Network Externalities and Critical Mass in the Mobile Telephone Network: a Panel Data Estimation

Baraldi, A. Laura
Second University of Naples

2008

Online at http://mpra.ub.uni-muenchen.de/13373/
MPRA Paper No. 13373, posted 12. February 2009 / 16:49
Network Externalities and Critical Mass  
in the Mobile Telephone Network:  
a Panel Data Estimation

A. Laura Baraldi  
Second University of Naples (Italy)  
December 2008

Preliminary version

Abstract
This paper develops a simple demand model with network externalities which allow us to identify the shape of the network externalities function in the mobile telephone market and to estimate the critical mass. If the mobile telephone network exhibits positive network externalities, we expect that the demand curve is not downward sloping everywhere but it has an increasing part, the critical mass of the installed base of subscribers. Once the critical mass is reached, the growth of the network is self-sustaining. We use a panel data of the 30 OEDC Countries from 1989 to 2006 for estimating the relationship between price of 3-minute cellular call and the installed base of subscribers; we find strong network externalities effects in mobile telephone market which drive the demand curve for this network good to be an inverted U function. Moreover, given that the concavity of the demand curve depends on the extent of network externalities, the idea is to identify some variables which could affect the intensity of network effects in the mobile telephone market, because the more concave the demand curve is, sooner the critical mass is reached for any price. This may have important implications for producers in terms of initial investment and marketing strategies which they have to do to attain the critical mass.

JEL Classification: D12, L96  
Keywords: Network Externalities, Mobile Telecommunication, Critical Mass
1. Introduction

The growth of the telecommunication network\(^1\) is one of the economic and social most relevant aspects of the recent years. Telecommunications growth means an ever larger number of users operators, connections and services. It means more and more bandwidth and longer periods of network utilisation.

Presently, the telecommunications sector is going through a revolutionary change. The rapid technological change in key inputs of telecommunications services and in complementary goods has reduced dramatically the costs of traditional services and have made many new services available at reasonable prices. Moreover, the sweeping digitization of the telecommunications and the related sectors contributed to the change. The underlying telecommunications technology has become digital and the consumer and business telecommunications interfaces have become more versatile and closer to multifunction computers than to traditional telephones.

Digitization and integration of telecommunications services with computers create lot of business opportunities and impose significant pressure on traditional pricing structures, especially in voice telephony.

The tremendous growth in demand of the mobile telecommunication network in 1994-95 was not driven as much by price reduction as it was by the “feedback” effect introduced by both past increases and anticipated future increase in the size of the installed base; in other words, “critical mass” of the installed base of consumers was reached. The concept of critical mass is linked to that of network externalities. Indeed, the telecommunication network is a typical one characterized by direct network externalities (see section 2) which play a fundamental role for the growth of this market.

Economic literature showed that consumption network externalities could have an important implication for size and structure of the telecommunication market. A major problem facing a producer interested in introducing a network good is the ability to attain the critical mass: a group of subscribers to startup. When we look at the new TLC services, it is difficult to know how to identify which of them will take off and become a new kind of network, and, indeed, what distinguishes the successful new network from those that fail to become realized.

\(^1\) A network is a market in which the benefit each consumer obtains from a good is an increasing function of the number of consumers who own the same or similar goods.
The existence of network externalities shapes the perception of good’s value during a network startup: prospective subscribers perceive more value as the subscribers base grows. Then, value increases with size, at least to the limit of the community with which any subscriber wants to communicate. This is why, in the start-up, as the quantity of subscribers rises, the price rises too, showing a positive slope of the demand for the network good.

Startup network differs in one key aspect from the mature network: the need of a critical mass. Once the critical mass is reached the system experiments its development and growth which characterizes its maturity. As a consequence, new subscribers to a mature network can join one after another rather than as a group. This is why we should not expect to find evidence of network externalities in mature network.

Rogers (2003) defines the critical mass as the minimal number of subscribers (adopters) of an interactive innovation for the further rate of adoption to be self-sustaining; that is, network effects can generate multiple stable equilibria separated by an instable one, the critical mass.

Economides & Himmelberg (1995) propose another concept of critical mass: the minimal non-zero sustainable equilibrium size (market coverage) of a network good or service (for any price); for many network goods, the critical mass is of significant size, and therefore for these goods smaller market coverage will never be observed. Common to both the definition is that consumers must be convinced that the market will be sufficiently large to justify their purchase (since utility of the network good’s consumers depends on the number of other consumers). Then customers must be convinced about the intentions of the others. Thus, common knowledge of beliefs is required to guarantee that critical mass will, indeed, be reached. And, once the critical mass is reached the network experiments an exponential growth.

During the start-up phase of new technologies, when network externalities do not play an active role in developing demand, the effort to overcome this structural inertia of the diffusion process and reach a critical mass will require supply incentives. The amount of the incentives, if these are financial incentives, is highly dependent on the critical mass level and this level is in its turn dependent on the interest an individual (or group of individuals) has in infrastructure itself. In other words, the amount of the financial incentives required to stimulate networks depends on the intensity with which network externalities play their role in the diffusion process.
From this last point come the idea which constitute the aim of the paper. It provides an empirical study on the extent of network effects in mobile telecommunications.

We developed a simple demand model with network externalities which allow us to identify the shape of the network externalities function in the mobile telephone market and to estimate the critical mass. The theoretical literature has showed that the demand curve in presence in network effects has an increasing part; the upward-sloping part of the inverted U consists of unstable equilibria and constitutes the critical mass of the good or services (for any given price). If the critical mass is exceeded, demand expands to the downward sloping part of the inverted U which consists of stable equilibria.

Existing empirical literature shows the difficulty in estimating the critical mass point; the present paper try to develop a methodology to do that and to measure the network externalities effects which could be useful for forecasting the diffusion of future telecommunication technologies, such as UMTS (3G) and 4G.

In our knowledge this is one of the very few panel empirical analysis on the network externalities effects; it uses the variables in the World Telecommunication/ICT indicators database of 30 OEDC Countries from 1989 to 2006.

