Telecommun Syst (2007) 36: 173–179
DOI 10.1007/s11235-008-9066-4

Nonparametric estimation of a hedonic price index for ADSL connections in the European market using the Akaike Information Criterion

C. Deligiorgi · C. Michalakelis · A. Vavoulas · D. Varoutas

Published online: 21 February 2008

Abstract It is evidence that the demand for Asymmetric Digital Subscriber Line (ADSL) connections increases day by day in all European countries as much as worldwide and it is well known that Internet's penetration is considerably amazing. At the same time, due to the fierce competition among ADSL connections providers, several packages are offered in attractive tariffs. As a product consists of various characteristics that consumers value, the question that arise could be summarized to the following: How should consumers' choices and preferences for ADSL connections affect tariffs and what are the more significant and powerful characteristics that shape tariffs of ADSL connections?

This paper provides a hedonic price analysis of ADSL connections for the European market. A problem which is posed is the selection of the best model. So, in order to estimate prices a sliced inverse regression (SIR) is performed, without knowing the shape of the function. Then by applying Local Polynomial Regression (LPR) a possible shape of the hedonic function is given. Among several candidate hedonic models and by applying Akaike Information Criterion (AIC), the best one is derived. In order to evaluate the proposed methodology, tariffs' data have been collected from 15 European countries over the period from 2003 to 2005. Apart from tariffs, information on characteristics such as supported data rate (DR), maximum consumed data volume (V) and maximum allowed minutes on line (T) have been collected and analyzed.

Keywords ADSL connections · Hedonic method · Local polynomial regression · Sliced inverse regression · Akaike Information Criterion

1 Introduction

Over the last 5 years, the worldwide demand for ADSL connections has considerably increased. Europe has also met a corresponding substantial increase in demand for the xDSL technology. The Internet penetration rate in whole Europe poses several problems to the market characteristics. At the same time, fierce competition has emerged among providers, whose market shares vary heavily even within a month's time as several service baskets are offered in attractive prices. This continuous market share change is related to the competitive market where providers wish to attract new customers while preserving the existing ones, resulting in the decrement of prices [1]. As far as ADSL connections are concerned, the corresponding service baskets are continually fitted to the changing preferences and increasing expectations and needs of customers. The latter arises the following question:

- *How corresponding tariffs are affected by consumers' choices and preferences for ADSL connections?*
- *Which are the more dominant characteristics that shape tariffs of ADSL connections?*

This paper provides an empirical analysis of ADSL connections, using a non parametric estimation of hedonic

C. Deligiorgi (✉) · C. Michalakelis · A. Vavoulas · D. Varoutas
Department of Informatics and Telecommunications, University of Athens, Athens 15784, Greece
e-mail: ntina@di.uoa.gr

C. Michalakelis
e-mail: michalak@di.uoa.gr

A. Vavoulas
e-mail: vavoulas@di.uoa.gr

D. Varoutas
e-mail: arkas@di.uoa.gr

prices. More specifically, tariffs' data have been collected over the period from 2003 to 2005. Apart from tariffs, data that specify ADSL connections, such as supported data rate, maximum consumed data volume and maximum allowed minutes on line have also been collected and analyzed. In order to identify the most appropriate function that describes the influence of these characteristics over price, a SIR is performed, and LPR with kernel smoothing is applied, over a number of candidate functions. The above procedure approximates the possible shape of the underlying hedonic function.

The paper is structured as follows: Section 2 describes a theoretical approach of hedonic functions analysis and non-parametric models. The market for ADSL connections is described in Section 3, which also provides more information on the data used. Section 4 presents the evaluation of hedonic regression models and finally Section 5 concludes.

2 Theoretical model

Hedonic methods refer to regression models in which product prices are related to product characteristics and the observed price of a product (service) is considered to be a function of these characteristics. Generally, hedonic methods are based on the idea that a service (product) is a bundle of characteristics and that consumers just buy bundles of product characteristics instead of the product itself. These methods can be used to construct a quality-adjusted price index of a service. Authors in [2] and [3] described an overview on hedonic price equations. In [4] it is presented that from a large amount of product varieties, consumer chooses without influencing prices. Therefore, consumers maximize utility and producers maximize profits. In hedonic studies it is possible to adjust the price of a service for its quality not quantity. All of them are based on some estimated coefficients that are inflicted on the characteristics of the products in both periods; m and $m+1$. Someone can estimate the coefficients for every year separately or can have observations of two or all years together and estimate a common set of coefficients. The advantage of this method is that the included calculations are easy and fast to implement. Hedonic methods are indeed very fast to apply but the disadvantage is that index price can change even if no new products are existed, or all prices remain the same.

