Analysing Wine Demand With Artificial Neural Networks

M. Gerolimetto[1], C. Mauracher[2], I. Procidano[2]
[1] Dep. of Statistical Sciences - Padova University, e-mail: gerolime@stat.unipd.it
[2] Dep. of Statistics and Section of Agricultural Economics and Politics –
Ca' Foscari University of Venice, e-mail: maurache@unive.it; isabella@unive.it

Paper prepared for presentation at the 11th Congress of the EAAE
(European Association of Agricultural Economists), Copenhagen,
Denmark, August 24-27, 2005

ANALYSING WINE DEMAND WITH ARTIFICIAL NEURAL NETWORKS

Abstract. In this paper we analyse wine demand in Italy using microdata. Instead of estimating a traditional parametric model (like AIDS) we employed artificial neural networks (ANN) and evaluate the elasticities using two different methods, specific for the non parametric framework. We compared the performances of the two methods to estimate elasticities and put in evidence the relevance of some demographic variables together with the usual economic ones, explaining the consumer's behaviour.
Keywords: Artificial Neural Networks, Demand Analysis, Wine, Elasticity; C14, C21, Q11, Q13

1. Introduction

Demand elasticities are of considerable interest for policy purposes. Market demand for various foods and beverages is an important component in the formation of agricultural, as well as other public policies in most industrialized countries. Demand functions are, for example, used to evaluate the effects of changes in target prices on farmers' income and also to forecast. Farmers, food processors and retailers need to forecast demand to plan their production and sales, and demand elasticities are of crucial importance.

In this paper the consumption behaviour in the Italian wine market is modelled using the Artificial Neural Network (ANN), in particular the demand curve and the elasticities are estimated. It is important to point out that among the input variables some socio-demographic variables are included, since it seems that together with the traditional economic variables (price and income) also some non economic factors, such as the demographic structure of the population, can affect the consumers' choices.

Unlike the more traditional demand systems models (AIDS, Rotterdam model) that are parametric, the neural network approach is non parametric and therefore more flexible since it is capable to capture complex non linear interactions between input and output variables in a system. Also when computing the elasticities, which is one of the aims of this paper, some important differences between the parametric and the non parametric approach are evident. Indeed, in the parametric scheme the elasticities can be computed directly from the model parameters, with a closed form solution, whereas following a ANN approach the elasticities are computed in a far completely different way, which is usually based on the definition of elasticities. In literature there are only a few proposals on how to compute the elasticities in the non parametric context and in this paper we considered the Subba Rao method (Subba Rao *et al*, 1998) and the Gruca method (Gruca and Klemz, 1998).

We estimate the demand curve and compute the elasticities following the Subba Rao method and then for a few specific cases we determine the elasticities with both procedures in order to make a comparison of the performances.

In the second section of this paper we describe the data set we used with an overview on the data and the wine consumption in Italy. In the third section we give some basic notions about ANN, in the fourth section we explain the two methods to compute the elasticities, in the fifth section we show the results we obtained by applying the ANN to the Italian wine market, sixth section concludes.

2. The data set

2.1 General description of the data

In this paper we work with microdata and the data set we are using consists of a sample of 4245 households, representative of the entire national universe.

Data come from Osservatorio AcNielsen about the Italian households expenditures in year 2003 and the survey does not consider the so-called "away from home expenditures" (hotels, restaurants and coffees) and the expenditure for wine obtained from particular sources, for example, directly from the producers.

Therefore, we are working on a data set that is constructed by monitoring the expenditures realized through the big distribution, cash and carry and traditional commerce. Data have been collected by a home scanning method, that registers (using the bar code) all the products bought by the family. The advantage of this method is that a big amount of details and information about the product is so available. For example, the colour of wine (red, white etc.) is known, as well as the specific type of wine (prosecco, chardonnay, malvasia, etc.).

Data are available in a geographically disaggregated version, meaning that we have specific data sets for North-West (Piemonte, Valle d'Aosta, Lombardia and Liguria), North-East (Veneto, Friuli, Trentino ed Emilia), Centre (Toscana, Umbria, Marche, Lazio e Sardegna), South (Abruzzo, Campania, Calabria, Basilicata, Molise, Puglia e Sicilia).

Some demographic information is also available, particularly some details about the reference person of the household, for example his\her education level, social position (employed, unemployed, retired etc.) etc.

Now we present some tables about the main socio-demographic characteristics of the sample, where with the short form R.P. we intend the reference person of the family.

Table 1. Gender of the R.P.

GENDER R.P.	Abs. frequency	Percentage
male	3.709	87,37%
female	536	12,63%

Table 2. Age of the R.P.

AGE R.P.	Abs. frequency	Percentage
Up to 44 years old	1.110	26,15%
from 45 to 60 years old	1.825	42,99%
over 60 years old	1.310	30,86%

Table 3. Education level of the R.P.

EDUCATION LEVEL	Abs. frequency	Percentage
high	556	13,10%
medium	1.941	45,72%
low	1.748	41,18%

In table 3 in the "high level education" class, people who obtained at least the short graduation diploma (laurea breve) are included, whereas in the "medium and low level education" classes belong to those who obtained, respectively, a high school diploma and middle school diploma.

Table 4. Social position of the R.P.

SOCIAL POSITION OF R.P.	Abs. frequency	Percentage
Employed	2.871	67,64%
Unemployed	60	1,41%
Other	1.314	30,95%

In the category "Other", housewives, students, people unable to work, retired and joining the army are included.

Table 5. Income pro capita

INCOME pro capita	Abs. frequency	Percentage
Less than 260 € per month	365	8,60%
from 260 to 420 €	972	22,90%
from 420 to 620 €	1.241	29,23%
over 620 €	1.667	39,27%

Table 6. Geographic repartition

GEOGRAPHIC AREA	Abs. frequency	Percentage
North-West	1.266	29,82%
Nord-East	708	16,68%
Centre	840	19,79%
South	1.431	33,71%

2.2 Wine consumption and demographic variables

Now we present a sequence of graphs about the different levels of wine consumption according to some socio-demographic variables.