Under the hypothesis that the best proxy for the future installed base of the mobile telephone network is the past installed base, we estimate a network externality function and we find strong network effects in mobile telecommunications. Given that, we estimate the demand for mobile telephone services and we show that the critical mass exists and depends on the extent of the network externalities: the more concave the inverted U is, sooner the critical mass is reached for any price. Then the idea is to identify some variables which can affect the critical mass. Reaching the critical mass point is fundamental in the start-up phase of the network good: the question is simply of getting beyond the critical mass point to have a diffusion path which is self-sustaining.

We found that variables such as the population density, the internet base of subscribers, the number of digital mainlines and the rate of schooling, affected the intensity of network externalities in mobile telephone market and then the critical mass. Knowing that, producers of network goods in telecommunication market can design their marketing strategies to reach the critical mass according to the presence of those variables.

This paper wants to enrich the poor empirical literature on network externalities effect, analyzing the market where probably these network externalities are more present.
The paper is organized as follows: in Section 2 summarize the literature on network externalities; Section 3 provide a description the TLC market; in Section 4 we make a descriptive analysis and in section 5 we discuss the empirical analysis and the results; Section 6 provides the concluding remarks and the lines for further investigations.

2. The Literature

The literature on network effects usually distinguishes among two types of network externalities: *direct network externalities* and *indirect network externalities*. Direct network externalities refer to the case where users directly benefit from the fact that there are large numbers of other users of the same network; that is, direct network externalities are generated through a direct effect of the number of agents, consuming the same good, on the utility function of agents themselves (through a creation of new goods that directly and positively affects the utility function of every participant to the network). The TLC network (fixed and mobile) is a typical one characterized by direct network externalities which, indeed, arises when the user can call a larger set of other users.

Indirect network externalities arise when the value of a good increases as the number, or variety, of complementary goods increases: the addiction of new varieties of one type of components affects positively but indirectly the utility of all participants through the reduction of prices. More generally, most markets with indirect network externalities are characterized by the presence of two distinct sides which benefit from the interaction among them. Typical examples are the PC market and the credit cards network.

Positive network externalities give rise to positive feedback; positive feedback makes the strong get stronger and the weak get weaker, leading to extreme outcomes. In a network the firm’s dominance is based on *demand-side economies of scale*. Customers value the firm’s good because it is widely used, the de facto industry standard; rival goods just don't have the critical mass to pose much of a threat. Unlike the supply-side economies of scale, demand-side economies of scale don't dissipate when the market gets large enough: if everybody else uses the firm’s good, that's even more reason for you to use it too. So marketing strategy designed to influence consumer expectations is critical in network markets.
With network externalities the fundamental relationship between price and quantity may fail\(^2\). For these goods, the willingness to pay for the last unit increases as the number expected to be sold increases. If expected sales equal actual sales, the willingness to pay for the last unit *may* increase with the number of units sold. Thus, for goods with network externalities, the (fulfilled expectations) demand-price schedule may not slope downward everywhere; in such markets, as costs decrease we may observe discontinuous expansions in sales rather than the smooth expansion along a downward sloping demand curve.

If the number of people who connect to the network is low, then the willingness to pay of the marginal individual is low, because there aren’t many other people out there that he/she can communicate with; if there are a large number of people connected, then the willingness to pay of the marginal individual is low, because everyone else who valued it more highly has already connected. This is why we can imagine a demand curve for network goods like that in figure 1.

![Figure 1](attachment:image.png)

**Figure 1**

The network starts at essentially zero, with a few small perturbations over time. As cost decreases over time (due to the technological progress), at some point it reaches a critical mass (the unstable equilibrium) that kicks us up past the low-level equilibrium and the system then zooms up to the high-level equilibrium (as shown by the arrows in figure 1). Then the two stable equilibria are zero and the highest level of the network size. The middle equilibrium is unstable because if one person decides to drops out of the network, then at least one of the remaining subscribers will find it unprofitable to belong and will leave (the value of the good is lower than the cost); but when this

\(^2\) For normal goods which do not exhibit network externalities, demand slopes downward; as price decreases, more of the good is demanded.
happens, at least another person will leave and so on until the network has no remaining members. If, on the other hand, one person decides to join, another member will find profitable to join to, and so on until the highest equilibrium level. Therefore, to get the high level equilibrium from the zero equilibrium, it would be not necessary for all consumers to agree in advance to join; all that would be needed is to achieve the critical mass, that is, the number needed to get just beyond the unstable equilibrium.

The stable equilibrium has a large number of people; here the price is small because the marginal person who purchases the good doesn’t value it very highly, even though the market is very large. As we can notice, the concept of critical mass is linked to this particular form of the demand curve.

According to this explanation, the definition of the critical mass which we agree with is the “point after which further diffusion becomes self-sustaining”\(^3\).

The concept of critical mass formalizes the "chicken and the egg" paradox that logically arises in such markets, namely: many consumers are not interested in purchasing the good because the installed base is too small, and the installed base is too small because an insufficiently small number of consumers have purchased the good.

Before analyzing the theoretical and empirical literature on network externalities, we briefly describe the sources of network externalities in telecommunication network. First, with rising number of users having subscribed to a network, it becomes more attractive for other people also to buy a mobile phone and subscribe to the same network. This is the “direct effect” as in fixed-line telecommunications: consumers value the installed base of subscribers, because they can satisfy more communication needs\(^4\). Second, network expansion drives the usage volume of people already using mobile telecommunication: we would expect the usage volume of existing subscribers increases with the total number of mobile telephone subscribers.

More recent economic literature (e.g. Granovetter and Soong, 1986; Becker, 1991; Lindbeck et al., 1999; Schoder, 2000) starts with the social interaction theory in order to show that another source of network externalities is a need of people to buy, consume, and behave like their follows; therefore we expect that consumption of mobile telephone service is influenced by such conformist behavior.

\(^3\) Rogers (2003).