The hedonic price indices are commonly used as approximations to the true cost-of-living indices (COLI) which indicate how much money a consumer would need in period $m+1$ relatively to the amount of money he needed in period m so as to keep the same level of utility [5]. The solution is to determine the consumer's profile and his reaction to a varied and fast-changing supply of products. But how can this profile be determined when everyone has different needs and requirements? No matter what profile is decided, it will be a hypothesis and an assumption that will correspond to a specific model. In addition to the above, someone can see that consumer's desire is not stable and this is not unreasonable because there is a great offer as the 'goods' of technology become more and more attractive. However according to this approach the price index is constructed only using the prices of products, which are available in two adjacent periods.

2.1 Hedonic method

The term 'hedonic methods' refers to the use of a 'hedonic function' $f(X)$ in economic measurement, where

$$P_i = f(X_i) \tag{1}$$

where P_i is the price of a variety (or a model) i of a product and X_i is a vector of characteristics associated with the variety. The hedonic function is then used, for different characteristics among varieties of the product, in calculating the price index. As soon as the characteristics to be considered are determined then, for N telecommunication products in period m and in period $m+1$ the following equations must be evaluated:

$$\ln(p_{im}) = b_0 + b_1 \cdot X_{1i} + b_2 \cdot X_{2i} + u_{im}, \quad i = 1, \dots, N \tag{2}$$

$$\ln(p_{im+1}) = b_0 + b_1 \cdot X_{1i} + b_2 \cdot X_{2i} + b_3 + u_{im+1}, \quad i = 1, \dots, N \tag{3}$$

where b_i are some coefficients that have to be estimated.

2.2 True and candidate models by using the hedonic function

In this approach there is a set of consumers who have preferences over some characteristics of a service. The construction of a price index is complicated by product-pricing limits such as different charges for various characteristics. The first thing to be done is to define a basket of services (products). The model which describes the attributes (characteristics) of a product and their prices is given by the function:

$$P_i = f(X_b) + u \tag{4}$$

where $X = (x_1, x_2, \dots, x_n)$ is an $n \times p$ matrix of random regressors values, x_i and b are $p \times 1$ vectors and $f(X_b)$ is an $n \times 1$ vector with i-th component $f(x_i{}'b)$ $(i = 1, 2, 3, \dots, n)$, u for given $X = x$ is distributed as $N(0, \sigma^2)$ and σ is an unknown scalar. A problem which is posed is the selection of the best model. For instance, models supported by high R-square satisfying an important statistical level of significance, are frequently doubted because of weak assumptions or irrational estimates of parameters [6]. Furthermore, we assume that f is an unknown function, which means that its shape is unknown and b estimators have a unit norm. In a similar way to the true model, we construct the candidate models by the definition of characteristics that 'play' a significant role to the construction of the price index. In order to estimate the distance between the true and candidate models, is necessary the single-index model to be described.

2.3 Single-index model

Single-index models can be estimated by using iterative or direct methods. In the iterative method nonparametric regression is applied in order to be calculated the mean regressor b. By using the iterative method the computation is difficult because a requirement exists for an estimate of nonparametric mean regression at each data point, in order to compute a representative function f. On the contrary, by using direct method the computation is easier because the b relative weights are estimated by SIR [7] and the estimation of function f is not necessary. Once the relative weights, b_i, are estimated, an index $z = X\hat{b}_{SIR}$ is constructed. Then, by applying LPR [8], the nonlinear link $price = \hat{f}(z)$ is estimated.

After obtaining an estimation for f and b and by employing an iterative or direct method we compute:

$$\hat{\sigma}^2 = \frac{\{Y - \hat{f}(X\hat{b})\}'\{Y - \hat{f}(X\hat{b})\}}{n} \tag{5}$$

Using $(\hat{f}, \hat{b}, \hat{\sigma}^2)$ the most suitable model, among a variety of candidate models via the AIC, is determined.