With reference to the geographic area (fig.1), it emerges that the biggest part of the consumption is in the North West of Italy (39%). As far as the income is concerned (fig. 2), the lowest average consumptions are exhibited by the first income class (less than 260 € per month), whereas the highest consumptions are in the fourth income class (over 620 €). Therefore, it seems that with the increasing of the p.c. income, also the consumptions tend to increase, demonstrating how they are related to the effective purchase capability.

Figure 1. Global consumption repartition per geographic area

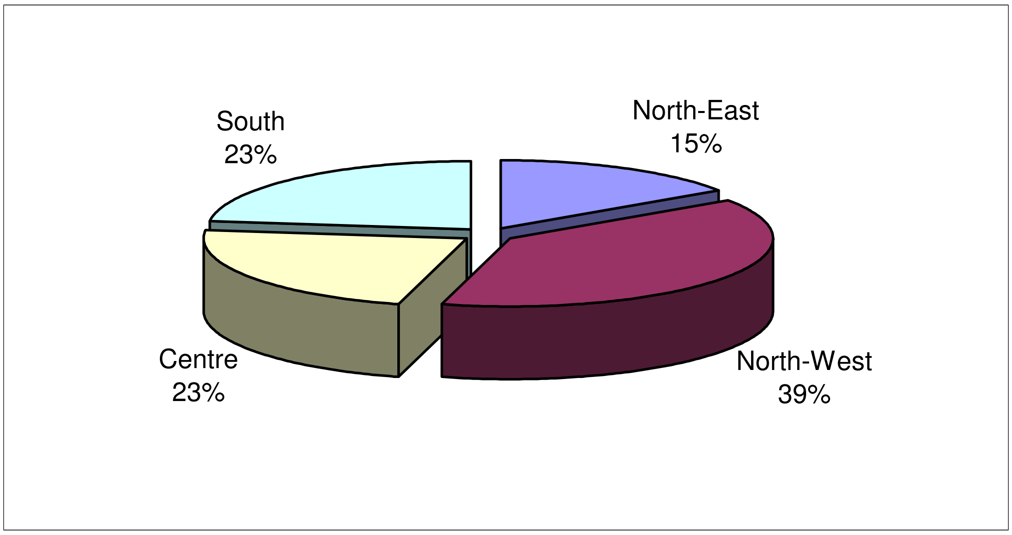
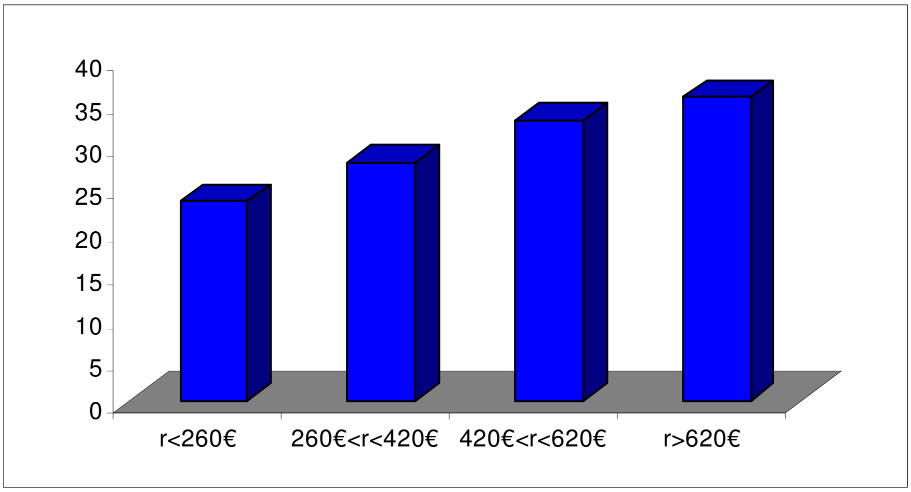


Figure 2. Household consumption repartition per income class (litres)



Legenda: *r* represents variable income.

Now we go into further details and consider the consumption repartition according to the colour of wine, to see whether some of the socio-demographic variables can affect the consumption choices. We have to point out that we cannot make comparisons among the consumptions along different classes of one single demographic variables, because we do not possess enough information to determine the pro capita average consumption. However it is possible, inside each class of the demographic variable to compare the consumption of the two typologies of wine, since in this case we observe the same group of households.

Figure 3 shows the analysis of the average wine consumption from the geographical point and we can observe that the white wine seems to be more appreciated in the North-East area, whereas in the North-West the red wine is more consumed. For the other two zones (Centre, South) the level of consumptions for the two kinds of wine is approximately the same.

Figure 3. Average wine consumptions per geographic area and typology of wine (quantity in litres)

