Technological Forecasting & Social Change 77 (2010) 541–557

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Innovation diffusion with generation substitution effects

Christos Michalakelis *, Dimitris Varoutas, Thomas Sphicopoulos

University of Athens, Department of Informatics and Telecommunications, Panepistimiopolis, Ilisia, Athens, 157 84, Greece

ARTICLE INFO

Article history:
Received 20 May 2009
Received in revised form 16 July 2009
Accepted 2 November 2009

Keywords:
Generation substitution
Innovation diffusion
Market potential
Intergeneration time
High technology market
Telecommunications
Dynamic diffusion model

ABSTRACT

Markets of high technology products and services, such as telecommunications, are described by fast technological changes and rapid generational substitutions. Since the conventional modeling approaches that are based on diffusion models do not usually incorporate this important aspect into their formulations, the accuracy of the provided forecasts is consequently affected. The work presented in this paper is concerned with the development of a methodology for describing innovation diffusion, in the context of generation substitution. For this purpose, a dynamic diffusion model is developed and evaluated, based on the assumption that the saturation level of the market does not remain constant throughout the diffusion process but is affected by the diffusion of its descendant generation, as soon as the latter is introduced into the market. In contradiction to the conventional diffusion models, which assume static saturation levels, the proposed approach incorporates the effects of generation substitution and develops a diffusion model with a dynamic ceiling. The importance of such an approach is especially significant for markets characterized by rapid technological and generational changes. Evaluation of the proposed methodology was performed over 2G and 3G historical data and for a number of European countries, providing quite accurate estimation and forecasting results, along with important information regarding the rate of generation substitution.

1. Introduction

Among the major characteristics of market evolution is that the corresponding products and services are differentiated in a number of successive generations, with each preceding generation affecting at a certain, positive or negative, level the diffusion characteristics of its ancestors. In today's competitive global environment, many organizations are facing tremendous pressures and others are experiencing exponential growth, mainly due to the frequent introduction of new technologies. As a result, diffusion of innovative products remains a research field facing a continuously high level of interest, since it is usually synonymous to heavy investments and critical business plans, targeting to meet market demand and competition. Industrial plans, rolled out in an attempt to attract new customers and retain the existing ones, must be accompanied by accurate demand forecasts, in order to avoid dramatic consequences in terms of supply, like oversupply and unneeded over-investments, or under-utilization of a firm's capacities.

The importance of the above considerations becomes even higher if the sector of high technology products is considered, with telecommunications being probably the most characteristic example, as they correspond to one of the most significant contemporary investments. Privatization and deregulation of the telecom market, together with the effects of increasing competition and the introduction of new services, resulted in the emergence of new problems regarding technology diffusion forecasting, under a high level of uncertainty and a need for risk management, mainly due to competition.

* Corresponding author. Tel.: +30 2107275318; fax: +30 2107275601.
E-mail addresses: michalak@di.uoa.gr (C. Michalakelis), arkas@di.uoa.gr (D. Varoutas), thomas@di.uoa.gr (T. Sphicopoulos).

0040-1625/$ – see front matter
doi:10.1016/j.techfore.2009.11.001

Modeling and forecasting the diffusion of innovations has attracted the interest of practical and academic studies since some decades ago, initiated by the pioneering works of Bass [1], Mansfield [2], Rogers [3] and others. The importance of estimating and forecasting innovations diffusion patterns is now extended to all kinds of high technology markets. The corresponding literature includes some important contributions, such as [4] where general diffusion patterns for information technology innovations are suggested, [5] where a thorough overview, together with a method for planning and implementing forecasts for disruptive and discontinuous innovations are presented and [6] containing an informative review regarding the research conducted towards diffusion modeling.

Despite the fact that the theory for diffusion estimation and forecasting is mature enough, new opportunities have emerged, mainly due to the nature of the high technology market, retaining the need for simple and accurate forecasting methods. One of the probably most important factors, very common to high technology, is the frequent substitution of technological generations. Generation substitution can be defined as a process by which an innovation is partially or completely replaced by another, in terms of its market share and this replacement can be either instantaneous or it may take some considerable time. As stated in [7], *"Each subsequent technology, if successful, tends to follow an S-shaped curve: it starts slowly due to initial resistance; then proceeds more rapidly as the competition between the new and the old technology grows keener and the new technology gains an upper hand over the old one; and finally, as the new technology becomes widespread, the pace of growth slows down."*

In general, a new generation usually appears before its predecessor has been fully diffused among its potential adopters. In addition and due to the competitive environment, the intergeneration time, the time interval that elapses between the introductions of two successive generations of a product, is becoming even shorter making the phenomenon of generation substitution more intensive. Since the process of introducing a new generation of a high technology product is usually translated into corresponding high levels of investments, not only in terms of research and development but also in terms of marketing and advertising efforts, the need for accurate predictions of the diffusion processes is a crucial factor for a firm's welfare. In addition to the process of generation substitution, the intergeneration time is by itself of paramount importance, since it seeks answers to the question of the best timing for introducing the next generation into the market. Conventional high technology products are usually substituted by their descendant generation usually after allowing a considerable time to elapse. However, as stated above, this is not the case for the high technology and especially the telecommunications market, where the corresponding time lag is usually quite shorter. Therefore, there is a need to focus on such kind of markets by studying the effects of rapid generation substitution and provide appropriate modeling approaches.

The latter consideration constitutes the main target of this work, which is the proposal of a methodology and a corresponding dynamic diffusion model, capable of capturing the effects of generation substitution into a single dynamic diffusion model. The main assumption is that both generations are considered as parts of the same process, describing the diffusion of the considered technology, which in turn consists of two stages, corresponding to the two generations of the product. In this way, the market dynamics of the preceding generation is reflected into the diffusion process of it descendant. Based on diffusion models, one could attempt to consider the diffusion estimation separately for each generation and consequently aggregate the results. However, the drawback of this approach is that the diffusion of the next generation would be in this way considered as a new separate process, or as a disruptive product, without taking into account the influence of its ancestor generation and the maturity of the market.

The rest of the paper is structured as follows: In Section 2 a conceptual overview of innovation diffusion and generation substitution literature together with the contribution of the developed methodology is presented followed by Section 3, where the proposed methodology and the development of the dynamic model are described. Evaluation of methodology in terms of estimation and forecasting is performed in Sections 4 and 5 respectively and, finally, Section 6 concludes, indicating directions for future research.

2. Conceptual overview

2.1. Diffusion theory

One of the main research interests of innovation diffusion is the mathematical modeling of different types of innovations and under different assumptions. The main findings can be summarized to the construction of bell-shaped curves, depicting the frequency of adoption against time and S-shaped curves, when the cumulative numbers of adopters are considered. During the first stages of the innovation life cycle (the introduction stage), the adoption rate is relatively low, followed by the next stage (the take off), described by a high rate of adoption, until the peak of the bell curve is reached, which corresponds to the inflection point of the cumulative adoption. After that time adoption rate decreases again, until the market saturation level is asymptotically met and the maximum number of adopters is reached. This corresponds to the end of the life cycle of the innovation, which in the cases of high technology products, is usually replaced by its descendant generation.

Apart from the early work of Gompertz [8,9], the work of Bass [1] represents the early contemporary efforts to capture the diffusion dynamics. These, together with the logistic family models [10], such as the linear logistic or Fisher–Pry [11], are the most widely used diffusion models for demand estimation and forecasting. The above models constitute the family of the aggregate diffusion models, which are mathematical formulations used to describe a diffusion process as a function of time. They are generally quite capable of providing reliable estimations regarding the adoption of innovations into a market of reference. One of their main contemporary fields of application is the sector of high technology and especially telecommunications. An informative review regarding the forecasting of demand in telecommunications can be found in [12]. Moreover, important research results regarding the development and evaluation of diffusion models are presented in [10,13–17].

Extensions of the simple diffusion models include diffusion models for capturing cross-national influence [18–20] and marketing variables [21,22], in an attempt to describe the process of adoption in more detail and provide literature with accurate modeling approaches. Most of the resulting models are derived by incorporating functional adjustments into the original formulation of a diffusion model's equation.

2.2. Multigenerational diffusion models

The work included into the papers presented above refers to modeling approaches mainly for single generation products, without taking into explicit account the effect of generation substitution. Although relevant literature in not too broad, however there are some important contributions towards the direction of modeling multigenerational high technology products, at the aggregate level.

An important approach, which is often used to model the generation substitution effect, is the work of Fisher and Pry [11] where, under the assumption that only two competing generations exist, the substitution effect is captured by the means of a logistic function formulation. The development of the corresponding model was based on Pearle's Law which states that: *"The fractional rate of fractional substitution of new for old is proportional to the remaining amount of the old, left to be substituted."* Following this, the main assumption made to built the model was that the log of the ratio of the market share of the succeeding technology to that of the first is a linear function of time or, in mathematical notation, $\ln[s/(1-s)] = kt$, where s is the fractional market share of the new generation at time t and k is a constant of proportionality. The work of Fisher and Pry was extended by Peterka [23] and Sharif and Kabir [24] to consider multilevel generational substitution. Additional to the above, is the model proposed by Mansfield [2] which was modified by Blackman [25] who proposed that $\ln[s/(S-s)] = a + bt$, where S is the upper limit on market share obtainable by the new technology, s the market share of the new generation at time t, a is a constant, and b is a linear function of investment and profitability. The main limitation of all these models is that the explicit volume of sales for each generation is not incorporated into the corresponding formulations but the substitution ratio is only considered, by expressing the share of sales for each of the generations.