2.4 Assumptions for AIC

In order to find out which is the best model (among a variety of candidates) that describe a product with a set of characteristics the following equation is used:

$$AIC = \log \hat{\sigma}^2 + \frac{1 + tr(\hat{H}_p + \hat{H}_{np} - \hat{H}_p\hat{H}_{np})/n}{1 - \{tr(\hat{H}_p + \hat{H}_{np} - \hat{H}_p\hat{H}_{np}) + 2\}/n} \tag{6}$$

The smaller AIC value the better model is described [11]. Because of the difficult computation, since there is an unknown function, without having great inclination, the following assumptions facilitate the computation of the AIC value:

Assumption 1 We have $\tilde{f}(Xb^*) \cong H_{np}Y$.

Assumption 2 Similarly we have $E_0\{\tilde{f}(Xb^*)\} \cong f_0(Xb^*)$.

Assumption 3 In addition to the above we have that

$$\hat{f}(X\hat{b}) - \tilde{f}(Xb^*) \cong \tilde{V}(\hat{b} - b^*) \cong H_p\{Y - \hat{f}(Xb^*)\}$$

where

$$H_p = \tilde{V}(\tilde{V}'\tilde{V})^{-1}\tilde{V}', \quad \tilde{V} = \partial\tilde{f}(Xb)/\partial b|_{b=b^*} = \tilde{f}(Xb^*)X.$$

3 The market for ADSL connections

3.1 Variables-characteristics

According to the economic theory described above regarding hedonic models, there are several characteristics that influence tariffs. Data by the meaning of attributes of a given product can be divided in three categories: spatial, physical and socio-economic [9].

The spatial division can be handled separate by location. Attributes or exogenous variables are made up of physical quality characteristics and socio-economic attributes. Measures and differences that compare geographical areas have implications for regional cost of living. This has a substantial effect not only to consumers' choices and preferences, but also to the governments' policies and market relocation as well. This is an index for decision-making affecting all involved parties.

The physical quality characteristics are the attributes which are unique and specify the product. For instance, a bundle of physical characteristics can be the offered data rate and the maximum consumed data volume. Therefore, physical characteristics define a basket of an ADSL connection and give a stream of services which a particular connection provides.

The socio-economic variables are characteristics that are not easily calculated, but it is assumed that influence the ADSL market expansion and penetration. Such characteristics are: the educational level, the age and the income of the subscribers, the technological infrastructure improvement, the year of telecommunications' liberalization etc. As to the customers' interests, they can be identified from the web pages they visit and the amount of time they spend on them. So it depends on personalization of customers and reflects their own interests, needs, knowledge background and goals. But it is extremely difficult to gather data, in order to understand customers' profile and characteristics, because on one hand the veracity of them is doubtful and on the other customers are not interested in providing information for privacy concerns. These variables can play a significant role on the ADSL market and especially on the tariffs' shape. A possible approach on this subject is to include in the model a hypothesis test with values zero and one, for 'having or not' a characteristic like the above. By including in the model the socio-economic variables implicit market's prices may arise for quality characteristics.

Table 1 Required bandwidth for typical broadband services

Service	Bandwidth (Mbps)
High Definition TV (HDTV)	16–20
Telemedicine	6
Video on Demand (VoD)	6–18
Internet access	1.5
Video conferencing	1.5
Telecommuting	3
Multiple digital TV	6–24