\(^4\) If the installed base of fixed-line subscribers is already huge, network effects could arise in mobile telecommunications when mobile customers can call the stationary numbers. However, short message service (SMS) – available only within mobile network – might help to generate the "traditional" network effects.
On-net call⁵ discounts offer another explanation for network effects in mobile telecommunications. Blonski (2002) call this effect as “endogenous network externality”: given that it is cheaper to call a mobile number from mobile telephone in the same network than from another network, larger mobile network implies - as before - lower monthly bill, hence higher attractiveness of mobile telephone service in general.

After the seminal article of Rohlfs (1974), and the influential papers of Katz and Shapiro (1985) and Farrell and Saloner (1985), the theoretical studies on network effects became more and more rich; but, empirical works in this area are still poor. Greenstein (1993) conducts the first research in that stream. He shows that compatibility with the installed base matters in the choice of the mainframe computer system. Gandal (1994, 1995), in order to test the hypothesis that the software markets exhibit network externalities, estimates hedonic price equations for spreadsheets and data base management systems finding that the consumer’s willingness to pay for software supporting a common file compatibility standard is increasing. Similar results are in the Brynjolfsson and Kemerer (1996) paper. Additionally, they find that a product’s installed base increases the price of spreadsheets. But those authors use a specification of hedonic price model which, in our mind, should be used in market with direct network externalities and not with indirect one, as in the spreadsheets network.

In the empirical part of the Economides and Himmelberg (1995) paper they estimate the demand for facsimiles in the U.S. over 1978-1991. The assumption that facilitates the estimation is that expected network size is a linear function of the past network size. Fulfilled expectations would then lead to a constant growth rate of the U.S. fax network, which is counterfactual and breaks the consistency of that structural model.

Others structural econometric works concerning network externalities include Gandal, Kende, and Rob (2000) for the CD industry and Rysmann (2002) for the Yellow Pages market. These authors concentrate on the indirect network effect and estimate two interrelated demand equations, for software and hardware, to model the complementarities between software and hardware.

While it is widely acknowledged that network effects are a key feature of telecommunications industries, and indeed that telecommunications networks provide perhaps the leading example of network effects, relatively few studies have analyzed the empirical importance and extent of network effects in the telecommunications market.

⁵ On-net calls are calls made to the same network; off-net calls are calls made to other network.
Empirical literature on mobile telecommunications concentrates on determinants of growth and competitiveness of the industry neglecting network effects in general. The study by Bousquet and Ivaldi (1997) is probably the first one which tests empirically for existence of network effects in the fixed-line telecommunications; the concept of network externality they use relies on received calls, which benefit subscribers without paying for them, rather than on installed base of subscribers. Next, Okada and Hatta (1999) specify demand for fixed-line and mobile telephone service adopting an Almost Ideal Demand System. They show that the number of mobile subscribers, as a quality measure for telephone service, has significant positive effect on share of telecommunications’ expenditures – both mobile and fixed-line – in households’ budgets. This result is an empirical evidence of network effects in demand for telephone service. Kim and Kwon (2003) show that consumers prefer mobile service providers with larger number of subscribers because of the intra-network call discounts and quality signaling effect. Directly related to our research is the study from Doganoglu and Grzybowski (2005) on network effects in the German mobile telecommunications market. They estimate a system of demand function for mobile subscribers in Germany in the period from January 1998 to June 2003 and find that network effects played a significant role in the diffusion of mobile services in Germany. Grajek (2003) specifies a structural model of demand for mobile telephone service and estimate this model for the Polish mobile telephone industry using quarterly panel data for the period 1996-2001; he provides empirical evidence on the extent of network effects and compatibility between networks in mobile telecommunications finding strong network effects, which give rise to upward-sloping demand, and, despite full interconnection of the mobile telephone networks, low compatibility.

3. Description of the telephone market

The fixed telephone market borns as a monopolistic one. The justification was the presence of economies of scale and of density which have leaded to a natural monopoly. This natural monopoly was a public one in Europe and a regulated one in USA. In Europe the market liberalization started in 1988; in USA it started in 1984 for the long distance communications and in 1996 (with the Telecommunication Act) for the local communications.
The technology research and development contributed to reduce fixed costs and to shoot down the barriers to entry. This fixed telephone market has been always regulated, either with only one firm or with a plurality of firms, to guarantee an efficient service to everyone, independently from their revenue.

For the mobile telephone market the history has been different because there were not natural monopoly technology conditions but the number of firms in the market has been decided by the number of licenses offered by governments.

In most countries, cellular phones were first available to end consumers in the 1980s with first-generation (1G) cellular networks, based on analogue signal transmission, which offered lower service quality. Analog cellular telephone systems were experiencing rapid growth in Europe, particularly in Scandinavia and the United Kingdom, but also in France and Germany. Each country developed its own system, which was incompatible with everyone else’s in equipment and operation. This was an undesirable situation, because not only the mobile equipment was limited to operation within national boundaries, which in a unified Europe were increasingly unimportant, but there was also a very limited market for each type of equipment, so economies of scale and the subsequent savings could not be realized. The Europeans realized this early on, and in 1982 the Conference of European Posts and Telegraphs CEPT, formed a study group called the Groupe Spécial Mobile GSM in order to study and develop a pan-European public land mobile system.

Second generation (2G) network, based on digital technology, appeared in the middle of 1990s offering greater network capacity and the SMS functionality, which enabled users to send short text messages to each other.

In 1989, GSM responsibility was transferred to the European Telecommunication Standards Institute (ETSI), and phase I of the GSM specifications was published in 1990, Commercial service was started in mid-1991 and by 1993 there were 36 GSM networks in 22 countries with 25 additional countries having already selected or considering GSM. Although standardized in Europe, GSM is not only a European standard. GSM networks are operational or planned in almost 60 countries in Europe, the Middle East, the Far East, Africa, South America and Australia. In the beginning of 1994 there were 1,3 million subscribers worldwide. By the beginning of 1995 there were over 5 million subscribers. The acronym GSM now aptly stands for Global System for Mobile communications.
Once introduced, mobile telecommunication in the US and Europe always was in strong demand. In the 1990’s the rapid and sustained growth rate was accompanied by profound changes in the telecommunications markets. What once was the usual way to call someone changed from using the telephone booth or a fixed telephone line to using a personal phone kept in the pocket or in the handbag.