Apart from the above, the work of Norton and Bass [26] constitutes another important methodological approach, the first to incorporate the actual sales of each generation. The proposed model was based on the original Bass model [1] of the diffusion of a single generation product, which was suitably extended in order to capture the effect of generation substitution. The model Norton and Bass proposed is described by the following set of equations, for the case of two competing generations: $N_1(t) = m_1F_1(t)(1-F_2(t-\tau))$ for the first and: $N_2(t) = (m_2 + m_1F_1(t))F_2(t-\tau)$for the second. $N_i(t)$ represents the adoptions in the i-th generation and τ the time of launch of the second generation. Diffusion and substitution effects are explicitly incorporated into the model, which is in addition able to simultaneously consider multiple technological generations. An extension of this model was proposed by Mahajan and Muller [27] by incorporating into a single methodology the processes of diffusion and generation substitution, aiming to predict sales and determine the optimal timing of new generations of a durable technological innovation. Evaluation of the model was performed over four successive generations of the IBM mainframe computers, providing quite accurate results. Following the latter approaches, Speece and MacLachlan [28] developed the model in a different way by incorporating price as an explanatory variable when modeling the diffusion of different technologies for milk containers. However, the main limitation of the Norton–Bass model, as well as most of the alternatives based on this approach, is that the coefficients of innovation and imitation are assumed to retain the same constant values across all the generations. This assumption was studied by Islam and Meade [29] by developing a modeling framework for simultaneous estimation of successive generations using a full information maximum likelihood procedure, which showed that, in most cases, the hypothesis of constant coefficients can be rejected. The constant coefficients hypothesis was evaluated over data describing diffusion of cellular phones in a number of European countries and the IBM mainframe computer data. According to the results derived by this study, use of a model with changing coefficients would considerably improve forecasting performance.

In addition to these approaches, there are also some more, worth mentioning, contributions. Among them, is the work of Rai [8], where a class of new substitution models is proposed. These models are non-linear in nature and are based on the arguments that the influential factor is a function of both market share and time. They have been applied to describe technological substitutions and the results obtained are compared with those of the Fisher–Pry model, evaluated over cases including television sets, domestic petroleum production, railway locomotives and the substitution of energy.

Moreover and as far as the intergeneration time is concerned, Pae and Lehmann [7] proposed that the time between the launch of two adjacent generations in the same broad product category is associated with the speed of diffusion of the later generation, concluding that the longer the intergeneration time, the lower the later generation's initial rate of adoption but the higher its subsequent rate of growth. An application of such models over the telecommunications market can be found in [30].

2.3. Research objectives and contribution

The proposed methodology is built upon the hypothesis that adopters of each successive generation consist of three categories of customers: Those who are switching from the earlier generation, those who adopt the new technology instead of the earlier one and potential customers who would only adopt the new technology. Moreover, the proposed methodology can accommodate repeat adopters, either those who previously purchased earlier generation products and proceed to the adoption of the new generation without discarding the earlier ones, or those who discard the earlier generation product. All these adopters' preferences and choices are depicted on the diffusion process of the product and consequently affect all of its successive generations.

Conventional high technology products are usually substituted by their descendant generation, usually after allowing a considerable time to elapse. However, in some cases the intergeneration time lag is rather short, which in turn leads to the necessity of focusing on the corresponding markets and study the effects of rapid generation substitution. In addition to this, in most of the diffusion modeling approaches the corresponding parameters are considered as constants throughout the process, although they are very probable to change during the diffusion life cycle. In the telecommunications sector for example, the saturation of a generation is not expected to reach the initially estimated level because, as a result of competition, its descendant generation will be rolled out into the market quite early in time. Therefore, there is a need to identify the mechanism by which the diffusion parameters of a product's generation are influenced. Important contributions in developing suitable frameworks for capturing the dynamic nature of the market potential, as this is affected by the growth of the population, constitute the works of [31,32]. The methodologies proposed in these works attempt to overcome the limitation of a fixed market potential throughout the diffusion process, by proposing a dynamic model where the saturation level is expressed as a function of the population size, at each point of time.

The major assumption made in the context of the present work is to consider the diffusion of a product as a single process, consisting of different stages, each one corresponding to a successive generation. Consequently, by the time the new generation appears the whole process is affected by the means of the model's parameters. Thus, instead of considering the diffusion process to have a fixed ceiling, the latter is represented as a function of the new generation's diffusion. In this way, the dynamic model derived by this assumption is capable of estimating and forecasting the whole process, incorporating on the same time the effects of the new generation over the product's life cycle. An alternative diffusion estimation approach, based on diffusion models as well, would be to consider the diffusion process of each one of the successive generations as standalone, apply the estimation procedure over each generation and aggregate the results. However, the main drawback of this direction is that the new generation is considered as a totally new process overlooking the so far dynamic of the market, due to the previous generation. On the contrary, the proposed methodology is based on the existing market dynamics, which is extended in order to accommodate the new generation. The estimation accuracy of this approach, as presented into the evaluation section, is much higher than of any other alternatives. Evaluation of the proposed model is performed over mobile telephony data, describing 2G and 3G diffusion across a number of European countries.

The contribution and the extension of the proposed methodology to the existing literature can be summarized into the following concepts:

Some of the existing approaches, especially the ones based on the Norton–Bass model [26], consider constant diffusion parameters across all generations of the innovation. This, in turn can be translated into the assumption that every new generation is expected to follow the same diffusion path. However, this assumption is not always the case as shown by Islam and Meade [29]. The model developed into the context of the present work does not consider a static diffusion process, in terms of the diffusion parameters but incorporates two, or more, generations into a single formulation, each one retaining its own characteristics, regarding the diffusion speed and the saturation level.

In addition, the modeling approaches based on the Fisher–Pry model [11] impose the limitation of not explicitly considering the level of adoption but they only model the rate of substitution. The proposed dynamic model is capable of explicitly estimating the substitution rate, via a single parameter, as well as the adoption level and market saturation for all generations considered and at each point of time.

Moreover, the model is expressed by a, relatively simple, formulation, without needing to resort to more complex ones. Estimation of the containing parameters is a relatively easy procedure, since it is based on conventional methods like NLS (Nonlinear Least Squares) or regression analogues.

3. Methodology development

A typical generation substitution process for a two generations product, the preceding and the successive (mentioned in the text as “first” and “second”, for simplicity reasons), is usually described by the following discrete steps:

1. The first generation is introduced into the market at some point of time.
2. The diffusion process is initiated and forecasts regarding the diffusion parameters are provided, in terms of diffusion rate and saturation level.
3. At a certain point of future time, t_o, the new generation of the product or service is rolled out, initiating its own diffusion process.

The diffusion process of the corresponding product is initiated again, as a result of the introduction of the new generation. The initially estimated saturation level is consequently expected to reach a higher level.

Based on the above assumptions, the main steps of the proposed methodology are described by the following algorithm:

1) Initial estimation of first generation's diffusion parameters (saturation level and diffusion rate), using a conventional diffusion model. This model can be used to describe the diffusion process until the new generation will be introduced into the market.
2) Estimation of the same parameters for the successive generation, as soon as it is introduced into the market.
3) Development of the mathematical relationship between the first and the second generation diffusion parameters. This can be achieved as a result of rational assumptions, regarding the relationship between the two generations' diffusion processes.

4) Derivation of the final model formulation, based on the assumptions of step 3, that describes the total diffusion process of the product, incorporating the generation substitution effect.

Steps 1) and 2) are performed by the means of a conventional diffusion model, such as the logistic and parameter estimates can be derived by applying nonlinear regression (NLS).

According to the concepts of diffusion theory, the diffusion of a product among a social system is proportional to the number of the existing adopters and to the number of the remaining market potential and is described by the following differential equation:

$$\frac{dN(t)}{dt} = rN(t)[\overline{N}-N(t)] \tag{1}$$

In Eq. (1), $N(t)$ refers to the cumulative number of adopters at time t, r is the rate of diffusion and $\overline{\mathrm{N}}$ is the saturation level, which corresponds to the maximum cumulative number of adopters that diffusion is expected to reach. The initial number of adopters, at the beginning of the process is assumed to be $N(t=0)=N_0$ and the quantity $(\overline{\mathrm{N}}-N(t))$ corresponds to the remaining market potential, while $\overline{\mathrm{N}}$ is initially assumed constant throughout the product's life cycle.

Following the assumptions regarding the dynamic nature of first generation's saturation level, $\overline{\mathrm{N}}$, instead of being considered as constant it is expressed as a function of the endogenous and exogenous parameters that affect it. Thus, in the general case, it can be expressed as:

$$\overline{N}(t) = f(U(t)) \tag{2}$$

where $U(t)$ represents the vector of the parameters affecting the saturation level.

In the context of the developed methodology, $\overline{\mathrm{N}}(t)$ is assumed to be affected only by the penetration level of the descendant generation, considering the rest parameters as constants, therefore it can be expressed as:

$$\overline{N}(t) = f(M(t)) \tag{3}$$

By incorporating Eq. (3) into Eq. (1), the latter can be expressed as:

$$\frac{dN(t)}{dt} = rN(t)[f(M(t))-N(t)] \tag{4}$$

In both the above equations, $M(t)$ refers to the penetration of the descendant generation.

Solution of the differential Eq. (4) gives the number of adopters of the product, at each point of time, t, as expressed by the following equation:

$$N(t) = \frac{\overline{N}(t)}{1+\left(\frac{\overline{N}_0}{N_0}-1\right)exp[-rP(t)]} \tag{5}$$

In the above equation, whose detailed derivation is included in Appendix A, $\overline{\mathrm{N}}_0$ is the initially estimated saturation level and in the absence of the next generation, N_0 represents the number of adopters at time t_o and $P(t)$ is given by: $P(t) = \int_{t_0}^{t}\overline{N}(t)dt$. In correspondence with the assumption of Eq. (1) and if the time of the new generation's introduction is considered to be the initial time, then N_0 corresponds to the initial number of adopters at this time, thus $N_0=N(t=t_0)$. Obviously, the formulation of Eq. (5) reduces to that of Eq. (1) if $\overline{\mathrm{N}}(t)=\overline{\mathrm{N}}$.

Diffusion of the next generation, introduced into the market at time t_o, can be described by the logistic model as:

$$\frac{dM(t)}{dt} = sM(t)[\ M-M(t)] \tag{6}$$

In Eq. (6), $M(t)$ reflects penetration of the new generation, at time t, $\overline{\mathrm{M}}$ is the corresponding estimated saturation level and s is the diffusion rate.