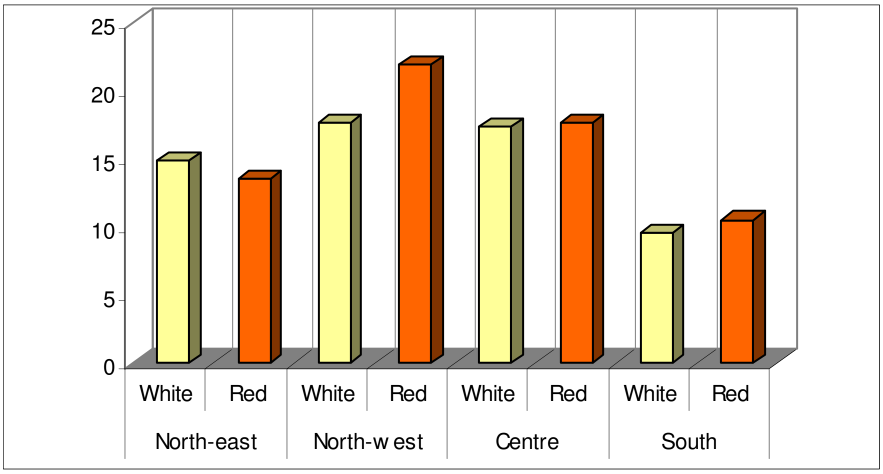
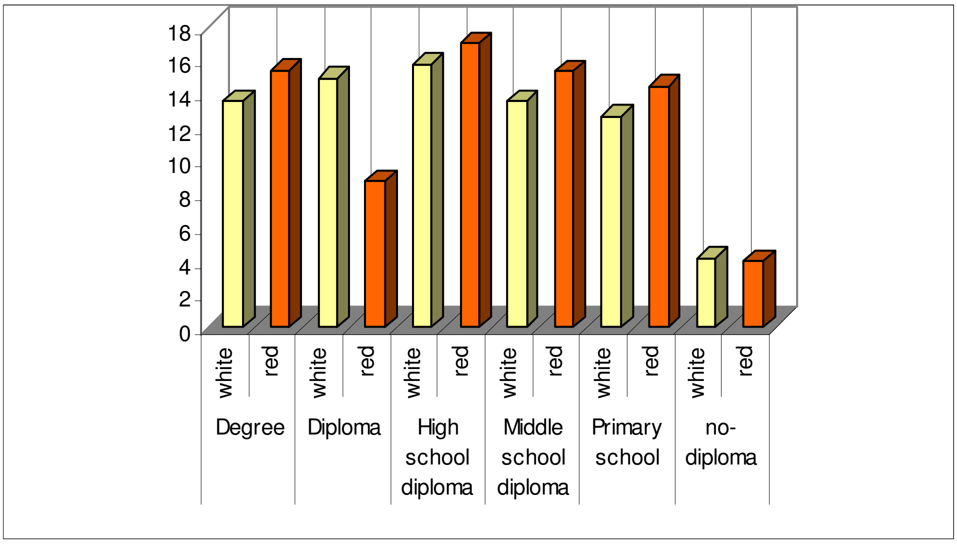


Figure 4. Average wine consumptions per level education and typology of wine (quantity in litres)



With reference to the education level of the R.P. (fig. 4) the only clear preference is the one exhibited by people belonging to the education level "high school diploma", whereas in the other classes the consumptions of the two typologies of wine are balanced.

3. Neural network basics

A very simple neural networks consists of four elements: a set of inputs (independent variables), a weight vector, a processing element/step and an output (dependent variable).
The processing element (referred to in the neural network literature as a "node") generate the output by weighting, summing and transforming the input variables. If a linear transformation function is used, this simple neural network model is identical to the familiar regression model. Usually, a tractable transformation function with the proper shape and bounds, as well as desirable analytical properties is the logistic function.

The major limitation of this simple network configuration is that the number of different input/output patterns that can be accurately modeled is severely limited. Moreover, the relationships between the inputs and the outputs is limited to smooth monotonic functions. To accommodate more complex relationships between inputs and outputs a more sophisticated network is needed.
By adding an intermediate layer of processing elements to the simple input-output model above, a 3-layer (input, intermediate or hidden, output) structure can model a large number of complicated input-output patterns.

Usually, the input variable and intermediate nodes are fully connected, that is every input variable is weighted and combined (summed and transformed) by every output node. While clearly violating the modeling principal of parsimony in model building, neural networks trade off elegance for the ability to model complex phenomena.

The weight vectors can be estimated in a variety of ways. The most common method to estimate these weights, is a learning algorithm called "back-propagation", which is, essentially a two-phase, gradient, descent optimization algorithm and the goal is to minimize the sum of squared errors

$$E = \sum_k (\mathbf{Q}_k - \hat{\mathbf{Q}}_k)^2 \qquad (1)$$

where $\mathbf{Q}_k$ is the actual output and $\hat{\mathbf{Q}}_k$ is the estimated by the network. The objective of the estimation process is to iteratively adjust the weight vectors in order to minimize *E*. And if the training is successful, one finds that *E,* given in (1), reduces.

In the training stage only a part of the data set is presented to the network and the holdback sample (that is the part that has not been used for training) is afterwards to check the performance of the neural network in reproducing the output. This validation part is necessary to be able to use the neural network model with any degree of confidence.

4. Elasticities

The usual summary measure of the effects of price and income (input) changes on quantity (output) is some form of elasticity. There are two forms of price elasticities to consider. The first is the self-elasticity which measures the effects of a variation of the price of the good *i* on the quantity of the same good. The second is the cross-elasticity which measures the effects of a variation of the price of the good *i* on the quantity of the good *j*.

The traditional methods for determining income and price elasticities involve estimating econometric models. The relationship between prices (income) and quantity is characterized by interactions, non linearities and asymmetric cross-effects, so it has been developed *i)* models which have very desirable properties to model also those characteristics as the Rotterdam, LES and AIDS models (Deaton and Muellbauer, 1980a, parametric approach) and *ii)* alternative techniques as ANN which are well suited to modelling complex relationships such as those found in wine market data. As we said in the introduction, in this work we follow the ANN non parametric approach and we decided to compute the elasticities with two different procedures.

The first identifies elasticities by observing the change in the individual choice probability for a change in the relevant system variable (Subba Rao *et al.*, 1998).
The following steps have been done:

1. run the calibrate neural network for the training sample by increasing the price (income) by different percentages;
2. consider the normalized output w_{ij} as equivalent to choice probability;
3. let w_{ij}^* be the output normalized for price *j* after P% change in the price of the good *j;*
4. use the following formula to compute a series of individual choice elasticity by increasing the price between 1% to 100%:

$$I\varepsilon_{i,j,k} = \frac{w_{ijk} - w_{ijk}^*}{w_{ijk}} \bigg/ \frac{P}{100} \qquad (2)$$

where *j* is the input generic variable, *i* is the generic hidden layer and *k* is the output;

5. calculate the probability-weighted average of these individual elasticities and obtain the aggregate demand elasticity.