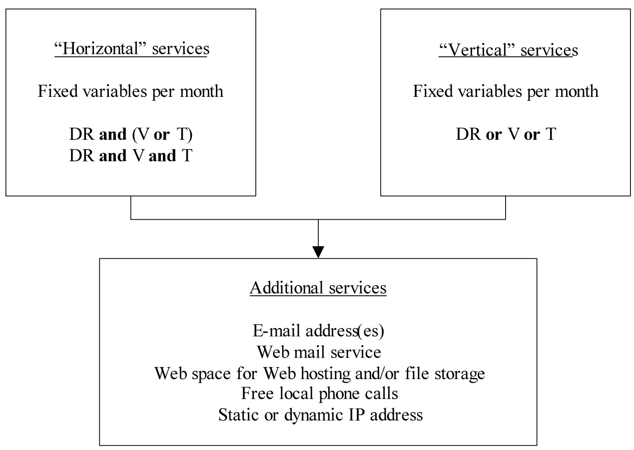


Fig. 1 Typical ADSL service basket

3.2 ADSL high-speed Internet connections

Broadband services and applications are classified according to the offered data rate. The domination of ADSL technology for broadband access across Europe during the last years demonstrated the high-speed Internet access and IP-telephone as the most common broadband services. However, a wide range of new services are now becoming present in the mass market with an increased penetration of new technologies (e.g. FTTx, ADSL2++, WiMAX). Table 1 summarizes typical broadband services that are expected to dominate in the near future.

In order to specify the basic basket of broadband services an extended survey of providers across European countries took place, focused on the services offered, the pricing policy as well as the development in broadband market. As a result, the typical ADSL service basket was determined as a combination of main and additional services according to Fig. 1. Main services are distinguished into "horizontal" and "vertical" according to the number of fixed variables among supported data rate (DR), maximum consumed data volume (V) and maximum allowed minutes on line (T). Additional services includes a number of email addresses, web space for web hosting and/or file storage and optional free local phone calls and static IP address. The choice of the appropriate combination for each operator is depending on the specific business plan as well as the techno-economic model parameters and assumptions.

The evolution of broadband technology offers new and challenging options. The EU Commission's "Broadband for all" policy is expected to grow the interest for broadband in the next years and to enforce the infrastructure competition among providers. As a result of this competitive environment the provision of enhanced broadband services with reduced tariffs is expected to increase significantly the number of broadband subscribers.

4 Evaluation of ADSL connections

4.1 Models with nonlinear functions

Let's have an example in which the characteristics of the product are determined, so as to show the quality of this product. As soon as the characteristics of a product are determined, applying the hedonic method someone estimates its price. Otherwise the price of a product can be estimated "manually" by comparing the new product with the most similar old one.

The main assumption is that ADSL connections have four main physical characteristics: supported DR (up and down), maximum V and maximum T. These four characteristics are widely used from telecom operators for valuating and selling leased lines across Europe.

First of all, the single-index model is estimated, described by (1). The data are sorted by P_i and then are divided into three slices. Without specifying the unknown link function we derive:

$$\widehat{b}_{SIR} = (0.581899, -0.78326, -0.00011, -0.21886)$$

The above results imply that the price is strongly related only with the downlink DR [10].

As it can be seen after inspecting the diagram, the described relationship can be easily expressed as a linear model.

By having four characteristics we take under consideration $2^4 - 1$ nested candidate models. For each of the nested models SIR estimates are obtained (Fig. 1) and then by applying the LPR (Fig. 2) the link function $\widehat{f}_k$ $(k = 1, 2, 3, \ldots, 15)$. Figure 2 also shows that some individual tariffs are existing which decline significantly from the main cluster.

In order to examine the relationship between the price of an ADSL connection and their main characteristics, such as the supported data rate and the maximum consumed data, several candidate link functions are applied. Across the candidate models in several shapes of link function, the