The analytical solution of the above differential equation, which is presented in Appendix A, results to:

$$M(t) = \frac{\overline{M}}{1+\left(\frac{\overline{M}}{Mo}-1\right)exp[-s\overline{M}(t-to)]} \tag{7}$$

In the above Eq. (7), Mo is the initial penetration of the new generation, at the time of introduction, t_o.

In the simple case, it can be assumed that $\overline{N}(t)$ is linearly dependent on the diffusion of the successive generation, $M(t)$, and this assumption is described by the following differential equation:

$$\frac{d\overline{N}(t)}{dM(t)} = a \tag{8}$$

Integration of both parts of Eq. (8) yields:

$$\int_{t_0}^{t} d\overline{N}(t) = a\int_{t_0}^{t} dM(t) \Rightarrow \overline{N}(t) - \overline{N}_0 = a[M(t) - M_0] \tag{9}$$

and finally:

$$\overline{N}(t) = \overline{N}_0 + aM(t) - aM_0 \tag{10}$$

Substituting Eq. (10) into Eq. (4) leads to the following formulation for first generation diffusion:

$$\frac{dN(t)}{dt} = rN(t)[\overline{N}_0 + aM(t) - aM_0 - N(t)] \tag{11}$$

Thus, based on the assumption of Eq. (8) that the saturation level of the preceding generation is linearly dependent on the diffusion rate of the successive one, the derived Eq. (10) provides the relation between them. In Eq. (9), the integrals are considered after the point of time t_0, when the new generation is rolled out into the market, after which the new generation is expected to affect its predecessor's so far estimated saturation level.

Substitution of Eq. (7) into Eq. (10) gives:

$$\overline{N}(t) = \overline{N_0} - aM_0 + a\left(\frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right)exp[-s\overline{M}(t - t_\partial)]}\right) \tag{12}$$

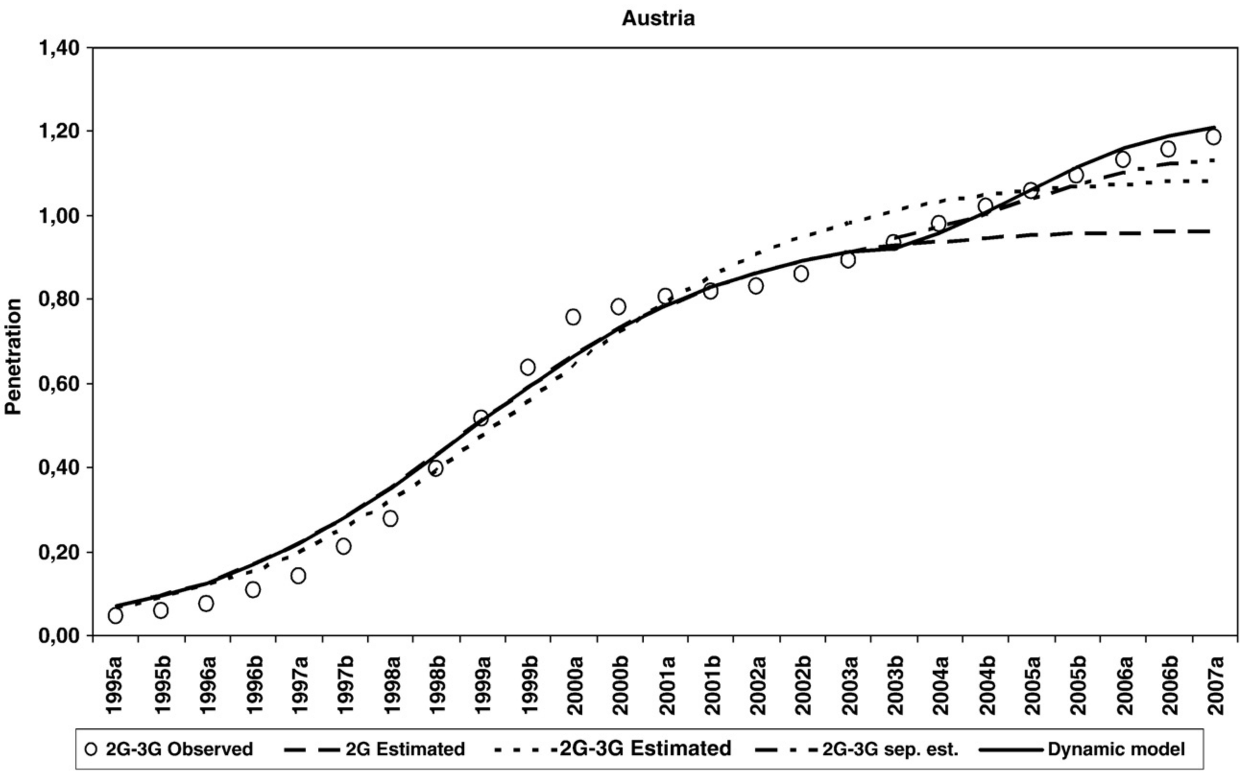

Fig. 1. Estimation of 2G and 3G diffusion in Austria, vs. observed values.

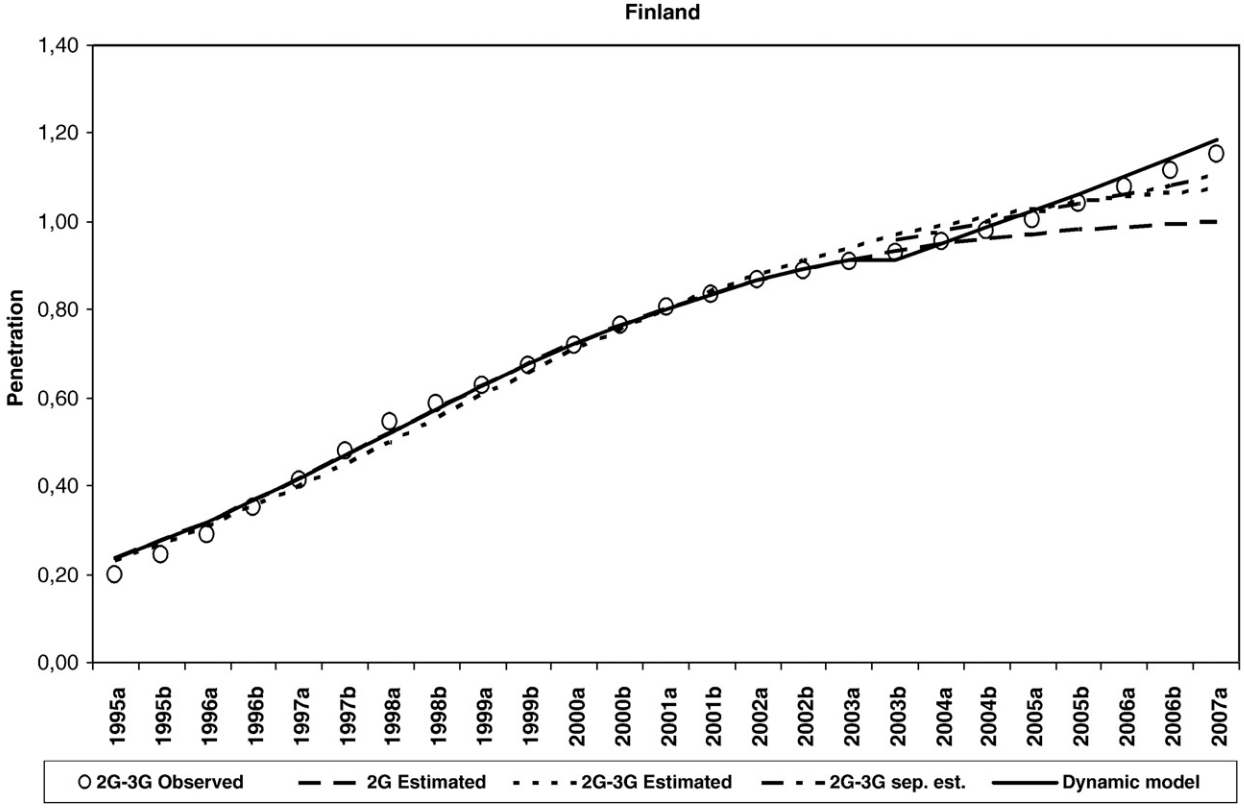

Fig. 2. Estimation of 2G and 3G diffusion in Finland, vs. observed values.