This method is definitely very powerful. Firstly, it allows constructing the reaction of the consumers to different price variations for every family of the sample by computing the individual elasticity. Secondly, it makes possible to obtain the sample average demand curve by computing the weighted average of the individual elasticities for every increase of the price.
The second method to compute the elasticity is based on the usual definition of elasticity itself (Gruca and Klemz, 1998). To evaluate this effect using ANN we proceed as the following scheme:

1. set all the input variables, except P_j (price of j-th good) to their mean levels;

2. for every observed value of P_j, estimate $\hat{Q}_k$ using the trained neural network;
3. order the prices P_j (increasingly) and construct the demand curve
4. from these two ordered data series P_j's $\hat{Q}_k$'s, estimate the elasticity using the standard elasticity formula:

$$\varepsilon_{i,k} = \frac{\Delta \hat{Q}_k}{\Delta P_j} \frac{P_j}{\hat{Q}_k} \qquad (3)$$

5. Empirical evidence

In this section we present some results obtained by applying the artificial neural network to the data set described in section 2.

First of all, some preprocessing of the data is necessary before presenting the input pattern to the neural network. This usually involves scaling or normalizing the input patterns to value in the range [0,1]. In literature there are many procedures to normalize the data (Azoff, 1994), in our case we decided to divide the data by the maximum value.

In this work the architecture of the neural network consists of 1 input layer (made by 22 input variables),1 hidden layer and 1 output layer. Table 7 lists the 22 input variables that are presented to the network.

The transformation function we used in our case is the so called Pos-line function, a linear positive function taking values in the interval [0,1].

With regards to the learning phase to train the network, we followed the back- propagation scheme and the weights have been determined with a Levemberg Marquard algorithm, that seemed to be the fastest one.

To show the performance of the ANN we employed to analyse the wine demand we report a sequence of graphs (5-7) about the global errors, always bearing in mind that the global error is the difference between the real output and the one measured by the ANN.

In the training phase the error decreases with a rate that tends to be lower and lower until it becomes almost constant (fig. 5). The error variance during the training phase (fig. 6) is equal to 0.0077 (after 38 training epochs), whereas the error variance evaluated on the examples (fig. 7) is bigger (0.016) since in this case the ANN works on cases that has not seen before.

Before estimating the demand curve and the elasticities, we need to explain the role of the demographic variables on the consumption choices of the family and therefore we evaluate the so called relative influence of the variables. Indeed, ANN have been sometimes criticized for suffering from the lack of methods for the interpretation of the significance of input variables. Following Subba Rao approach, we employ a partitioning-of-weights algorithm (Garson, 1991) to interpret input parameters. The connection weights from the input layer to the hidden nodes and from the hidden to the output layer can be used to partition the relative share of the output prediction associated with each input variable. The method essentially involves partitioning the hidden output connection weights of each hidden neuron into components associated to the input neuron. The computation process to partition the weights follows an idea of Antony (1994).

Table 7. Input variables presented to the Artificial Neural Network

Variable	Name of the input variable	Description
sex	Gender	*0- Female* and *1- Male*
e1, e2, e3, e4	Age	Dummy: *up to 25years old* *from 25 to 44 years old* *from 45 to 60 years old* *over 60 years old*
High, medium, low	Education level	Dummy: *high* (degree, short degree diploma) *medium* (high school) *low* (middle school, primary school, no diploma)
E, O, C, S	Geographic area	Dummy : North West North East Centre South
Empl, Unempl, Other	Social position	Dummy: *employed* *unemployed* *other*
r1, r2, r3, r4	Economic status	Dummy: income less than 260 € income between 260 and 420 € income between 460 and 620 € income over 620 €
pb	$Price_i$	Price of the analyzed good (e. g. white wine)
pr	$Price_j$	Price of the other good (e.g. red wine)
expend	Income	Total wine expenditure is considered an approximation of the income.