Using a phone increasingly meant using a mobile phone instead of a fixed, a change that started in 1993. Global mobile communication - in all EU member states - is subject to regulation by an independent national regulatory authority (NRA). For the broader market only the European Commission targets the wholesale market, hence the retail market is essentially a national market (EC, 2006).

The third generation (3G) networks allows the data transmission and is the technology in usage nowadays. For the 2G network, operators focused on capturing the mass market, that is, on reaching the critical mass of consumers. They adopted lots of strategies, as penetration pricing (taking losses for some years), or handset subsidies, giving handset away “for free” if the consumers signed up for a long-term contract. The best strategy implemented by cellular phone operators, which justified the rapid increase in the diffusion speed, was the prepaid contracts, which involved a per-minute cost instead of a monthly fee. After the explosion of the market, the number of tariffs has proliferated enormously.

In the case of cellular telephony, direct network effects may operate across multiple operators and technologies (since users of a particular network can call users from other networks and even fixed line numbers)

4. Descriptive analysis

Figure 2 shows the mobile cellular telephone subscribers and the price of 3-minute cellular call\(^6\) from 1989 to 2006 (we call, in the picture, \textit{base} the mobile cellular telephone subscribers and \textit{price} the price of 3-minute cellular call). Our analysis begins in 1989 because before 1989 the data on mobile subscribers (and price) were not available given the absence of the cellular telephone market. The exponential growth of this market started from 1993-1994, when the GSM technology replaced the TACS

\(^6\) We calculate the mean of the mobile cellular telephone subscribers and of the 3-minute cellular call across over the 30 OECD Countries (Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxemburg, Mexico, Netherland, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, UK, US) for every year.
technology and when consumers started using mobile phone instead of fixed one. As the figure shows, after 1993-1994 the demand for mobiles began to accelerate drastically. In 1997 it exploded to more than double the previous year, and in the following years it did the same; in 2004 the installed base has grown to more than 30 million subscribers. The plot for time series for price and subscribers reveals that the number of users connected to the network is initially small, and increases only gradually until the critical mass is reached, when the network growth takes off dramatically following a rapid decline in price. This picture seems to confirm the prediction of the theory.

![Figure 2](image1.png)

**Figure 2**

Interesting is Figure 3 which shows the mobile cellular subscribers over the population.

![Figure 3](image2.png)

**Figure 3**

The data speak for themselves: the fact that ‘everyone has a mobile phone’ is not very far away. Actually there are Countries that have a market penetration of more than 100
per cent because some individuals have multiple subscriptions, for example one at work and one for private use and temporary subscriptions in foreign countries. An important driver of demand is price decreases. Additional support to the strong market growth came from the Internet revolution and worldwide liberalisation, privatisation and deregulation of the telecommunication markets. The mobile phone has become a symbol of status and fashion, the use of a mobile phone is also a part of young people’s consumption style, incidentally to a large part paid by their parents (Wilska, 2003).

The number of mobile network operators increased considerably as a result of two processes. The first is the liberalisation of fixed and mobile telephony which started in the 1980s; the second process is the incorporation and partial privatisation of the former incumbent public telecom operators in the 1980s and 1990s. As a consequence of competition, prices went down while traffic volumes increased. We can say that beside the introduction of prepaid cards and new services, such as the Short Message Service (SMS) and wireless application protocol (WAP), network effects are the most important force causing such tremendous growth rate.

5. **Econometric model, analysis and results**

The explosive growth of the cellular network during the 90s was fueled by both realized and anticipated increases in the size of the installed base. This because of the network externalities.

We developed a simple demand model with network externalities; using the mobile telephone market data, we test the shape of the network externalities function and we estimate the critical mass in this network.

We define a network externalities function which captures the influence of network size expectations on the willingness to pay for the good provided through the network: $h(n^e)$, where $n^e$ is the expected size of the network (we specify and estimate the shape of this function later on). Network externalities are positive, so $h’>0$ (larger expected sizes of networks give higher individual utility) and $h’’<0$ (the marginal network externality is decreasing in network size). Consumers are uniformly distributed between 0 and 1; $G(y)$ is the density function, where $y$ is the willingness to pay for one unit of the good in a network of expected size $n^e$. We use an addictive utility specification: for a consumer

---

7 $G’(y)$ is positive.
8 All consumers receive the same benefit from the same network (Katz and Shapiro (1985), Cabral (1990), Economides (1995)).
indexed by \( y \); the utility function for consuming the good in a network of expected size \( n^e \) is
\[
u(y, n^e) = y + h(n^e).
\]

Given expectations \( n^e \) and price \( p \), every consumer with
\[
u(y, n^e) = y + h(n^e) \geq p
\]
will purchase the good; then we can impose the indifferent condition
\[
u(y, n^e) = y + h(n^e) = p
\]
to find \( y^* = p - h(n^e) \): all consumers with \( y > y^* \) will purchase the good. Given the uniform distribution of types, we find that
\[
n = 1 - y^*
\]
and then
\[
n = 1 - p + h(n^e).
\]

We can invert the (2) and find the inverse demand function for the network good
\[
p = 1 - n + h(n^e)
\]
(3)
The \( h(n^e) \) gives the shape of the above inverse demand function. The hypothesis here is that the future (unobserved) installed base is approximated by the lagged network size: consumers care about the lagged network size in their decision about joining the network\(^9\). In our mind, the two most likely specifications of the \( h(n^e) \) (given the constraints \( h' > 0 \) and \( h'' < 0 \)) are
\[
h(n^e) = k + b_1 n_{-t} + b_2 (n_{-t})^2 \quad \text{and} \quad h(n^e) = k + b_1 n_{-t} + b_2 \log(n_{-t})
\]
where \( k \) is the stand-alone value of the network good.