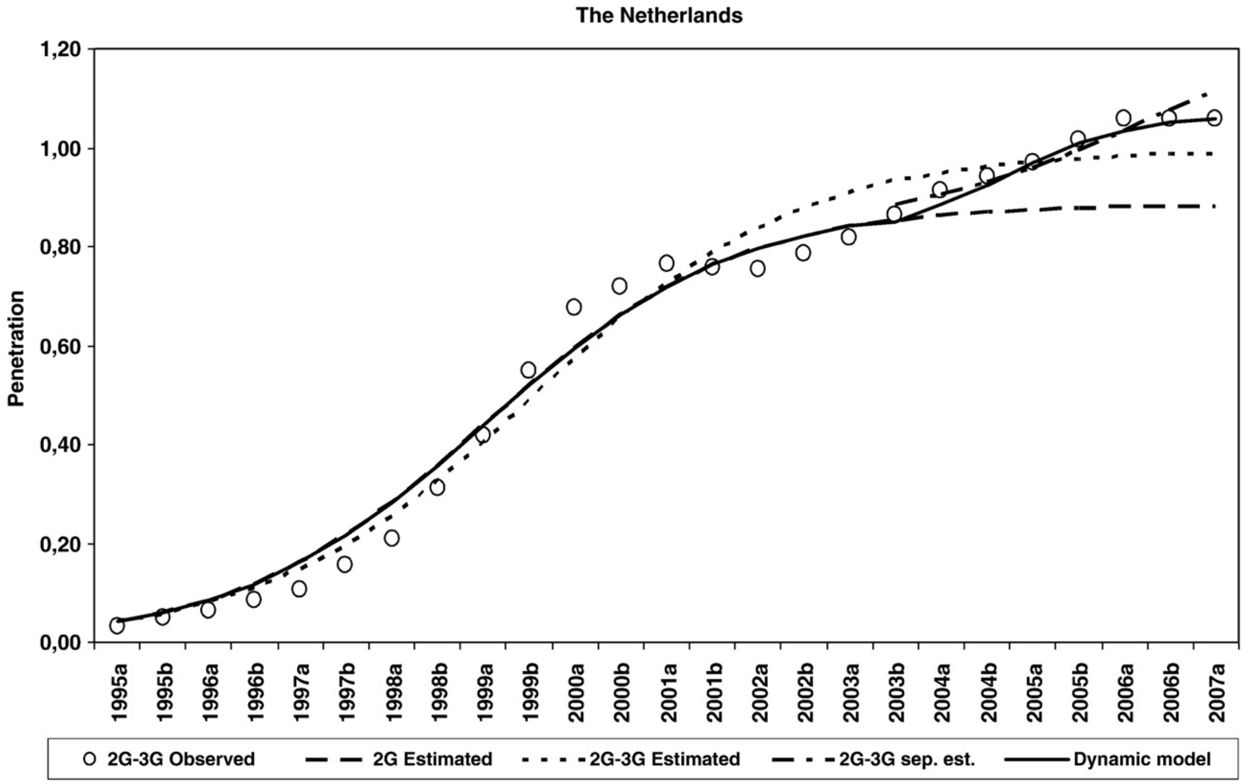

Fig. 3. Estimation of 2G and 3G diffusion in The Netherlands, vs. observed values.

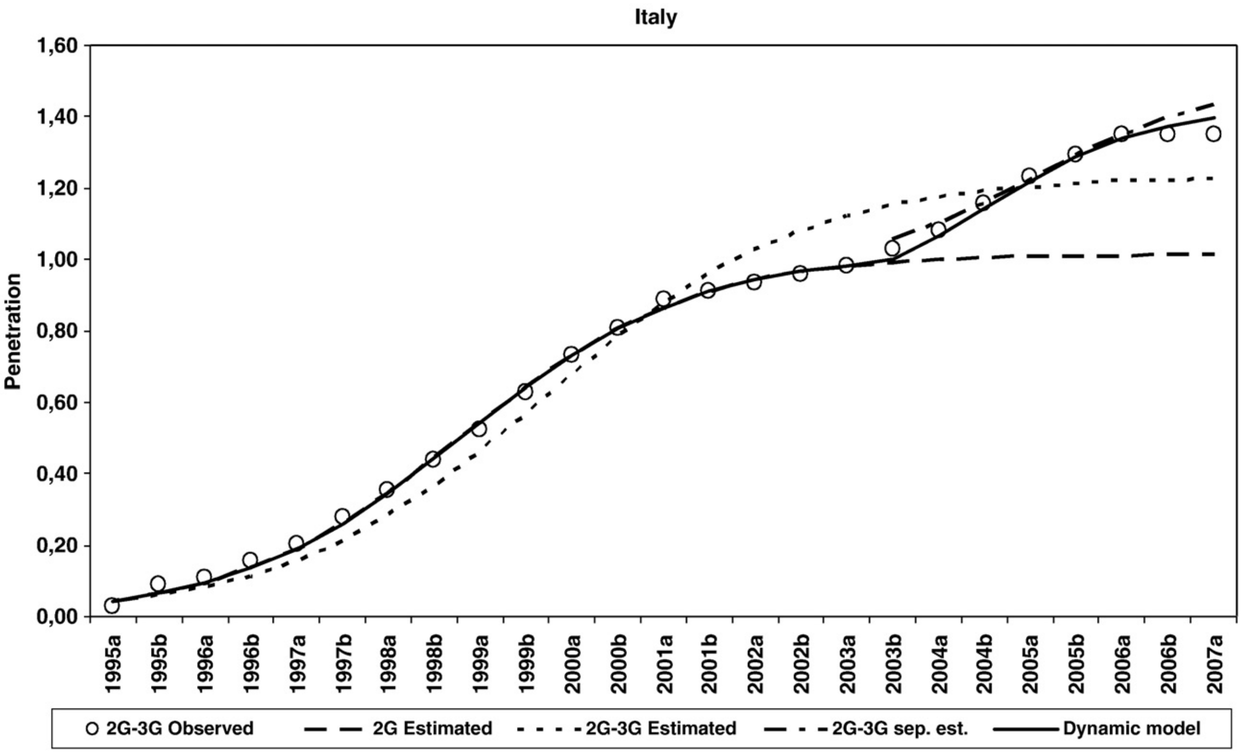

Fig. 4. Estimation of 2G and 3G diffusion in Italy, vs. observed values.

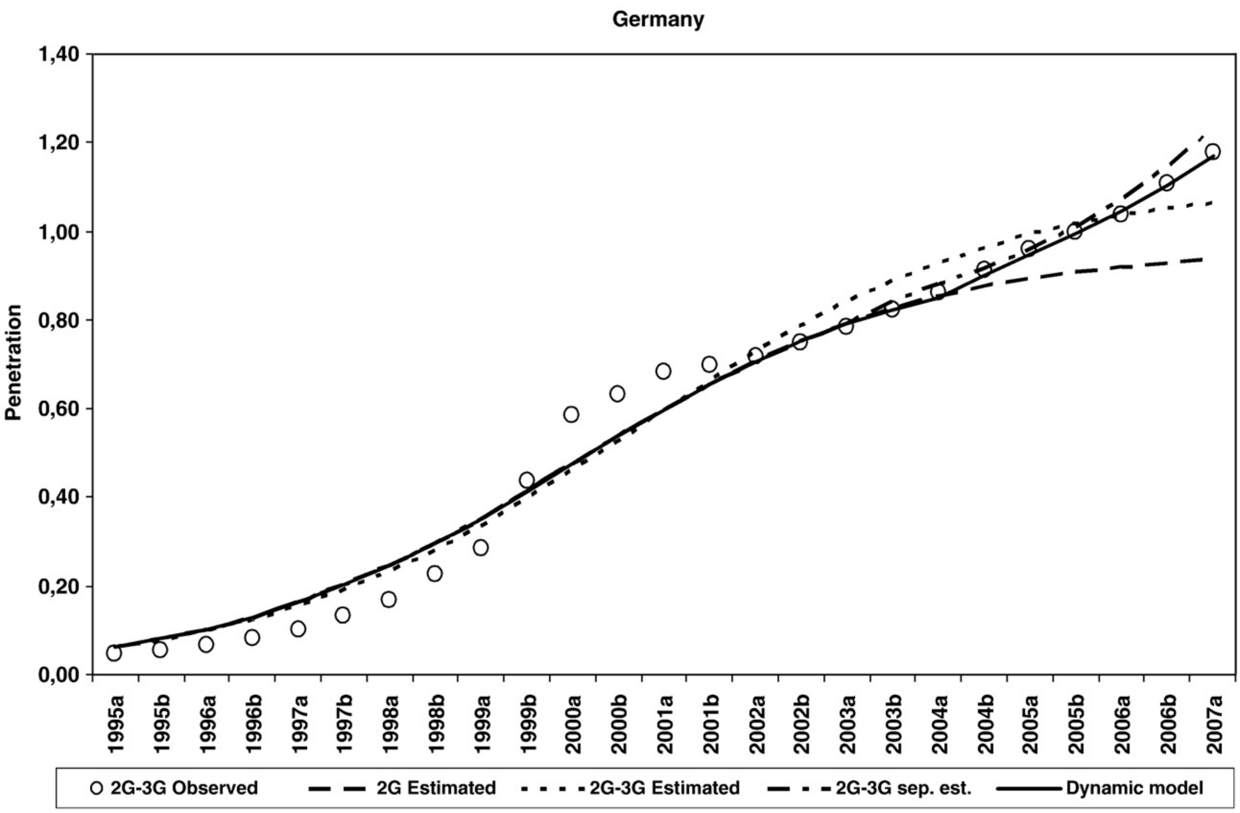

Fig. 5. Estimation of 2G and 3G diffusion in Germany, vs. observed values.

Table 1
Statistical measures of accuracy and substitution level.

		a	R^2	MSE	MAPE
Austria	Dynamic model	1.120	0.983	1.72E-03	0.164
	2G–3G sep. est.		0.978	1.92E-03	0.166
Finland	Dynamic model	0.875	0.974	3.21E-04	0.029
	2G–3G sep. est.		0.971	3.99E-04	0.030
The Netherlands	Dynamic model	1.132	0.986	1.39E-03	0.131
	2G–3G sep. est.		0.977	1.52E-03	0.133
Italy	Dynamic model	1.233	0.981	3.02E-04	0.058
	2G–3G sep. est.		0.968	5.55E-04	0.059
Germany	Dynamic model	0.987	0.976	2.37E-03	0.190
	2G–3G sep. est.		0.939	2.62E-03	0.195

Moreover, since$P(t) = \int_{t_0}^{t} \overline{N}(t)dt$, using Eq. (12) the following expression is derived (Appendix A contains the detailed derivation):

$$P(t) = (\overline{N}_0 - aM_0)(t - t_\partial) + \frac{a}{s} ln\left(\frac{x(t)}{x(t_0)}\right) \tag{13}$$

where:

$$x(t) = \left(\frac{\overline{M}}{M_0} - 1\right) + exp(s\overline{M}t) \tag{14}$$

Incorporating Eqs. (12) and (13) into Eq. (11) and solving the corresponding differential equation lead to the final expression for the diffusion of the first generation:

$$N(t) = \frac{\overline{N}_0 - aM_0 + a\left(\frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[-s\overline{M}(t - t_0)]}\right)}{1 + \left(\frac{\overline{N}_0}{N_0} - 1\right) exp\left[-r\left((\overline{N}_0 - aM_0)(t - t_0) + \frac{a}{s}\left[\frac{x(t)}{x(t_0)}\right]\right)\right]} \tag{15}$$