Figure 5. Mean global error in the training phase

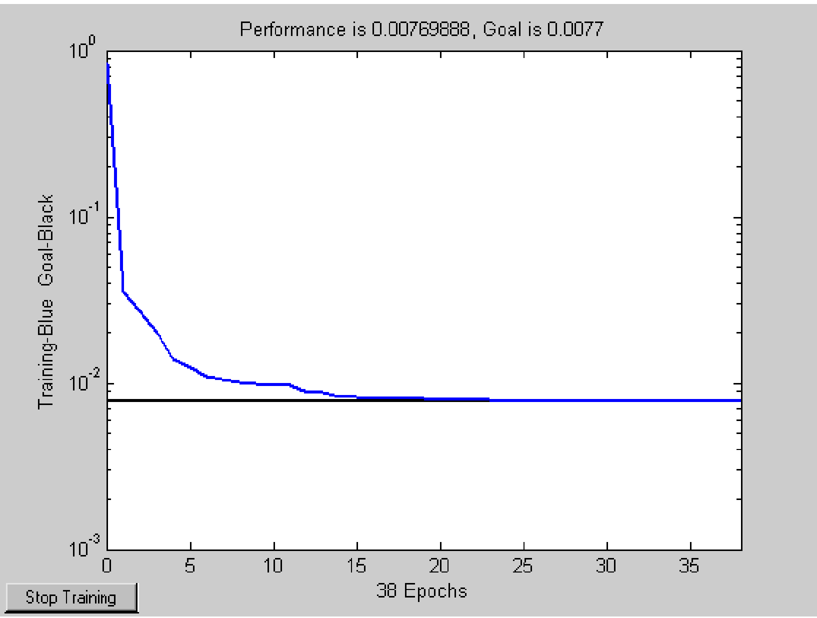
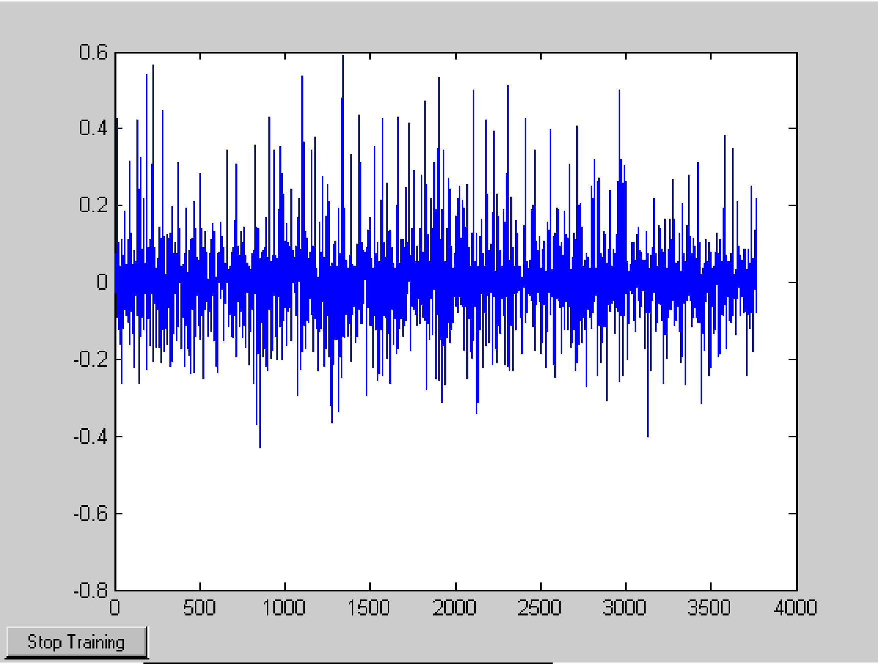
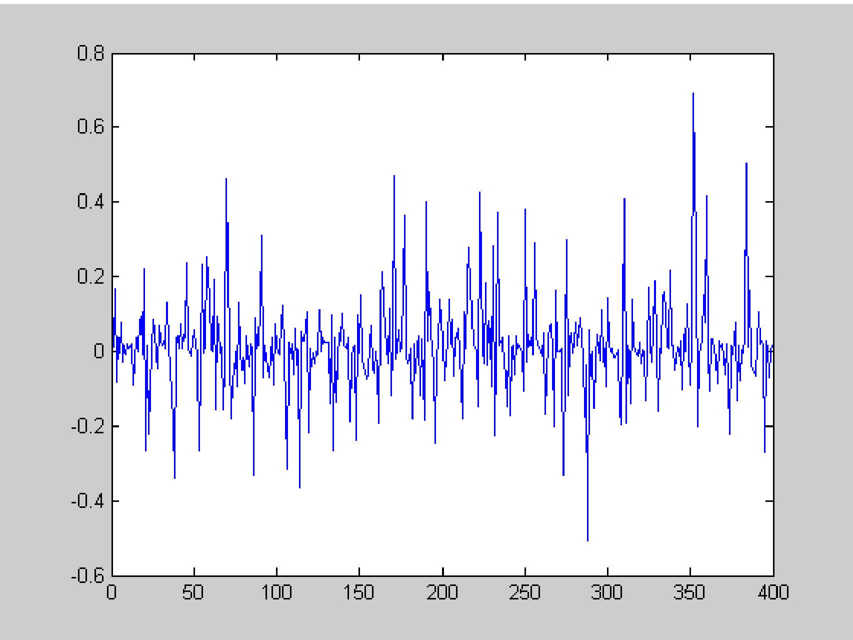


Figure 6. Error in the *training set*



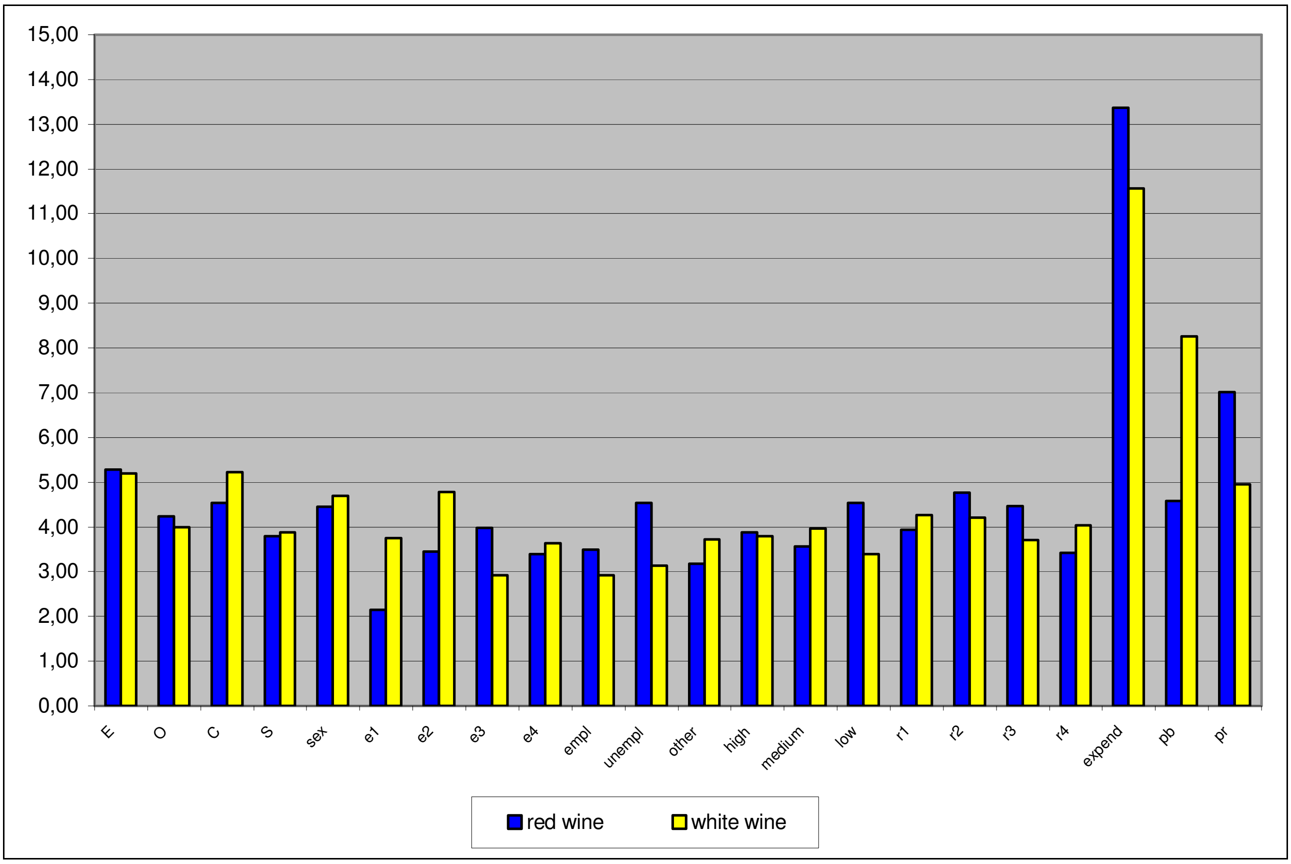
Var of the error_in	0.0077
Std of the error_in	0.0874
Mean of the error_in	0.0075