To choose among them, we estimate both the functions
\[
n_{it} = k + a_1 n_{i,t-1} + a_2 (n_{i,t-1})^2 + \epsilon_{i,t}
\]
(4)
and
\[
n_{it} = k + b_1 n_{i,t-1} + b_2 \log(n_{i,t-1}) + \epsilon_{i,t}
\]
(5)
where \( n_{i,t} \) is the Mobile cellular telephone subscribers (Post-paid + Pre-paid), for Country \( i \) at time \( t \); it refers to the use of portable telephones subscribing to a mobile telephone service and provides access to Public Switched Telephone Network (PSTN) using cellular technology. This can include analogue and digital cellular systems. This also includes subscribers to IMT-2000 (Third Generation, 3G). Given the compatibility of the different mobile operators in the same Country, we can consider the network size (the installed base) as composed by the sum of the number of subscribers of each network operator in every Country. These data come from the ITU database; the sample contains

\(^9\) In the estimated equation the lagged network size corresponds to the lagged dependent variable.
the 30 OECD Countries from 1989 to 2006. We start from 1989 because before this year cellular almost didn’t exist. We have chosen the OECD Countries because of the same evolution of the telephone market. In our estimation we call \( n_{ij} \) as \( base_{ij} \). The results are showed in table 1.

It is well known by now that the standard approaches to panel data analysis are inappropriate in a dynamic setting. Both fixed and random effects estimators lead to biased and inconsistent estimation results in the presence of a lagged dependent variable (Baltagi, 1995). To remove this bias, it is necessary to provide a valid set of instruments for this lagged dependent variable. Arellano & Bond (1991)\(^{10}\) offer a solution to this problem by treating the model as a system of equations (viz. one for each time period) and developing a Generalized Method of Moments estimator that exploits the moment conditions for the equations in first differences. Specifically, the estimator is based on taking first differences of the model (to remove Countries-specific effects) and then instrumenting the lagged dependent variable in first differences with suitable lags of its own levels\(^{11}\).

The coefficient showed in columns (a) and (b) of table 1 are estimated by the Arellano-Bond GMM robust estimator. However, an important obstruction to using GMM is that the lagged values of the dependent variable may be only weak instruments in the differenced regression. This could lead to severe finite-sample bias, especially when the series is very persistent (see Blundell & Bond, 1998). Given this, we employ system GMM estimation (Arellano & Bover, 1995; Blundell & Bond, 1998). This method combines the moment conditions for the equations in first differences exploited in the difference GMM estimator with additional moment conditions for the equations in levels. The introduction of these additional moments increases the efficiency of the

---

\(^{10}\) Linear dynamic panel-data models include \( p \) lags of the dependent variable as covariates and contain unobserved panel-level effects, fixed or random. By construction, the unobserved panel-level effects are correlated with the lagged dependent variables, making standard estimators inconsistent. Arellano and Bond (1991) derive a consistent generalized method of moments (GMM) estimator for the parameters of the model

\[
y_{i,t} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + \beta_1 x_{i,t} + \beta_2 w_{i,t} + \nu_i + \epsilon_{i,t}
\]

where \( \alpha_j \) are \( p \) parameters to be estimated, \( x_{i,t} \) is a vector of strictly exogenous variables, \( w_{i,t} \) is a vector of predetermined variables, \( \beta_1 \) and \( \beta_2 \) are parameters to be estimated, \( \nu_i \) are the random effects that are i.i.d. over the panel with variance \( \sigma^2 \), and \( \epsilon_{i,t} \) are i.i.d. over the whole sample with variance \( \sigma^2_{\epsilon}. \)

\(^{11}\) The estimator developed by Arellano & Bond (1991) is generally called difference GMM (or GMM-DIF). It is ideal for short time series (such as ours).
estimation. Equations (c) and (d) in table 1 uses Blundell-Bond one-step GMM robust estimator.

If the specification of the externalities function is the (4), to be \( h' > 0 \) must be \( n_{i,t-1} < a_1/2a_2 \) which is not possible for the values of the installed base in the dataset\(^{12}\). Looking at the \( h(n') \) specification in the (5) the estimated coefficients in columns (b) and (d) show that \( h' > 0 \) and \( h'' < 0 \) hold. In the estimation presented in columns (b) and (d) we control for the per capita GDP in US dollar (called \( GDP_{it} \)) and for population (called \( Pop_{i,t} \)), both for Country i at time t. The results are that the mobile telephone market exhibits strong positive network externalities: the estimated coefficients \( b_1 \) and \( b_2 \) are positive and highly significant. Moreover the marginal network externalities is decreasing: the wider the installed base is, the weaker the network externalities. The signs of the \( GDP_{it} \) is positive (as expected) and significant; the same holds for the population. The last raw of the table 1 shows the p-value of the second order autocorrelation of residuals of the Arellano-Bond test for zero autocorrelation in first-differenced errors\(^{13}\): in equation (b) at 5% there is no autocorrelation, in equation (d) there is.

Once specified the shape of the network externalities function, we can write and estimate the inverse demand function for the mobile telephone service, as derived in equation (3).

Stating from (3), we substitute
\[
h(n') = k + b_1n_{i,t} + b_2\log(n_{i,t})
\]
and write
\[
p_i = 1 - n_i + k + b_1n_{i,t} + b_2\log(n_{i,t});
\]
in equilibrium \( n_i = n_{i,t} \) and we find the willingness to pay of consumers for the mobile telephone services which we are going to estimate
\[
p_{it} = a + \beta_1 base_{it} + \beta_2 \log(base)_{it} + \beta_3 X_{it} + e_{it}
\]
(6)

To be an inverted U-shaped function, must be \( \beta_1 < 0 \) and \( \beta_2 > 0 \)\(^{14}\): when the installed base is small, the positive effect of the network size expectation on the willingness to pay of consumers is stronger than the negative effect of the network size; as soon as the

---

\(^{12}\) To prove that we use the estimated coefficients \( a_1 \) and \( a_2 \) in equations (a) and (c) of table 1.