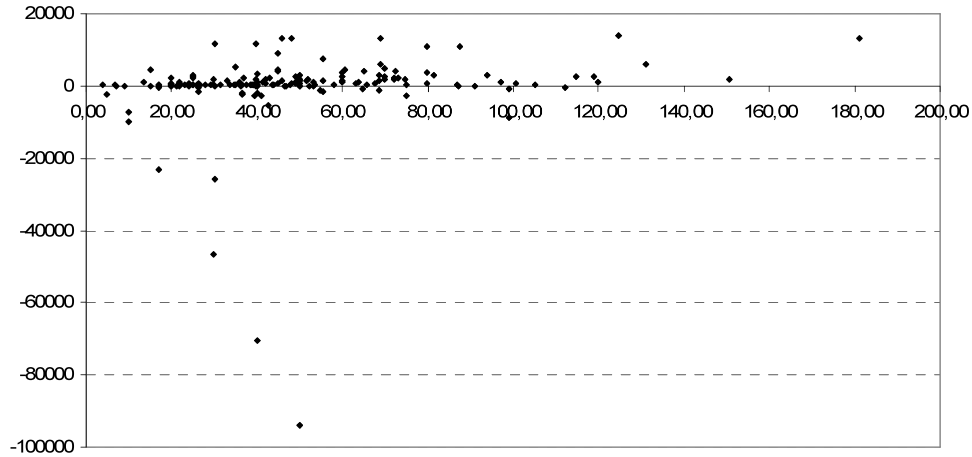


Fig. 2 Plot of P_i against SIR directions

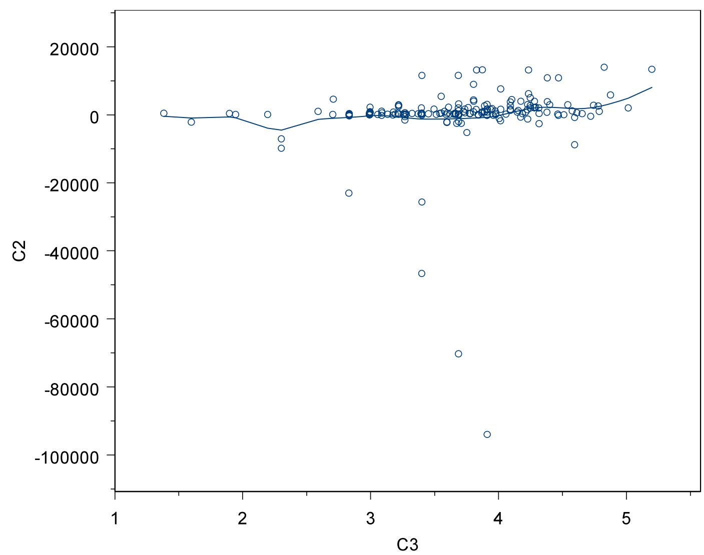


Fig. 3 Local polynomial regression with kernel smoothing

one which gives the smallest AIC (0.2125746 vs 7.846617) value from (4), is described by the equation:

$$\ln(p) = b_0 + b_1 \cdot DR(down) + b_2 \cdot DR(up) + b_3 \cdot (V) + b_4 \cdot (T) \tag{7}$$

Even link functions such as hyperbolic sine or hyperbolic cosine give almost the same results with the non linear functions. Although linear model on logged price is not comparable with all the other models, because of the AIC value, linear model on logged scale has an advantage.

Working in the logarithmic scale using a linear model results shows a better fit than all the other models because first of all the residuals from the log-linear model are all around zero (Fig. 3) [10] and less standard error (0.5845 vs 26.58).

Using data such as those presented in Table 2 for the case of Europe, it can been observed that it is not easy to compare prices for different data rate and consumed volume and allowed time on line, but there are similarities and patterns

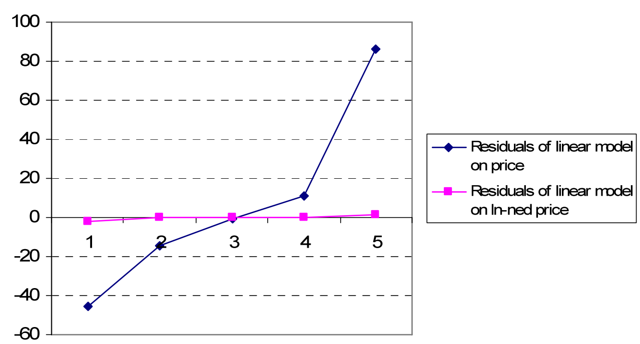


Fig. 4 Residuals

that must be evaluated. It is obvious that the more a consumer demands for a product with upper services, the more the prices of this product are increased. Because of this consumers' 'behavior' there is no implicit prices for all characteristics.

