and Economics*, 1-26.
- Saloner, G., and Shepard, A., 1995, Adoption of technologies with network effects: an empirical examination of the adoption of automated teller machines, *RAND Journal of Economics*, 13, 479-501.
- Stoneman, P., 2001, *The Economics of Technological Diffusion*, Oxford, Basil Blackwell, September.
- Verboven, F., 1996, International price discrimination in the new car market, *RAND Journal of Economics*, 27, 2, 240-268.
- Zettelmeyer, F., and Stoneman, P., 1995, Testing alternative models of new product diffusion, *Economics of Innovation and New Technology*, 2, 283-308.

Table 1

Technology	System	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	
Analog	AMPS	Countries	0	1	2	3	5	9	15	22	27	34	46	52	60	82	91	96	96	96
		Operators	0	7	11	12	14	18	24	31	38	49	63	81	102	181	209	225	229	226
		Subscribers	0	0	0.2	0.6	1.1	2.1	3.5	5.7	9	13	15	22	34	48	66	70	73	70
	NMT	Countries	0	0	0	1	3	5	7	12	16	16	21	21	26	27	27	27	25	21
		Operators	0	0	0	2	4	6	8	15	19	19	24	54	59	60	58	58	57	52
		Subscribers	0	0	0	0	0.1	0.3	0.5	1	1.7	2.4	3.4	5.4	9.5	15	18	16	14	12
	TACS	Countries	0	0	0	1	3	5	7	12	16	16	21	21	26	27	27	27	25	21
		Operators	0	0	0	2	4	6	8	15	19	19	24	54	59	60	58	58	57	52
		Subscribers	0	0	0	0	0.1	0.3	0.5	1	1.7	2.4	3.4	5.4	9.5	15	18	16	14	12
	C-450	Countries	0	0	0	1	2	2	2	3	3	3	3	3	3	3	3	3	3	1
		Operators	0	0	0	1	2	2	2	3	3	3	3	3	3	3	3	3	3	1
		Subscribers	0	0	0	0	0	0.1	0.1	0.2	0.3	0.6	0.8	0.8	0.8	0.7	0.6	0.5	0.4	0.2
Comvik	Countries	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
	Operators	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	
	Subscribers	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
NTT	Countries	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	Operators	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	
	Subscribers	0	0	0	0.1	0.1	0.1	0.2	0.4	0.6	1	1.3	1.6	1.9	2.7	2.1	0.6	0	0	
RC2000	Countries	0	0	0	1	1	1	1	1	1	1	1	1	1	2	2	1	1	0	
	Operators	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
	Subscribers	0	0	0	0	0	0	0.1	0.2	0.2	0.3	0.3	0.3	0.3	0.2	0.2	0.1	0	0	
RTMS	Countries	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
	Operators	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	
	Subscribers	0	0	0	0	0	0	0	0.1	0.1	0.1	0	0	0	0	0	0	0	0	
Digital	GSM	Countries	0	0	0	0	0	0	0	0	0	0	7	16	38	64	94	107	120	131
		Operators	0	0	0	0	0	0	0	0	0	0	14	29	64	124	205	273	327	364
		Subscribers	0	0	0	0	0	0	0	0	0	0	0.2	1.4	5	13	33	71	138	255
	TDMA	Countries	0	0	0	0	0	0	0	0	0	0	1	4	7	16	24	31	37	42
		Operators	0	0	0	0	0	0	0	0	0	0	1	4	9	21	34	49	77	91
		Subscribers	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.7	2.7	6.4	16	35
	CDMA	Countries	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	7	18	26
		Operators	0	0	0	0	0	0	0	0	0	0	0	0	0	1	11	19	36	59
		Subscribers	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	7.4	22	49
	PDC	Countries	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
		Operators	0	0	0	0	0	0	0	0	0	0	0	1	4	5	6	6	6	5
		Subscribers	0	0	0	0	0	0	0	0	0	0	0	0	0.5	3.3	14	27	38	45

LEGEND: Countries = number of countries that have adopted standard k by year t .
Operators = number of mobile phone operators using standard j in year t .
Subscribers = number of subscribers using standard j in year t (in millions).