\(^{13}\) The null hypothesis of the Arellano-Bond test for zero autocorrelation in first-differenced errors is no autocorrelation. First-order autocorrelation in the differenced residuals does not imply that the estimates are inconsistent but second-order autocorrelation would imply that the estimates are inconsistent.

\(^{14}\) \( p' = \beta_1 + \beta_2/n = 0 \rightarrow n = - \beta_2/\beta_1 \).

\( p^* = - \beta_2/n^2 < 0 \) if \( \beta_2 > 0 \); then must be \( \beta_1 < 0 \).
installed base goes beyond the critical mass, the strength of the two effects is reversed and the inverse demand function slopes downward.

The matrix $X_{ij}$ contains some control variables. The critical mass, that is the up-ward sloping of the willingness to pay of consumers, depends on the values of the parameters $\beta_1$ and $\beta_2$: the strength of network externalities; more concave is the curve, sooner the critical mass is reached for any price (see figure 4).

As said in the introduction, the idea here is to give a methodology to estimate the critical mass in a network characterized by network externalities and to find some variable which could influence the concavity of the inverse demand curve in the mobile telephone market, and then, the critical mass.

![Figure 4](image)

This could have some important implications. In presence of network externalities a prospective subscriber will actually decide to join only if some minimum number of the other prospects also decide to join. Then individuals base their decision on what they expect the others to decide. The individual decision depends on what the perception decision of the group is. During the startup of the network, firm faces losses since cost exceeds price, while in maturity, when price exceeds cost, the network provider gets higher profits. Then, the network provider, to overcome the initial inertia and reach the critical mass, will has to do huge amount of investments in terms of subsidies, to coordinate the purchasing decisions of consumers.

In this contest, subsidy needs to create a shared expectation that subscription will be larger than critical mass and, if it occurs, they then will trigger the growth remaining to reach the full network maturity. The incremental shift in expectation is realized by the creation of a new value considering as necessary what was earlier viewed as novelty.
The amount of those subsidies depends on the extent of network externalities which, as said before, influence the critical mass.

Since the paper focuses on the demand side of the market and, in particular, on the identification of the network effects and critical mass, we do not impose any structure on the supply side\textsuperscript{15}; from an econometric point of view, the endogeneity problem of price can be solved by the instrumental variable techniques.

Moreover, given that we focus on the identification of network effects, we restrict pricing behaviour of the providers by assuming that competition in the mobile telephone industry results in setting equal hedonic prices across brands over time for every Country. This assumption seems natural, as consumers’ preferences are not brand specific. As a consequence, in each instance of time consumers are indifferent toward brands.

In columns (a) and (b) of table 2 are showed the results of the estimation of equation (6). Before looking at that, let’s describe the dependent variable “price” in the estimated equation. “price” is the mobile cellular price of 3-minute local call (peak and off-peak) in US dollars, for Country \(i\) at time \(t\). The price of a 3-minute peak and off-peak rate call refers to calls from a mobile cellular telephone to a mobile cellular subscriber of the same network\textsuperscript{16}. One could think that the initial “connection charge” for a mobile network was the most appropriate price in estimating the willingness to pay of consumers for mobile telephone services; but we decided to choose the price of calls because it grasps a wider kind of network externalities which we call “externality of use”, that includes the “network externalities” to which we have referred to until now (as a benefit, for existing users, of a new user in the network).

If a mobile cellular telephone market exhibit a positive critical mass point, we expect to observe a negative sign of \(\beta_1\) and a positive sign of \(\beta_2\); these two coefficients measure the strength of network effects on the willingness to pay: the stronger the network externalities, the sooner the critical mass is reached. More in detail, the higher (the less negative) \(\beta_1\) and \(\beta_2\) are, the more concave the demand curve.

Equation (a) present a fixed effect robust estimation; we control, as before, for the per capita GDP in US dollar (\(GDP_{i,t}\)) and for population (\(Pop_{i,t}\)), both for Country \(i\) at time

\textsuperscript{15} The realistic assumption here is of an oligopolistic competition in mobile telecommunication market.

\textsuperscript{16} Given the compatibility of network in the same Country and among Countries, we can consider the price of 3-minute call from a mobile cellular telephone to a mobile cellular subscriber of the same network as a good proxy of the price of call from a mobile cellular telephone to a mobile cellular subscriber of a different network (even to a fixed telephone network).
t. We introduce the variable $\ln(\text{quantity})_{i,t}$, which is the natural logarithm of the number of mobile telephone calls. We didn’t find this data on the ITU database, but we derived it from the Revenue from mobile communication in US dollar\textsuperscript{17}; we treat this variable as an endogenous one. The interpretation of its coefficient is an absolute change in price of 3-minute calls due to a relative change in the quantity of calls; we expect a negative sign for this variable. The error term is interpreted as the mean value of the consumer’s valuations for unobserved product characteristics, such as product quality for instance. The variable $\text{base}$ has been described above.

Let’s comment the results of equation (a). The sign of the $\text{base}_{i,t}$ is negative and that of $\ln(\text{base})_{i,t}$ is positive as expected (both are highly significant): the willingness to pay of costumers for mobile calls is up-ward sloping, it reaches a maximum and then it slopes downward; then we empirically showed, for this network, that a critical mass point exists.

The coefficient of $\ln(\text{quantity})_{i,t}$\textsuperscript{18} is negative and significant meaning that a 1% increase in the quantity of calls implies a price of calls reduction of 0.53. Always as expected, the signs of GDP$_{i,t}$ and Pop$_{i,t}$ are positive and the coefficients are highly significant. The constant term (as $\text{Constant}$) is positive and significant: consumers of mobile telephone derive network benefits also from fixed line network, thus the constant term captures the utility to communicate with a fixed telephony, which we expect to be positive, such as it is.

The estimation in column (a) presents autocorrelation of residuals\textsuperscript{19}. To solve this problem we choose a dynamic specification (introducing among regressors one lag of the dependent variable $\text{Price}$) and we estimate by using the Arellano-Bond technique. The result are showed in column (b); we can notice that the results do not change; the last raw of column (b) present the p-value of the Arellano-Bond test for autocorrelation: we do not reject the null hypothesis of no second-order autocorrelation of residuals.