Figure 7. Error in the *validation set*



Var of the error_valid	0.0162
Std of the error_valid	0.1273
Mean of the error_valid	0.0116

Figure 8. Relative influence of the variables for white and red wine



Legenda: E: north-east; O: north-west; C: center; S: south; e1: up to 25 years old; e2: 25 to 44; e3: 46-60; e4: over 60; r1: less than 260 €; r2: 260-420 €; r3: 460-620 €; r4: over 620 €; pb: white wine price; pr: red wine price.

In fig. 8 the relative influence of the variables for our specific case are reported. Firstly, we can notice that variable expenditure (income) is definitely the most important one for both the typologies of wine, slightly more for the red wine (13.4%) than for the white wine (11.6%). The price is also very relevant, more in details for each kind of wine (for instance white wine) both the price of the white wine itself (pb) and the price of the red wine (pr) are influent.

Although, as expected, the traditional basic variables of the economic theory (income and prices) are still the most important ones, some demographic variables manage to be sometimes more important, often at least as influent as the traditional economic factors. For instance, variable geographical repartition is very important, especially North East (E) and Centre (C), both for white and red wine. This is an interesting confirm that the demographic variables need to taken into account to understand the consumers' behaviours.

These results confirm that all the variables considered as input variables are relevant in the consumption choices of the family and a further confirm comes from the elasticity analysis that follows.

Using the first method described in section 4 (Subba Rao *et al*, 1998) new output have been produced with the trained ANN by augmenting the price of increasing percentages, keeping constant all the other input variables. In this way, the demand curve of the single families has been determined. In the following fig. 9, 10 we report the demand curve obtained, for every price increment, as average of the estimated individual curves.

Once the demand curve is known, it is interesting for the economic operator to evaluate how the consumptions vary with the price. To this aim, the elasticity is quite a strong economic tool.
As we pointed out in section 4, the elasticity is a measure of the sensitivity of one variable when another variable varies and, in particular, considered in absolute value, it is the percentage variation of the quantity with regards to percentages variations of price.