Table 2
Descriptive Statistics

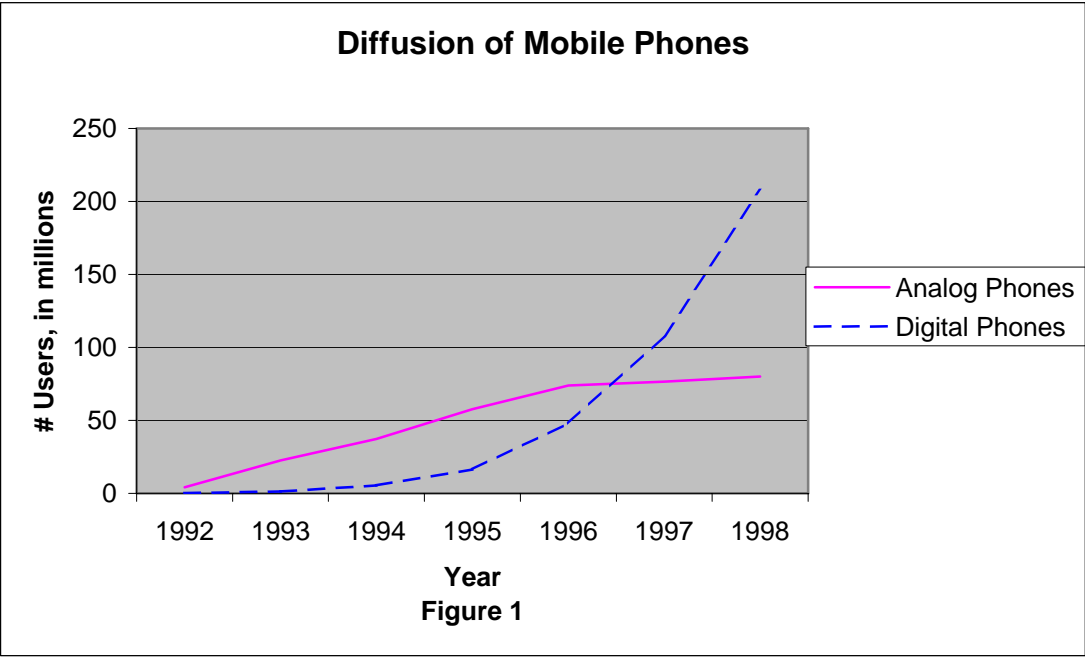
Variable	Mean (Standard Deviation)
Pop. Population (000's).	53469.210 (155987.918)
Urb. Proportion of urban population.	0.665 (0.218)
Pop65. The proportion of over 65-year olds.	0.097 (0.050)
Agedep. Age-dependency ratio.	0.563 (0.128)
Surf. Geographical area (100 000 sq. km).	0.101 (0.235)
YEAR. Calendar year – 1991 divided by 7 in order to normalize onto the unit interval.	0.727 (0.234)
Time (analog). Years from introduction of analog phones.	9.421 (3.540)
Time (digital). Years from introduction of digital phones.	2.880 (1.634)
License. Number of mobile phone licenses.	0.259 (0.140)
HHI3. Herfindahl index of three largest firms (max. 1).	0.620 (0.277)
Split. An indicator variable for geographically split licenses.	0.101 (0.302)
Tech_A. Number of analog standards in use. Normalized by dividing by the maximum number in the sample (3).	0.411 (0.169)
Tech_D. Number of digital standards in use. Normalized by dividing by the maximum number in the sample (3).	0.398 (0.165)
Cpp. Indicator variable for calling party pays-contracts.	0.421 (0.494)
Prepaid. Indicator variable for prepaid contracts.	0.250 (0.434)
Subsat1. Number of analog phone subscribers (000's), lagged by one year.	949.935 (3448.811)
Subsdt1. Number of digital phone subscribers (000's), lagged by one year.	562.717 (2022.969)
Dsubsa. Change in the number of analog phone subscribers from end of year t-1 to end of year t.	156.894 (791.843)
Dsubsd. Change in the number of digital phone subscribers from end of year t-1 to end of year t.	644.420 (1781.731)
Pen_A. Penetration rate of analog phones (# subscribers/population).	0.027 (0.034)
Pen_D. Penetration rate of digital phones (# subscribers/population).	0.026 (0.052)
NMT. An indicator variable for the NMT analog phone standard.	0.408 (0.493)
GSM. An indicator variable for the GSM digital phone standard.	0.845 (0.363)

NOTE: All variables defined on a year-country basis. 316 year-country observations, 80 countries. Number of observations per country varies between one and seven.