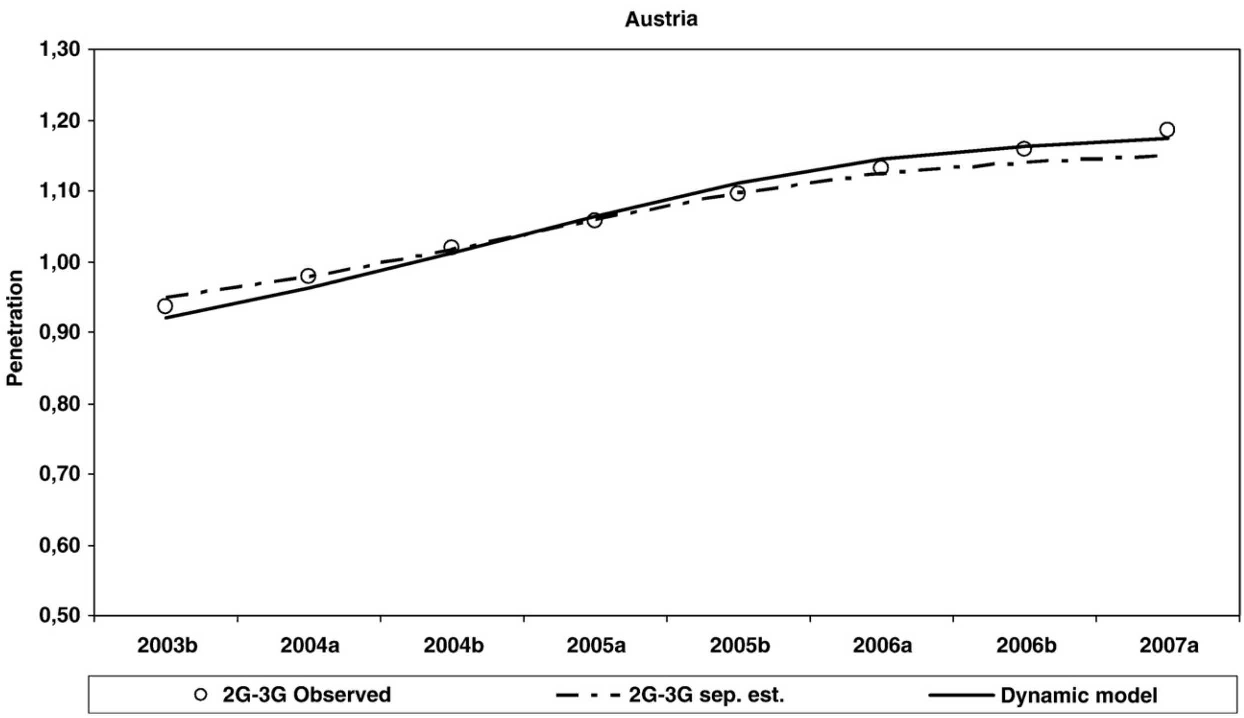

Fig. 6. Estimation of 2G and 3G diffusion in Germany, vs. observed values.

4. Evaluation

4.1. Evaluation procedure

Evaluation of the proposed methodology was performed over a number of five European countries (Austria, Finland, Netherlands, Italy, and Germany), based on historical data for 2G and 3G mobile telephony diffusion, as extracted from ITU's (International Telecommunications Union) database. Each dataset consists of two parts, corresponding to 2G and 3G diffusion data, respectively and the diffusion of mobile telephones is evaluated, in total.

The evaluation procedure for each country considered is performed as described by the following steps:

1. The diffusion process for 2G telephony is estimated based on the logistic model and the initial values of the saturation level and diffusion rate are used to describe the diffusion process until the next generation appears.
2. Based on the year that corresponds to 3G phones introduction in each particular country, t_0, the diffusion process for 3G phones is estimated, in terms of saturation level and diffusion rate Eq. (7).
3. The dynamic saturation level is estimated, as a function of the next generation's diffusion, by incorporating the results of the previous step into Eq. (12), thus deriving the diffusion of mobile telephony, including 3G as well. It is assumed that the 2G diffusion would follow the previously estimated path. This is a safe assumption, since estimation was based on a large number of observations, including the inflection point. Moreover, the market of the previous generation is almost saturated by the time the new generation is introduced.
4. The value of parameter a, appearing in Eq. (10), is estimated by using the regression analogue of Eq. (11) (derivation appearing in Appendix A) as: $\ln N(t) = r[\overline{N}_0 + aM(t) - aM_0 - N(t)]$.
5. Dynamic forecasts are provided by Eq. (15).

4.2. Results and discussion

Results for the evaluated cases are graphically illustrated in Figs. 1–5, where all the alternative estimation and forecasting approaches mentioned before are also included for comparison reasons. More specifically, the circles correspond to the observed values of mobile telephony (named as 2G–3G observed), the heavy dashed lines correspond to the estimation results based on the observed data of only 2G (2G estimated), the light dashed lines represent results using the conventional logistic model over the whole dataset (2G–3G estimated) and the dotted–dashed lines correspond to results of separate application of the logistic model over 2G and 3G actual data (2G–3G sep. est.). Finally, the solid lines illustrate the estimation results of the proposed dynamic model (Dynamic model).

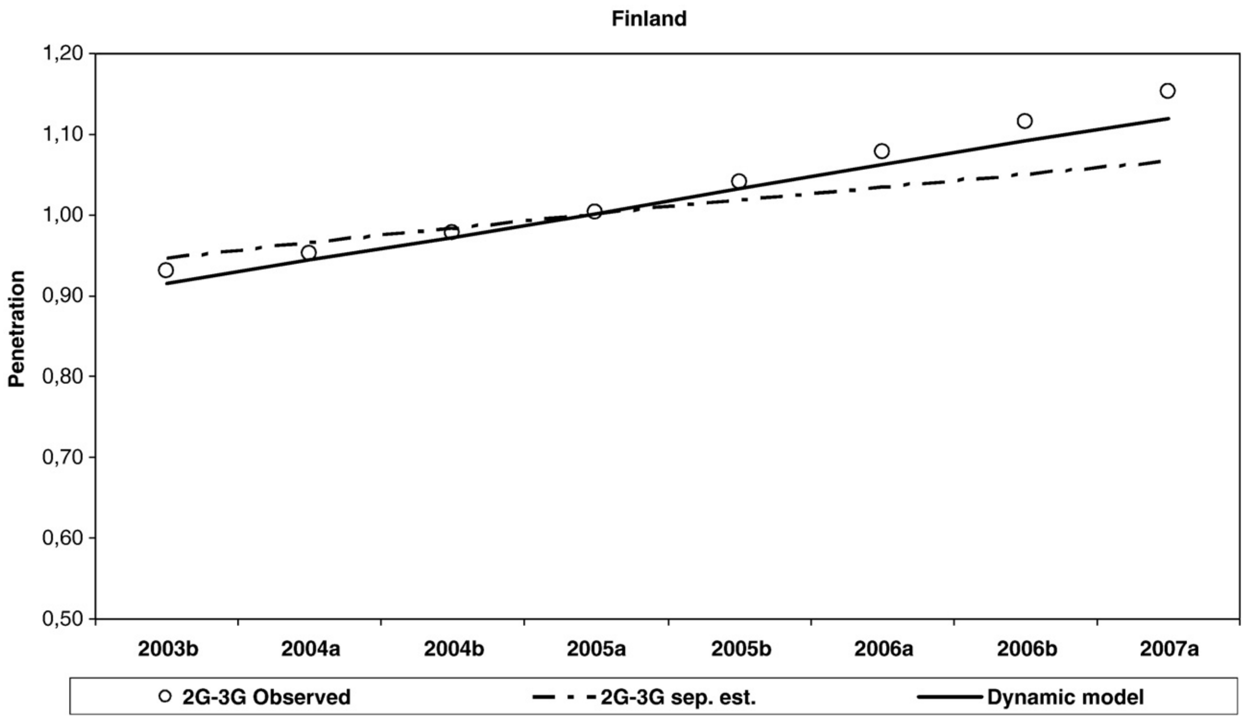

Fig. 7. Estimation of 2G and 3G diffusion in Germany, vs. observed values.

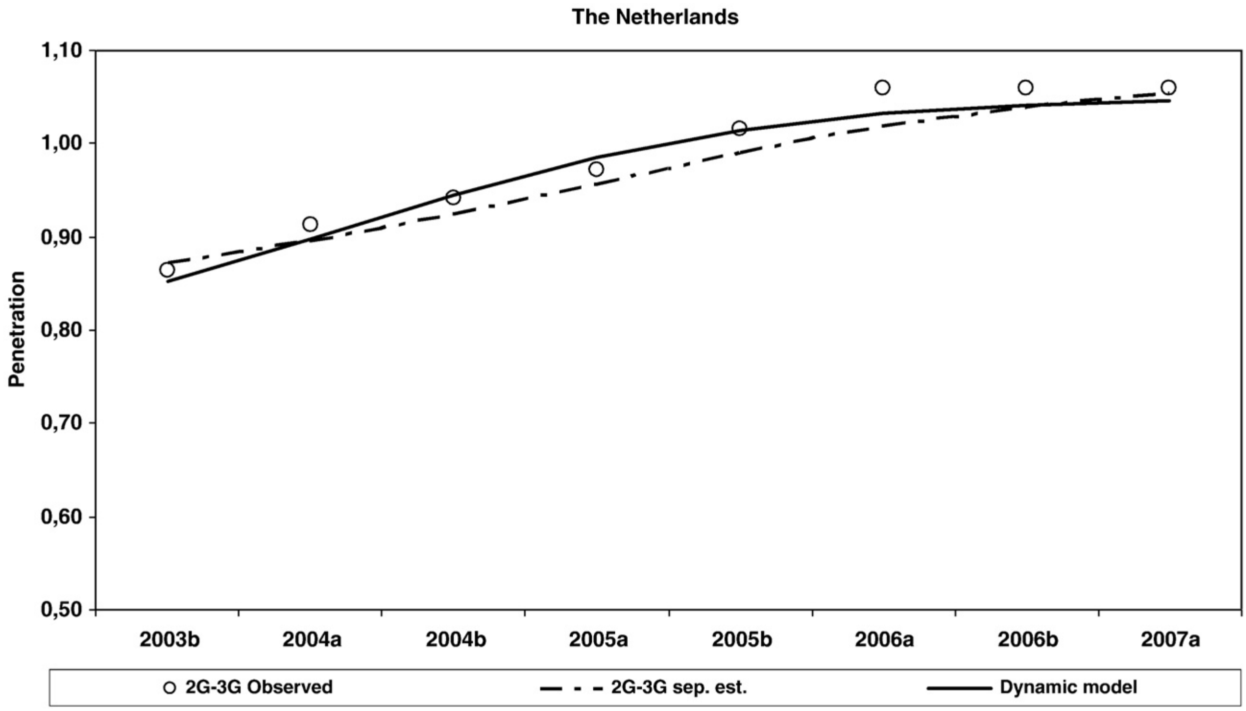

Fig. 8. Estimation of 2G and 3G diffusion in Germany, vs. observed values.