Figure 9. White wine demand curve

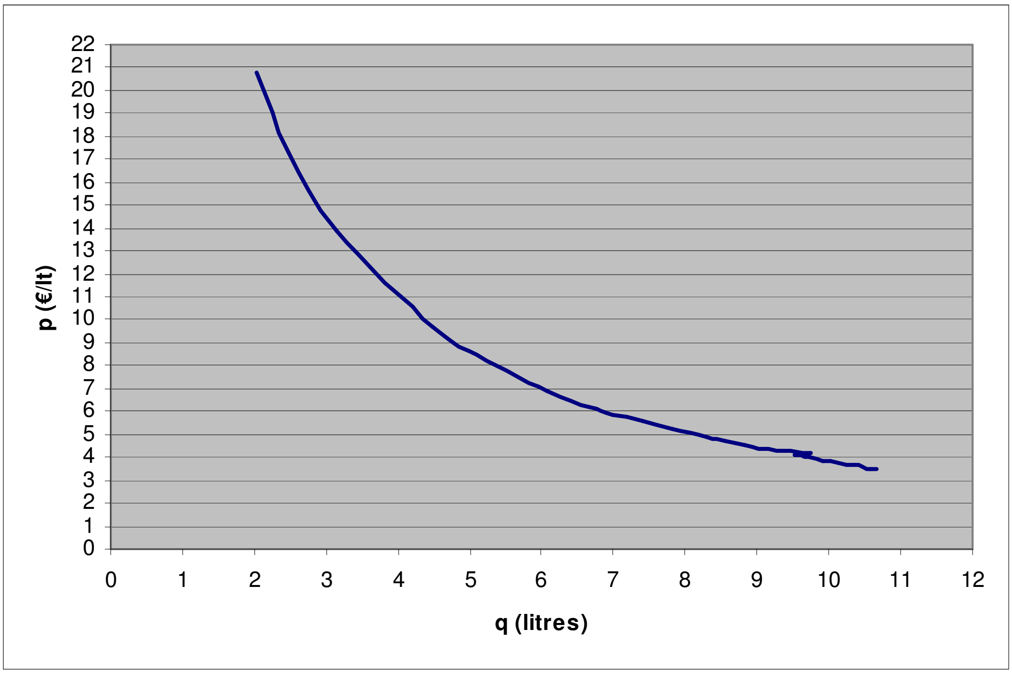
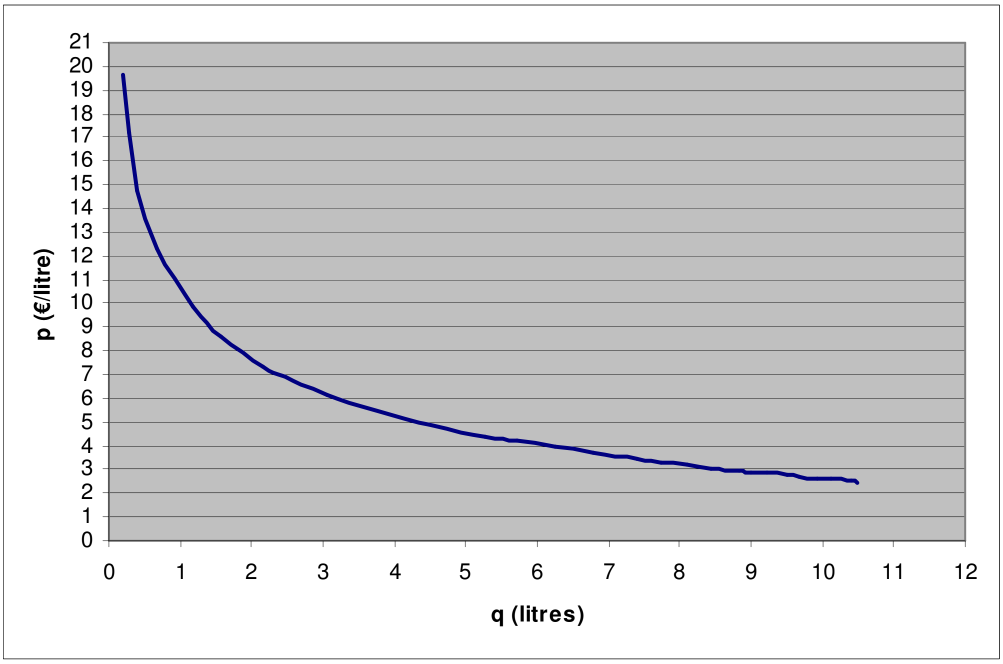


Figure 10. Red wine demand curve



In the following table 8 we report the weighted average elasticities independently of the socio-demographic characteristics of the family.

Table 8. Weighted average price elasticities per typology of wine

	Price Elasticity	
	self elasticity	cross elasticity
White wine	-0.67	0.66
Red wine	-0.13	0.85

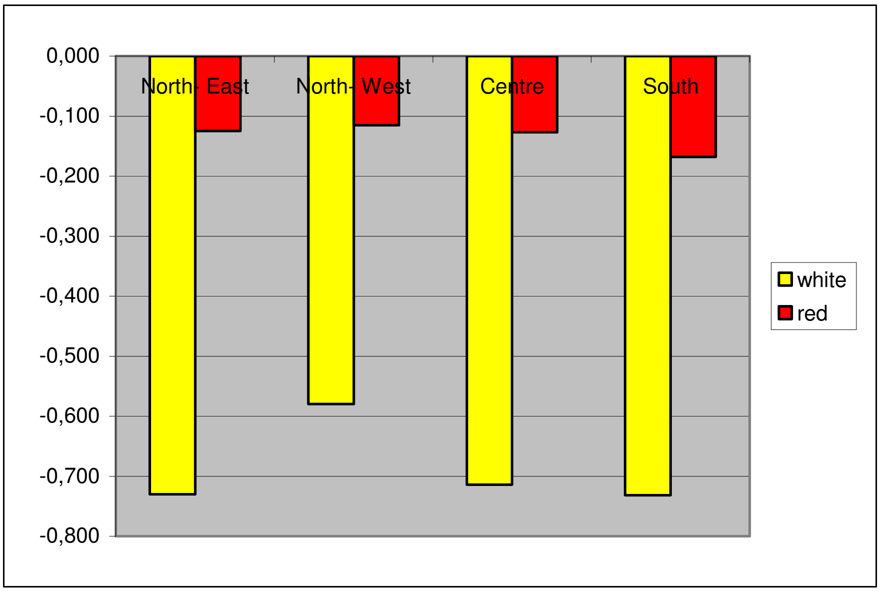
Considering the values we obtained (all smaller than 1)[1], the demand can be supposed to be basically an-elastic.

Then we computed the elasticities for single demographic variables with the first method described in section 4 (Subba Rao *et al.*, 1998) and we report in fig. 11-14 the results we obtained.

Among the most interesting results we can observe the different behaviour of the consumers of area North-West from the other areas and also consistently with previous empirical works variable gender clearly affects the elasticity. Indeed, if the reference person is female, it is bigger the reduction in the purchased quantity in case of increase of the price.

As far as variable age, it seems that consumers in the last class of age tend to be less sensitive to variations in the price of red wine, then consumers of different age class. Finally we briefly comment variable income, where as expected, the lowest income classes are more sensitive to price variations than the highest.

Figure 11. Self elasticities per geographic areas



[1] Before commenting the results, we need to point out that the probability distribution of the elasticities is not known, therefore nothing can be said about the significativity of the values

Figure 12. Self elasticities per gender of the R.P.