Table 3

Non-Linear Model Results			
Variable	Generic	Analog	Digital
Learning coeff. $g(.)$			
Constant	-0.222 (0.270)	-0.042* (0.025)	0.264 (0.226)
iTIME	0.076 (0.063)	0.016*** (0.004)	-0.034 (0.077)
iTIME2	-0.003 (0.003)	-0.001*** (0.0001)	0.002 (0.006)
Population Proportion $f(.)$			
Demographic and Economic Variables			
Constant	-7.103*** (1.289)	-9.591*** (1.795)	-20.401*** (1.705)
Gdpcap	41.662** (18.699)	105.453*** (39.523)	92.637*** (4.809)
Gdpcap2	-86.151** (41.800)	-302.713 (121.722)	-198.804*** (5.190)
Urb	-3.409*** (1.098)	10.030* (5.532)	-5.757*** (1.689)
Pop65	-7.592 (5.633)	20.723** (10.548)	-7.935 (6.597)
Agedep	0.591 (1.455)	-12.592 (8.743)	5.168*** (1.677)
Surface	1.012 (0.685)	-2.720*** (0.953)	3.266*** (1.270)
Technological and Cost Variables			
GSM	0.019 (0.400)	-	-0.430 (0.849)
NMT	0.164 (0.176)	-0.932* (0.538)	-
Tech_M	-0.987*** (0.345)	-	-
Tech_A	-	3.371 (2.463)	0.0978 (0.764)
Tech_D	-	0.456 (1.700)	-0.379 (0.351)
Cpp	0.868 (0.767)	-4.644* (2.479)	4.128*** (0.482)
Prepaid	-0.097 (0.116)	-1.158*** (0.370)	-0.451 ^a (0.278)
License	-0.006 (0.234)	1.367 (1.806)	3.708*** (1.460)
Split	0.436 (0.364)	0.790 (0.550)	0.039 (1.004)
HHI3	-0.820 ^a (0.515)	2.060* (1.234)	0.206 (0.819)
Year	4.488*** (1.796)	12.647*** (2.830)	18.232*** (4.303)
Year2	-0.400 (0.991)	-8.608*** (2.248)	-6.588*** (2.608)
Network Variables			
Pen_M	8.057* (4.960)	-	-
Pen_M2	-13.495 (12.384)	-	-
Pen_A	-	57.551*** (21.452)	-10.618** (4.738)
Pen_A2	-	432.965* (239.181)	123.171*** (3.650)
Pen_D	-	-134.340 (119.390)	23.924*** (6.004)
Pen_D2	-	-11065.236*** (5424.169)	-2.654 (11.317)
Pen_fixed	10.572*** (3.957)	-48.853** (23.268)	-3.341* (1.992)
Pen_fixed2	-15.881*** (3.782)	61.793** (30.877)	-6051.084*** (0.044)
R ²	0.9500	0.8993	0.9632
LogL.	-2357.570	-2194.350	-2291.442

Notes: reported numbers are coefficient and (s.e.). iTIME is generation specific. ***, **, *, and ^a refer to significant coefficients at 1, 5, 10, and 11% level. S.e.'s heteroskedasticity and autocorrelation (1st order) robust.