Using the estimated coefficient in column (b) we draw the price of the 3 minute call only taking account of the network externalities effect on the willingness to pay of consumers. In order to do that we calculate the price of call in this way

\textsuperscript{17} Revenue from mobile communication is the revenues from the provision of all types of mobile communications services such as mobile cellular, private trunked radio and radio paging. We derive the number of telephone call by dividing this revenue for prices of 3-minute call. Then this is just a proxy of the quantity of calls.

\textsuperscript{18} In equation (a) the variable $\ln(\text{quantity})$ is taken at time t-1 for endogeneity problems.

\textsuperscript{19} We didn’t show the test.
\[ \text{Price}_t = -(8.14e-09) \text{base}_t + 0.33 \ln(\text{base})_t \]

where \( \text{base}_t \) is the mean of the mobile installed base of subscribers across over the Countries for every year (from 1989 to 2006). Figure 5 presents the simulation.

![Price=8.14e-09base+0.33ln(base)](image)

**Figure 5**

It is evident the up-ward sloping part of the demand curve for mobile services and the existence of the critical mass.

As said above, the idea is to find a set of variables which could affect the extent of network externalities and, then, the critical mass, through the concavity of the demand curve. In empirical terms, we introduce in the equation (6) some variables in this way

\[
\text{price}_{it} = a + \beta_1 \text{base}_{it} + \beta_2 \log(\text{base})_{it} + \beta_3 \text{Variable}_{it} + \beta_4 \text{Variable}_{it} \times \log(\text{base})_{it} + \text{Variable}_{it} + e_{it} 
\]

(7)

The interaction terms (such as \( \text{Variable}_{i,t} \times \text{base}_{i,t} \) and \( \text{Variable}_{i,t} \times \ln(\text{base}_{i,t}) \)) affect the concavity of the demand curve; the term \( \text{Variable}_{i,t} \) shifts the curve up and down.

The first variable is the *population density* (called Density in table 2): a higher population density means a higher level of human interaction and an easier use of telecommunication service. Then, we expect that in Countries with higher population density the inverse demand curve for mobile telephone will be more concave. The result of the estimation is showed in column (c) of table 2.

The significance of the estimated coefficient of the interaction variable \( \text{Density}_{i,t} \times \ln(\text{base})_{i,t} \) confirms that the strength of network externalities positively
depends on the population density: the higher the population density, the greater the concavity of the inverse U.

The second variable is the internet base of subscribers (called \( Internet_{i,t} \))\(^{20} \). The sign here is not predictable. Internet may be an important communication rival to cellular phones and consumers can substitute internet access for mobile subscription. But, on the other hands, a higher level of internet penetration may proxy for a telecommunication policy environment which generally encourages the adoption of new technologies, including mobile phones.

The result of the estimation are showed in column (d) of table 2. The interpretation depends on the extent of the internet subscribers. Indeed, as long as the coefficient of \( base \) is greater (in absolute value) of the coefficient of \( Internet*base \) and the coefficient of \( ln(base) \) is greater than the coefficient of \( Internet*ln(base) \), an increase of the internet subscriber base leads to a greater network externalities effect on the willingness to pay for cellular phone. The sign of the \( Internet \) is positive and significant, meaning that the demand curve shifts up.

Another variable which we think that could affect the extent of network externalities on the willingness to pay of costumers for mobile services is the number of digital mainlines (called \( Digital_{i,t} \)). This variable refers to the per cent of main lines connected to digital exchanges. This percentage is obtained by dividing the number of main (fixed) lines connected to digital telephone exchanges by the total number of main lines. This indicator does not measure the percentage of exchanges which are digital, but the percentage of inter-exchange lines which are digital or the percentage of digital network termination points\(^{21} \). The result of the estimation is in column (e) of table 2. The sign of the interaction variable \( Digital_{i,t}*ln(base)_{i,t} \) is negative and significant, meaning a lower concavity of the willingness to pay function. The reasons could be: the more the digital mainlines the more each \( n^{th} \) consumer is unable to be in contact with the other (\( n-1 \)) existing users; moreover, the ability of users to have audio/chat services, completely replaces the capability of mobile services.

At the end, we thought of the rate of schooling (called \( School_{i,t} \)). We use the ratio of total enrollment, regardless of age, to the population of the age group that officially

\(^{20} \) This variable comes from the ITU database; it is defined as the number of total Internet subscribers with fixed access, which includes dial-up, total fixed broadband subscribers, cable modem, DSL Internet subscribers, other broadband and leased line Internet subscribers. Only active subscribers that have used the system within a reasonable period of time is included.

\(^{21} \) ITU database.
corresponds to the secondary level of education. More educated consumers tend to use
the mobile phone for business purposes more than less educated one.

The result is presented in column (f) of table 2. The coefficients of the interaction
variables $School_{i} \times base_{i},$ and $School_{i} \times \ln(base)_{i,1}$ are both positive meaning that the
greater the rate of schooling in a Country, the greater the network externalities effect on
the price of calls, but these coefficient are not significant, probably because of lots of
missing values in the series.

6. Concluding remarks

In this paper we construct a demand model to estimate the network externality effect on
the mobile telephone network and to check the existence of the critical mass point. Once
verified that mobile telephone network exhibits strong positive network externalities,
we expect the willingness to pay for that good be an inverted U-shaped function of the
installed base of subscribers; this allow us to identify the critical mass point after which
the network growth becomes explosive. We used a Arellano-Bond and Blundell-Bond
dynamic panel data estimators (for the 30 OEDC Countries from 1989 to 2006) to prove
that the network effects played a significant role in the growth of the mobile
telecommunication market and that we can think of a set of variables, such as the
population density, the internet base of subscribers, the number of digital mainlines and
the rate of schooling, which affect the strength of network externalities and then, the
level of the critical mass. This intuition has some important implications for mobile
services producers in terms of initial (start-up) investments and marketing strategies to
reach the critical mass, after that the growth becomes self-sustaining. This paper wants
to give a methodology to estimate the critical mass that can be applied to every network
in which network externalities play a significant role. Moreover, it wants to give the
intuition that there could be some other variables which may affect the intensity of
network effects in every network and then, its critical mass; this last point is
fundamental from the point of view of the network good providers.