As observed, the results provided by the dynamic model outperform, in comparison with the alternative approaches. Even the separate application of a conventional diffusion model over 2G and 3G data, although it provides adequately accurate results, it does not manage to capture the existing market dynamics reflected over the whole process by 2G diffusion.

The accuracy of the results was validated by the statistical measures presented in Table 1. Since the application of the logistic model over either only 2G or the whole 2G–3G data provides quite inaccurate results, statistical measures for only the dynamic

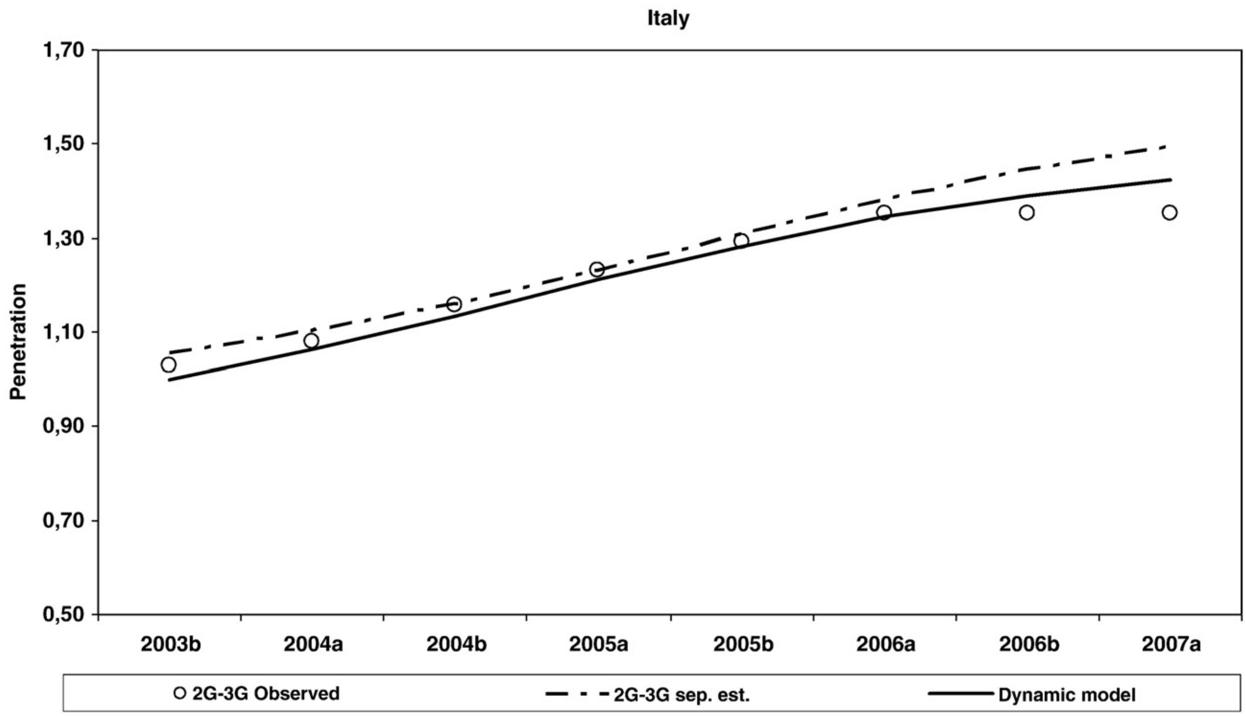

Fig. 9. Estimation of 2G and 3G diffusion in Germany, vs. observed values.

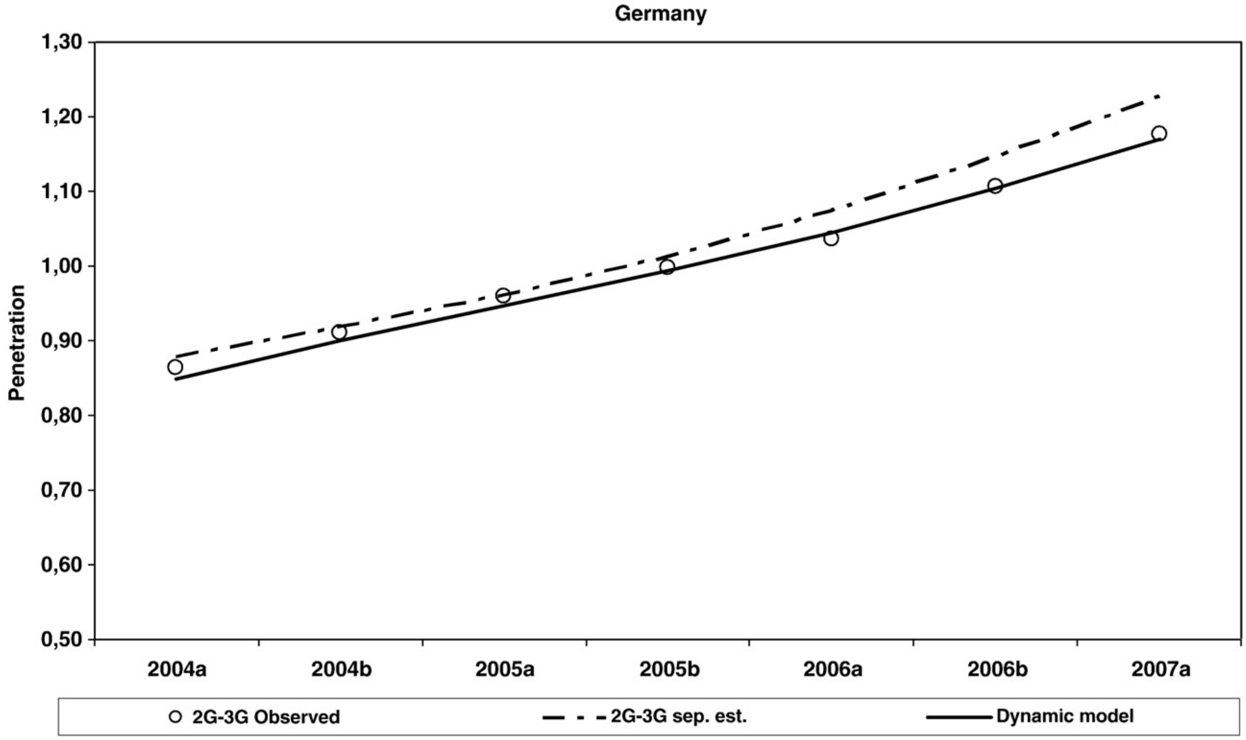

Fig. 10. Estimation of 2G and 3G diffusion in Germany, vs. observed values.

model and the separate application of the logistic model are calculated, for simplicity reasons (Dynamic model and 2G–3G sep. est., respectively). Moreover, in Table 1 the magnitude of generation substitution is presented, as this is captured by the parameter a of Eq. (8) (or, equivalently, Eq. (10)). According to the value of that parameter, the generation substitution was more intensive in Italy, followed by The Netherlands and Austria. In Germany and Finland the effect did not merit a high rate of substitution.

Apart from the results themselves, that show that the proposed methodology is able to provide highly accurate estimates, its superiority is not limited to the accuracy of the results only but is extended to a number of more and important aspects. Among them, the estimation of the substitution rate of the old by the new generation is included. This information is considered quite important, since it can provide valuable insights to the decision making process, regarding the expected diffusion of the new generation and the magnitude of substitution of the old one. This kind of knowledge should be incorporated into managerial plans for allocation of resources and infrastructure development, in order to meet the expected demand for all existing generations.

Moreover, as the rate of substitution is closely related to the timing of introduction of the next generation into the market, a careful study of the results will provide the necessary information for deciding the right timing. Indeed, it is obvious that an early introduction of the new generation will probably lead to low level of adoption, since the earlier technology would not have yet diffused into the market. On the other hand, a too late introduction would not necessarily succeed into obtaining a high diffusion level and rate, since the users of the older technology are used to it and would not be very willing to switch to the new one. Therefore, accurate approximation of the right time to introduce the new technological generation of the innovation would help avoiding dramatic consequences in supplying and investing.

Table 2
Statistical measures of accuracy for forecasting results.