Table 2 Data for ADSL connections

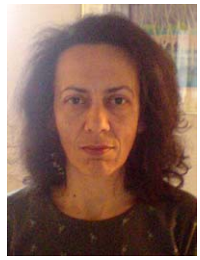
Subscription price (in €)	Data rate (down) (in Kbps)	Data rate (up) (in Kbps)	Volume (in KB)
59.95	10,000	512	15,728,640
14.95	256	64	314,572.80
21.95	1536	256	1,572,864
29.95	800	256	–
74.95	12,000	1024	8,388,608

5 Conclusions

The purpose of this analysis is to show the parameters that influence the prices of telecommunication services and the model which is the most appropriate for telecommunication services (products) and especially ADSL connections over time across all countries of Europe, applying a hedonic method for some defined characteristics. ADSL connections have important and specific characteristics indeed and their prices vary slowly over time where as their characteristics change in a fast way. The application of these econometric methods, following the definition of products' characteristics, provides a reliable and accurate method able to produce an exact estimate of prices both for new products and over next years. The validity of the model and the appropriate selection of the functional form that has been chosen to relate price to characteristics can be validated over next years and more observations.

References

1. Dewenter, R., Haucap, J., Luther, R., & Rotzel, P. (2004). *Hedonic prices in the German market for mobile phones* (Discussion paper No. 29). University of the Federal Armed Forces Hamburg.
2. Berndt, E. R. (1991). *The practice of econometrics: classic and contemporary*. Reading: Addison-Wesley.
3. Triplett, J. E. (2000). *Draft copy Handbook on quality adjustment of price indexes for information and communication technology products*. Paris: OECD.
4. Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, *92*, 34–55.
5. Jonker, N. (2001). *Constructing quality adjusted price indexes: a comparison of hedonic and discrete choice methods*. De Nederlandsche bank, Econometric Research and Special Studies Department.
6. Gao, X., Asami, Y., & Chung, C. J. (2002). *An empirical evaluation of hedonic regression models* (CSIS discussion paper No. 46). Center for Spatial Information Science.
7. Li, K. C. (1991). Sliced inverse regression for dimension reduction (with Discussion). *Journal of American Statistical Association*, *86*, 316–42.
8. Simonoff, J. S. (1996). *Smoothing methods in statistics*. Berlin: Springer.
9. Chowhan, J., & Prud'homme, M. (2004) *City comparisons of shelter costs in Canada: a hedonic approach*. Prices division.
10. Triplett, J. E. (2004). *Hand book on hedonic indexes and quality adjustments in price indexes* (OECD science, technology and industry working papers, 2004/9).
11. Naik, P. A., & Tsai, C.-L. (2001). Single-index model. *Biometrica Trust*, *88*(3), 821–832.



C. Deligiorgi holds a Mathematics degree and M.Sc. in Information Technology and Telecommunications from the University of Athens. She is a Ph.D. candidate in the area of telecommunications and technoeconomics in the Department of Informatics and Telecommunications at the same University. She works in the IT department of the Greek Ministry of Internal Affairs. She actively participates in several projects in the fields of telecommunications, broadband services, software and hardware installations, network architectures and services under the Greek Information Society Programme. Her research interests include technoeconomic evaluation of telecommunications as well as telecommunications price indexing. She is a member of Greek Computer Society and Hellenic Mathematical Society and a certificated adult educator for Information Technology.

C. Michalakelis holds a Mathematics degree (University of Athens, Department of Mathematics) and an M.Sc. degree in Software Engineering (The University of Liverpool, UK). He is a high school teacher of Informatics and Computer Science and he is currently working on his Ph.D. 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 supervisor. 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 participated in the development of a number of information systems and applications.
His research interests include the study of diffusion models over the technology products and services, the application of economic theories over the telecommunication and wider informatics area, price modeling and indexing. He has submitted for publication a number of 3 papers and he has contributed in the writing of two book chapters.

A. Vavoulas holds a B.Sc. degree in physics and M.Sc. degree in electronics and radiocommunications from the University of Athens, Greece. He is currently working toward the Ph.D. degree at the same University, in the area of radio resource allocation techniques. He has participated in many National and European R&D projects in the areas of telecommunications and technoeconomics. His research interests include physical and MAC layer issues for fixed broadband wireless access systems and technoeconomic evaluation of network services. He is a student member of Communications and Vehicular Technology Societies of IEEE and serves as reviewer in several journals and conferences.

D. Varoutas holds a Physics degree and M.Sc. and Ph.D. 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 modelling 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. He 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.