One could ask if, for example, 3G and 4G technologies have reached their critical mass,
and could apply the presented methodology to check the importance of other variables
in affecting the critical mass for those technologies.
<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(e)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>base</strong></td>
<td>0.98</td>
<td>0.92</td>
<td>1.16</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>(25.01)</td>
<td>(33.4)</td>
<td>(31.1)</td>
<td>(26.8)</td>
</tr>
<tr>
<td><strong>(base_{i,t-1})^2</strong></td>
<td>-3.61e-10</td>
<td>-7.64e-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.61)</td>
<td>(-3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Log(base_{i,t-1})</strong></td>
<td>309870.2</td>
<td>1016051</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.14)</td>
<td>(3.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GDP_{i,t}</strong></td>
<td>6.48</td>
<td>7.69</td>
<td>4.05</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>(2.8)</td>
<td>(3.5)</td>
<td>(1.00)</td>
<td>(0.4)</td>
</tr>
<tr>
<td><strong>Pop_{i,t}</strong></td>
<td>0.71</td>
<td>0.73</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(3.9)</td>
<td>(5.12)</td>
<td>(2.10)</td>
<td>(1.8)</td>
</tr>
<tr>
<td><strong>N.obs</strong></td>
<td>480</td>
<td>480</td>
<td>510</td>
<td>510</td>
</tr>
<tr>
<td><strong>Prob &gt; z (2 order)</strong></td>
<td>0.0574</td>
<td>0.0549</td>
<td>0.0422</td>
<td>0.0471</td>
</tr>
</tbody>
</table>

In parentheses are standardized normal z-test values. * significant at 1% level; ** significant at 5% level; *** significant at 10% level. Equation (a) and (b): Arellano-Bond one-step robust estimation; equations (c) and (d): Blundell-Bond one-step robust estimation. We include two lags of the dependent variable as instruments. Prob > z (2 order) is the p-value of the Arellano-Bond test for zero autocorrelation in first-differenced errors.
### Table 2

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
<th>(d)</th>
<th>(e)</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price</strong></td>
<td>7.9' (5.9)</td>
<td>8.52 (3.6)</td>
<td>4.48' (2.43)</td>
<td>9.4 (2.43)</td>
<td>6.62 (2.8)</td>
<td>7.77'' (2.2)</td>
</tr>
<tr>
<td><strong>Price_{t-1}</strong></td>
<td>0.148 (2.78)</td>
<td>0.14 (2.5)</td>
<td>0.14 (2.8)</td>
<td>0.15 (3.09)</td>
<td>0.08 (1.22)</td>
<td></td>
</tr>
<tr>
<td><strong>base_{t}</strong></td>
<td>-9.07e-09** (-2.22)</td>
<td>-8.14e-09** (-1.9)</td>
<td>-1.39e-08** (-2.9)</td>
<td>-1.04e-08*** (-1.8)</td>
<td>-1.04e-08*** (-2.5)</td>
<td>-6.31e-08 (0.8)</td>
</tr>
<tr>
<td><strong>In(base)_{t}</strong></td>
<td>0.155 (3.22)</td>
<td>0.33 (2.3)</td>
<td>0.30 (3.5)</td>
<td>0.65 (4.36)</td>
<td>0.57 (3.7)</td>
<td></td>
</tr>
<tr>
<td><strong>In(quantity)_{t}</strong></td>
<td>-0.533 (-7.4)</td>
<td>-0.72 (-5.7)</td>
<td>-0.398 (-5.4)</td>
<td>-0.72 (-5.4)</td>
<td>-0.74 (-5.6)</td>
<td>-0.92 (-6.5)</td>
</tr>
<tr>
<td><strong>GDP_{t}</strong></td>
<td>0.0000307' (23.91)</td>
<td>0.0000259' (7.72)</td>
<td>0.0000258' (7.1)</td>
<td>0.0000253' (8.4)</td>
<td>0.0000268' (8.6)</td>
<td>0.000031' (2.8)</td>
</tr>
<tr>
<td><strong>Pop_{t}</strong></td>
<td>4.33e-08** (2.16)</td>
<td>6.02e-08 (1.6)</td>
<td>9.59e-08 (2.4)</td>
<td>4.43e-08 (0.9)</td>
<td>6.86e-08 (1.8)</td>
<td>9.29e-08 (1.35)</td>
</tr>
<tr>
<td><strong>Density_{t},In(base)_{t}</strong></td>
<td>0.000517 (2.46)</td>
<td>3.85e-07** (1.78)</td>
<td>1.67e-16*** (1.8)</td>
<td>-2.15e-08** (-1.78)</td>
<td>-0.0009222 (2.3)</td>
<td></td>
</tr>
<tr>
<td><strong>Internet_{t}</strong></td>
<td></td>
<td>5.56e-10 (0.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Internet_{t},base_{t}</strong></td>
<td></td>
<td>5.0000016 (0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>N.obs</strong></td>
<td>510</td>
<td>480</td>
<td>480</td>
<td>479</td>
<td>438</td>
<td>157</td>
</tr>
</tbody>
</table>

In parentheses are standardized normal z-test values. * significant at 1% level; ** significant at 5% level; *** significant at 10% level. Equations (a): fixed effects robust estimation, the variable ln(quantity) is taken at time t-1 for endogeneity problems; R²=0.7. Equations (b), (c), (d), (e), (f) and (g): Arellano-Bond one-step robust estimation. ln(quantity)$_{t-1}$ is the endogenous variable in every equation. Prob > z (2 order) is the p-value of the Arellano-Bond test for zero autocorrelation in first-differenced errors.
Bibliography