		MSE	MAPE
Austria	Dynamic model	1.72E-03	0.1612
	2G–3G sep. est.	1.75E-03	0.1620
Finland	Dynamic model	2.62E-04	0.0271
	2G–3G sep. est.	7.36E-04	0.0324
The Netherlands	Dynamic model	1.38E-03	0.1310
	2G–3G sep. est.	1.45E-03	0.1328
Italy	Dynamic model	4.97E-04	0.0596
	2G–3G sep. est.	1.43E-03	0.0630
Germany	Dynamic model	2.37E-03	0.1900
	2G–3G sep. est.	2.62E-03	0.1949

The most important information that is probably provided by the dynamic model derives by the main assumption it is built upon, that diffusion of all generations is considered as a single process, incorporating multiple stages, each one corresponding to a successive generation which in turn inherits the dynamics of its predecessors, while retaining its own characteristics on the same time. Thus, the model succeeds in accurately estimating and forecasting the diffusion process, especially if compared to the alternatives presented in this evaluation. Indeed, the separate estimation of each generation leads to observable divergences between the actual and estimated adoption levels, especially at the beginning of the process of the new generation's diffusion. As the new generation diffuses usually slowly at the beginning, estimating the process separately, as if it was a standalone one, leaded to either underestimating or overestimating the actual recorded values. This modeling inaccuracy is observed during the process of forecasting as well, as presented in the next section.

5. Forecasting ability evaluation

Apart from the evaluation of the proposed dynamic model in terms of its estimation ability, its forecasting performance is also evaluated. Forecasting ability evaluation is based on leaving out a portion of the historical dataset (the "holdback" or "holdout" sample), consisting of the most recent observations, using the remaining part dataset to train the model's parameters (the "training sample") and try to forecast the holdback sample values.

Following the procedure described in the preceding paragraph and for the cases evaluated in this paper, the evaluation procedure is described by the following steps:

1) The three last observations are left out. These constitute the holdback sample.
2) All of the participating models are trained, based on the rest data and their corresponding parameters are estimated.
3) Based on the estimated parameters the left out data are forecasted, as if they were not already known.
4) Statistical measures of accuracy, for the training and the hold out sample as a whole, are calculated.
5) Following the results, corresponding conclusions are drawn.

The evaluation procedure was performed over the 3G data, apart from the dynamic model for the logistic model as well. Since forecasting refers mainly to the 3G data, results are illustrated for the corresponding periods of time and for each country. Results are graphically shown in Figs. 6–10 and illustrate the dynamic model's superior performance, managing to capture the market dynamics quite early in time and providing highly accurate forecasts. On the contrary, application of the logistic model over only 3G data under or overestimates the diffusion process. The corresponding calculated statistical measures of accuracy are presented in Table 2. Obviously, the proposed dynamic model is capable of providing quite accurate forecasts, quite early in time, when the diffusion of the new generation is still in the introductory stage.

As observed by the calculated statistical measures, presented in Table 2, the dynamic model can provide more accurate forecasting results, especially during the first stages of the new generation's diffusion, when not enough actual data are available for proper modeling. Moreover, if the incremental trend of the 3G diffusion will continue to exist, forecasting based on conventional approaches will probably provide more divergent results, at least until the inflection point of the 3G diffusion is met. This is not the case for the proposed model, which incorporates the dynamics of the preceding generations into the new one, and therefore it does not need too many observations to rely on, in order to provide accurate results.

6. Conclusion

Based on relevant literature, this paper has presented a methodology for studying the generation substitution effects, for the cases of multigenerational high technology products. After reviewing the relevant literature, the development of the proposed methodology is presented, followed by the corresponding evaluation, in terms of estimation and forecasting, providing accurate results, especially as compared with the rest considered approaches.

The major contribution of this work is the development of a dynamic diffusion model, capable of describing the diffusion process of a multigenerational high technology product, by incorporating the generation substitution effect. The main assumption made was that the saturation level process does not remain constant but it is affected by the diffusion of the next generation, as soon as the latter is introduced into the market. Consequently, the saturation level is expressed as a function of the next generation diffusion, leading to the formulation of the dynamic model. Moreover, the diffusion shapes of all generations of the product are considered as part of a single process, with each successive generation inheriting part of the potential of its predecessors. Evaluation of the methodology showed that the proposed model is capable of providing quite accurate results, not only in terms of diffusion estimation but also for forecasting purposes.

However, this methodology possesses some certain limitations which, in turn, indicate corresponding directions for future work. Apart from extending the proposed model in order to capture the effects of multiple generations, the proposed model does not take into account the effects of pricing and advertising strategies over the generation substitution process. Although this kind of diffusion models is able to capture the observed adopters' reactions at the macro level, incorporation of such decision variables would provide important information at the micro level, as to how corresponding managerial decisions influence the potential adopters' preferences. Moreover and despite the dynamic nature of the developed model, it still follows a deterministic approach. Therefore, considerations for stochastic extensions would give an interesting research direction, as incorporation of stochastic terms and randomness into the model would provide a set of the possible situations of the diffusion and substitution process.

Another important extension, regarding multigenerational products is the study of the intergeneration time itself. The study of the time lag between introductions of a product's successive generations is of great importance, in terms of forecasting future sales and adoptions. Development of a corresponding methodology for determining the appropriate timing for introducing the next generation into the market would provide important insights, helpful for managerial planning.

Appendix A

The diffusion model that describes penetration for the first generation is given by:

$$\frac{dN(t)}{dt} = rN(t)[\overline{N}(t) - N(t)] \tag{A – 1}$$

In Eq. (A-1) it holds that $\overline{N}(t) = f(M(t))$, based on the assumptions of the present work.
Solution of Eq. (A-1), for the time period after the introduction of the next generation, is derived as:

$$\begin{aligned}
&\frac{dN(t)}{N(t)[\overline{N}(t) - N(t)]} = rdt \Rightarrow \\
&\frac{dN(t)}{N(t)} + \frac{dN(t)}{\overline{N}(t) - N(t)} = r\overline{N}(t)dt \Rightarrow \\
&dlnN(t) - dln[\overline{N}(t) - N(t)] = r\overline{N}(t)dt \Rightarrow \\
&\int_{t0}^{t} dlnN(t) - \int_{t0}^{t} dln[\overline{N}(t) - N(t)] = r\int_{t0}^{t}\overline{N}(t)dt \Rightarrow \\
&lnN(t) - lnN_0 - ln[\overline{N}(t) - N(t)] + ln(\overline{N}_0 - N_0) = r\int_{t0}^{t}\overline{N}(t)dt
\end{aligned}$$

By setting $P(t) = \int_{t_0}^{t} \overline{N}(t)dt$ the last equation becomes:

$$\begin{aligned}
&ln\left[\frac{N(t)}{\overline{N}(t) - N(t)}\right] = ln\left[\frac{N_0}{\overline{N}_0 - N_0}\right] + rP(t) \Rightarrow \\
&\frac{N(t)}{\overline{N}(t) - N(t)} = \frac{N_0}{\overline{N}_0 - N_0} exp[rP(t)] \Rightarrow \\
&\frac{N(t)}{\overline{N}(t)} = 1 + \frac{N_0}{\overline{N}_0 - N_0} exp[rP(t)]
\end{aligned}$$

which leads to the solution:

$$N(t) = \frac{\overline{N}(t)}{1 + \left(\frac{\overline{N}_0}{N_0} - 1\right) exp[-rP(t)]} \tag{A – 2}$$

Regarding the next generation, corresponding penetration is described by the following equation:

$$\begin{aligned}
&\frac{dM(t)}{dt} = sM(t)[\overline{M} - M(t)] \Rightarrow \\
&\frac{dM(t)}{M(t)[\overline{M} - M(t)]} = sdt \Rightarrow \\
&\frac{dM(t)}{M(t)} + \frac{dM(t)}{\overline{M} - M(t)} = s\overline{M}dt \\
&dlnM(t) - dln(\overline{M} - M(t)) = s\overline{M}dt \Rightarrow \\
&ln(M(t)) - lnM_0 - ln(\overline{M} - M(t)) - ln(\overline{M} - M_0) = s\overline{M}(t - to) \Rightarrow \\
&ln\left(\frac{M(t)}{\overline{M} - M(t)}\right) = ln\left(\frac{M_0}{\overline{M} - M_0}\right) + s\overline{M}(t - to) \Rightarrow \\
&\frac{M(t)}{\overline{M} - M(t)} = \frac{M_0}{M - M_0} exp[s\overline{M}(t - t_0)]
\end{aligned}$$

and finally:

$$M(t) = \frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[-s\overline{M}(t - t_0)]} \quad (A-3)$$

Furthermore, it was assumed that $\overline{N}(t)$ is linearly dependent on $M(t)$ therefore:

$$\begin{aligned} &\frac{d\overline{N}(t)}{dM(t)} = a, \ \alpha: \text{constant} \\ &\text{So } \int_{t0}^{t} d\overline{N}(t) = a\int_{t0}^{t} dM(t) \\ &\overline{N}(t) - \overline{N}_0 = a[M(t) - M_0] \Rightarrow \\ &\overline{N}(t) = \overline{N}_0 + aM(t) - aM_0 \end{aligned} \quad (A-4)$$

By substituting Eq. (A-3) in Eq. (A-4):

$$\overline{N}(t) = \overline{N_0} - aM_0 + a\left(\frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[-s\overline{M}(t - t_0)]}\right) \quad (A-5)$$

Since $P(t) = \int_{t_0}^{t} \overline{N}(t)dt$

$$\begin{aligned} P(t) &= \int_{t_0}^{t}\left[\overline{N_0} - aM_0 + a\left(\frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[-s\overline{M}(t - t_0)]}\right)\right]dt \\ &= (\overline{N_0} - aM_0)(t - t_0) + a\int_{t_0}^{t}\left[\frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[-s\overline{M}(t - t_0)]}\right]dt \end{aligned} \quad (A-6)$$

Using the following transformation:

$$x(t) = 1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[s\overline{M}(t - t_0)] \quad (A-7)$$

the second part of Eq. (A-6) becomes:

$$\frac{a}{s}\int_{x(t_0)}^{x(t)} \frac{dx(t)}{x(t)} = \frac{a}{s} ln\left(\frac{x(t)}{x(t_0)}\right) \quad (A-8)$$

Finally, $P(t)$ is described by the following formulation:

$$P(t) = (\overline{N}_0 - aMo)(t - to) + \frac{a}{s} ln\left[\frac{x(t)}{x(t_0)}\right] \quad (A-9)$$

where $x(t)$ is given by Eq. (A-6).