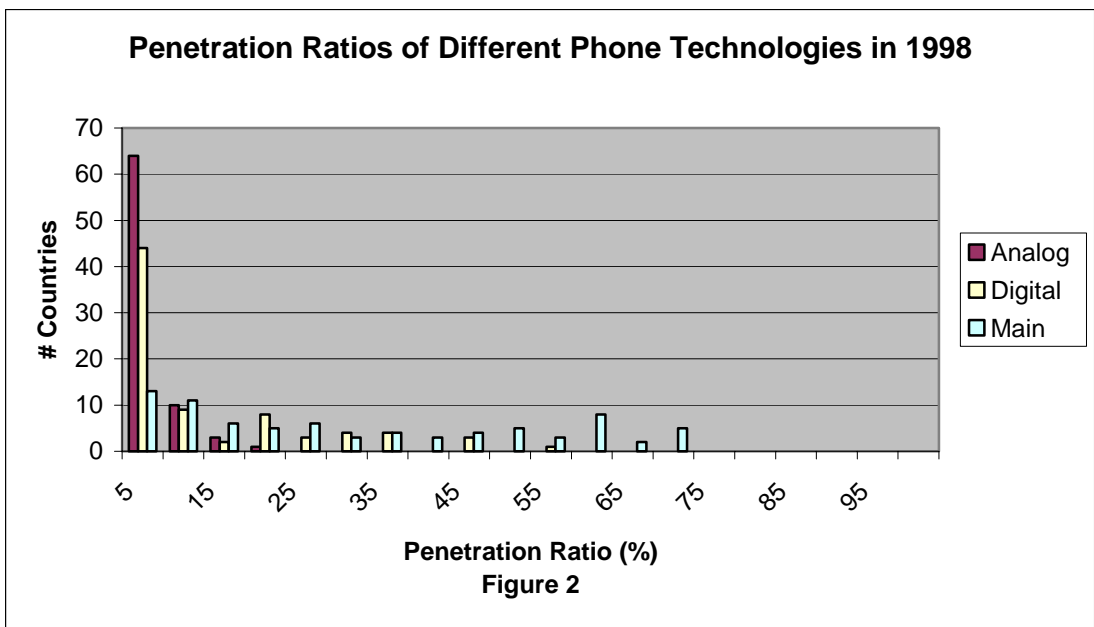


Figure 2

¹ Karshenas and Stoneman (1995) actually raise the possibility of studying technology diffusion generation by generation, taking into account interdependencies.

² Earlier studies that measure network effects empirically include Saloner and Shepard (1995), Gandal (1994) and Brynjolfsson and Kemerer (1998).

³ The world-wide diffusion curves are similar in shape

⁴ A referee has correctly pointed out that there exist provider level network externalities through price differentiation depending on whether the call takes place within the same provider's network or not.

⁵ Of course, the differences between 1G and 2G, and 2G and 3G are not identical.

⁶ See <http://www.emc-database.com/>.

⁷ Note that the difference between the sample and world-wide statistics is explained by the twin facts that 1) there are more countries in the world-wide statistics and 2) countries adopting GSM were fast in adoption, yielding more country-year observations. One cannot therefore compare the sample statistics and Table 1 directly.

⁸ The linear model we use regresses the growth rate of the stock of adopters on the same explanatory variables we use in the nonlinear model. The only exception is that instead of using penetration ratios, we use the numbers of adopters, as that is better in line with a linearization of equation (1).

⁹ Our observations are annual. In other words, we measure the change in the stock of adopters between 31st December in years t and $t-1$. The approach that GV takes presumes that discrete (annual) observations can be used to estimate a continuous time model. Our approach allows an easier incorporation of measures of network externalities.

¹⁰ We experimented with using a more general polynomial form, but LR-tests indicated that we could not reject the shown formulation against more general ones. None of the omitted variables (interaction terms of linear and squared penetration rates) ever obtained a significant coefficient.

¹¹ This is visible e.g. in Finland, where the companies providing fixed line services have started to replace damaged fixed lines with mobile connections in the country side.

¹² Price data would be particularly difficult to obtain and handle in the case of mobile phones. First, there are hundreds of phones available. Second, their prices to consumers vary considerably within and between countries, and within a year. In addition, in some countries, phones are bought bundled with a contract for mobile phone service (from a mobile phone operator); in others, no bundling takes place. Finally, operators offer several types of contracts, meaning that there are elements of price discrimination.