Finally, incorporation of the above equations into Eq. (A-2) provides the final model formulation:

$$N(t) = \frac{\overline{N_0} - aM_0 + a\left(\frac{\overline{M}}{1 + \left(\frac{\overline{M}}{M_0} - 1\right) exp[-s\overline{M}(t - t_0)]}\right)}{1 + \left(\frac{\overline{N}_0}{N_0} - 1\right) exp\left[-r\left((\overline{N}_0 - aM_0)(t - t_0) + \frac{a}{s}\left[\frac{x(t)}{x(t_0)}\right]\right)\right]} \quad (A-10)$$

The regression analogue of Eq. (11) is derived as:

$$\begin{aligned}
&\frac{dN(t)}{dt} = rN(t)[\overline{N}_0 + aM(t) - aM_0 - N(t)] \Rightarrow \\
&\frac{dN(t)\big/dt}{N(t)} = r[\overline{N}_0 + aM(t) - aM_0 - N(t)] \Rightarrow \\
&\frac{dlnN(t)}{dt} = r[\overline{N}_0 + aM(t) - aM_0 - N(t)] \Rightarrow \\
&lnN(t+1) = lnN(t) = r[\overline{N}_0 + aM(t) - aM_0 - N(t)]
\end{aligned} \qquad (A-11)$$

Using Eq. (A-11), the value of parameter a can be estimated, based on the known values of the cumulative number of adopters at time t, for both generations, the diffusion rate of the first generation, r and the already estimated values of $\overline{N}_0$ and M_0.

References

[1] F.M. Bass, A new product growth model for consumer durables, Manag. Sci. 15 (1969) 215–227.
[2] E. Mansfield, Technical change and the rate of imitation, Econometrica 29 (1961) 741–766.
[3] E.M. Rogers, Diffusion of Innovations, The Free Press, New York, 1962.
[4] J.T.C. Teng, V. Grover, W. Guttler, Information technology innovations: general diffusion patterns and its relationships to innovation characteristics, IEEE Trans. Eng. Manage. 49 (Feb 2002) 13–27.
[5] J.D. Linton, Forecasting the market diffusion of disruptive and discontinuous innovation, IEEE Trans. Eng. Manage. 49 (Nov 2002) 365–374.
[6] N. Meade, T. Islam, Modelling and forecasting the diffusion of innovation — a 25-year review, Int. J. Forecast. 22 (2006) 519–545.
[7] J.H. Pae, D. Lehmann, Multigeneration innovation diffusion: the impact of intergeneration time, J. Acad. Mark. Sci. 31 (2003) 36–45.
[8] L.P. Rai, Appropriate models for technology substitution, J. Sci. Ind. Res. 58 (Jan 1999) 14–18.
[9] B. Gompertz, On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, Phil. Trans. Roy. Soc. London 115 (1825) 513–585.
[10] R. Bewley, D.G. Fiebig, A flexible logistic growth-model with applications in telecommunications, Int. J. Forecast. 4 (1988) 177–192.
[11] J.C. Fisher, R.H. Pry, A simple substitution model of technological change, Technol. Forecast. Soc. Change 3 (1971) 75–88.
[12] R. Fildes, P. Kumar, Telecommunications demand forecasting—a review, Int. J. Forecast. 18 (2002) 489–522.
[13] P.A. Geroski, Models of technology diffusion, Res. Policy 29 (Apr 2000) 603–625.
[14] A. Jain, L.P. Rai, Diffusion-models for technology-forecasting, J. Sci. Ind. Res. 47 (Aug 1988) 419–429.
[15] V. Mahajan, E. Muller, Innovation diffusion and new product growth-models in marketing, J. Mark. 43 (1979) 55–68.
[16] V. Mahajan, E. Muller, F.M. Bass, New product diffusion-models in marketing — a review and directions for research, J. Mark. 54 (Jan 1990) 1–26.
[17] C.H. Skiadas, 2 simple-models for the early and middle stage prediction of innovation diffusion, IEEE Trans. Eng. Manage. 34 (May 1987) 79–84.
[18] V. Kumar, T.V. Krishnan, Multinational diffusion models: an alternative framework, Mark. Sci. 21 (Sum 2002) 318–330.
[19] V. Kumar, J. Ganesh, R. Echambadi, Cross-national diffusion research: what do we know and how certain are we? J. Prod. Innov. Manag. 15 (May 1998) 255–268.
[20] C. Michalakelis, G. Dede, D. Varoutas, T. Sphicopoulos, Impact of cross-national diffusion process in telecommunications demand forecasting, Telecommun. Syst. 39 (2008) 51–60.
[21] E. Ruiz-Conde, S.H.P. Leeflang, E.J. Wieringa, Marketing variables in macro-level diffusion models, J. Betriebswirtsch. 56 (November 2006) 155–183.
[22] V. Mahajan, R.A. Peterson, First-purchase diffusion-models of new-product acceptance, Technol. Forecast. Soc. Change 15 (1979) 127–146.
[23] V. Peterka, Macrodynamics of Technological Change: Market Penetration by New Technologies, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.
[24] M.N. Sharif, C. Kabir, A generalized model for forecasting technological substitution, Technol. Forecast. Soc. Change 8 (1976) 353–364.
[25] A.W. Blackman, Market dynamics of technological substitutions, Technol. Forecast. Soc. Change 6 (1974) 41–63.
[26] J.A. Norton, F.M. Bass, Evolution of technological generations — the law of capture, Sloan Manage. Rev. 33 (Win 1992) 66–77.
[27] V. Mahajan, E. Muller, Timing, diffusion, and substitution of successive generations of technological innovations: the IBM mainframe case, Technol. Forecast. Soc. Change 51 (Feb 1996) 109–132.
[28] M.W. Speece, D.L. Maclachlan, Forecasting fluid milk package type with a multigeneration new product diffusion-model, IEEE Trans. Eng. Manage. 39 (May 1992) 169–175.
[29] T. Islam, N. Meade, The diffusion of successive generations of a technology: a more general model, Technol. Forecast. Soc. Change 56 (Sep 1997) 49–60.
[30] N. Kim, D.R. Chang, A.D. Shocker, Modeling intercategory and generational dynamics for a growing information technology industry, Manag. Sci. 46 (2000) 496–512.
[31] V. Mahajan, R.A. Peterson, Innovation diffusion in a dynamic potential adopter population, Manag. Sci. 24 (1978) 1589–1597.
[32] V. Mahajan, R.A. Peterson, A.K. Jain, N. Malhotra, New product growth-model with a dynamic market potential, Long Range Plan. 12 (1979) 51–58.

Christos Michalakelis holds a Mathematics degree (University of Athens, Department of Mathematics), an MSc. degree in Software Engineering (The University of Liverpool, UK) and an MSc. degree in “Administration and Economics of Telecommunication Networks” from the National and Kapodistrian University of Athens (Department of Informatics and Telecommunications, Interfaculty course of the Departments of Informatics and Telecommunications and Economic Sciences). He is a high school teacher of Informatics and Computer Science and he is currently working on his PhD degree, in the area of technoeconomics and especially in demand estimation and forecasting of high technology products. He has been working for 7 years with the Greek Ministry of Education, in the Managing Authority of Operational Program for Education and Initial Vocational Training, as an IT manager. Mr. Michalakelis has participated in a number of projects concerning the design and implementation of database systems and he now participates in several technoeconomic activities for telecommunications, networks and services such the CELTIC/ECOSYS project pricing and regulation. He has also developed or had a major contribution in the development of a number of information systems and applications. Mr Michalakelis's research interests include the study of diffusion of high technology products and services, the application of economic theories over the telecommunication and the wider area of informatics and price modeling and indexing. He has authored a number of papers presented to conferences and he has four published papers into scientific journals. He has also contributed as a co-author in two book chapters.

Dimitris Varoutas holds a Physics degree and M.Sc. and PhD diplomas in communications and technoeconomics from the University of Athens. He is a lecturer on telecommunications technoeconomics in the Department of Informatics and Telecommunications at the University of Athens. He has been participating in numerous European R&D projects in the RACE I &II, ACTS, Telematics, RISI and IST frameworks in the areas of telecommunications and technoeconomics. He actively participates in several technoeconomic activities for telecommunications, networks and services such the ICT-OMEGA and the CELTIC/CINEMA projects, as well as the Conferences on Telecommunications TechnoEconomics. He also participates in or manages related national activities for technoeconomic evaluation of broadband strategies, telecommunications demand forecasting, price modeling etc. His research interests span design of optical and wireless communications systems to technoeconomic evaluation of network architectures and services. He has published more than 40 publications in refereed journals and conferences in the area of telecommunications, optoelectronics and technoeconomics, including leading IEEE Journals and conferences. Dr. Varoutas is a member of LEOS, Communications, Circuits and Systems, Education and Engineering Management Societies of IEEE and serves as reviewer in several including IEEE journals and conferences.

Thomas Sphicopoulos received the Physics degree from Athens University in 1976, the D.E.A. degree and Doctorate in Electronics both from the University of Paris VI in 1977 and 1980 respectively, the Doctorat Es Science from the Ecole Polytechnique Federale de Lausanne in 1986. From 1976 to 1977 he worked in Thomson CSF Central Research Laboratories on Microwave Oscillators. From 1977 to 1980 he was an Associate Researcher in Thomson CSF Aeronautics Infrastructure Division. In 1980 he joined the Electromagnetism Laboratory of the Ecole Polytechnique Federal de Lausanne where he carried out research on Applied Electromagnetism. Since 1987 he has been with the Athens University engaged in research on Broadband Communications Systems. In 1990 he was elected as an Assistant Professor of Communications in the Department of Informatics and Telecommunications, in 1993 as Associate Professor and since 1998 he has been a Professor in the same Department. His main scientific interests are Optical Communication Systems and Networks and Technoeconomics. He has lead about 50 National and European R&D projects. Professor Sphicopoulos has more than 150 publications in scientific journals and conference proceedings and he is an advisor in several organizations.