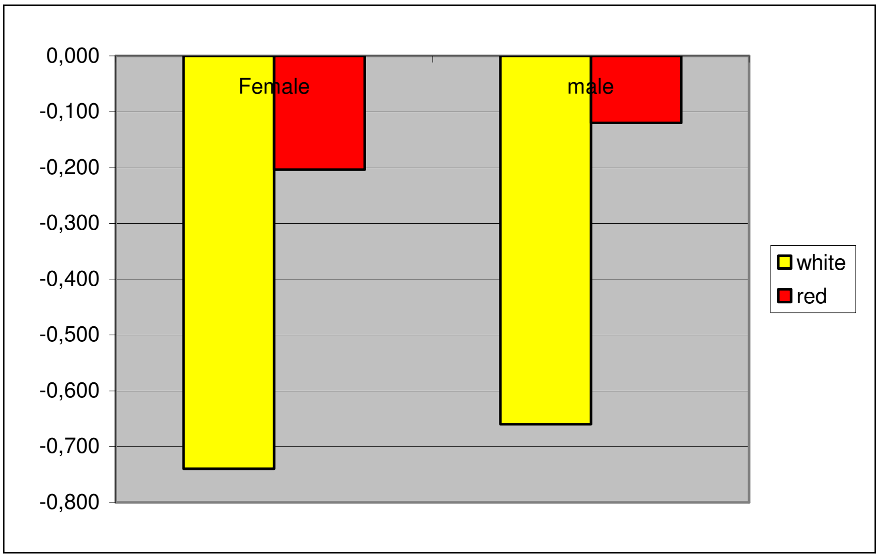


Figure 13. Self elasticities per social position of the R.P.

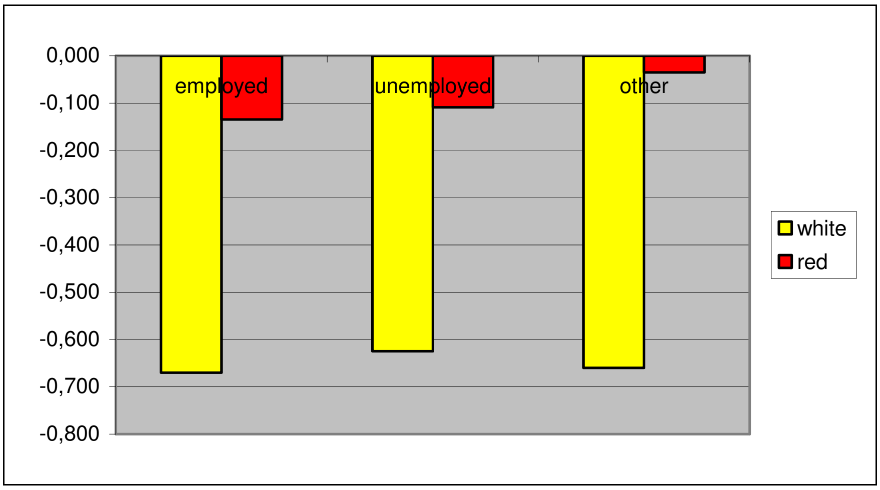
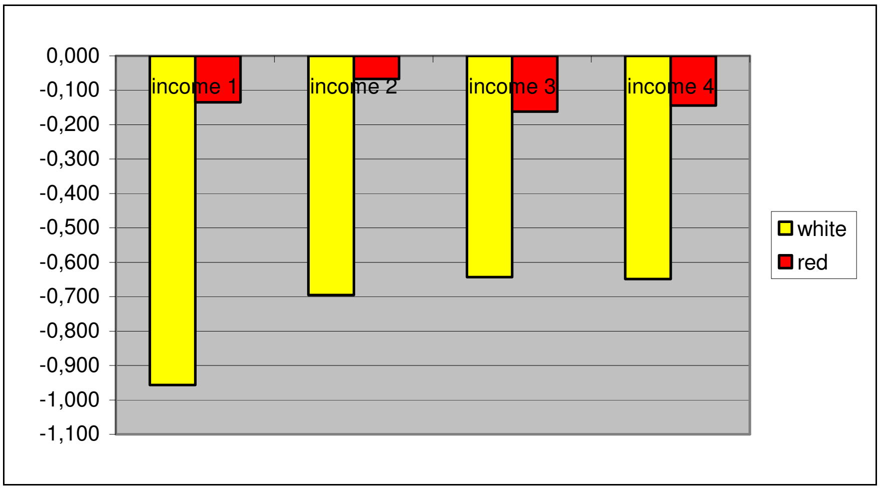


Figure 14. Self elasticities per p.c. income



Then we evaluated the joint effect of some demographic variables on the elasticities with both the first method (Subba Rao) and the second method (Gruca) in order to make a comparison of the performance of the two methods. In particular to do that we identify two kinds of family, family I (high education level, third class of age, high income) and family II (low education level, first class of age, low income).

Table 9. Self elasticities (white wine) per family typology

Family typology	Gruca Method	Subba Rao Method
Family I	-1.18	-0.67
Family II	-3.21	-0.97

Observing the table we notice that the values are different but the reaction of the consumers are in the same direction. This difference can depend on a somehow major robustness of the Subba Rao method in presence of cross sectional data, with many source of heterogeneity as in our case. For this reason we think that Subba Rao method is probably more reliable.

6. Conclusions

The main objective of this paper is to model the Italian wine demand market using the non parametric approach of the Articificial Neural Network, instead of the usual parametric techniques (AIDS, Rotterdam model). The use of the ANN to model demand market is quite an innovative topic, especially in food and beverage sector.

The results we obtained evidence an interesting capability of the ANN to generalize from the examples that have been presented in the training phase to new cases. This means that the network is able to give results also for cases (families, in our context) that have not been presented before. This flexibility of the ANN together with its relative easy implementation, once the architecture has been built, makes their use very enticing.

However, there are also some downsides, mainly because the ANN are a very recent methodological instrument and not enough investigation has been done about it. First of all, there is no precise method to optimally choose the network among all the possible architectures. Also, by now it not possible to build the network taking into account the economic theory assumptions and especially this second topic is among our future research lines.

In summary, in this work we used successfully the ANN method to estimate the demand curve and the elasticities for wine market in Italy that have been computed with two procedures: the Subba Rao method and the Gruca method. Even though to our knowledge there are no results about microdata, our numerical results are consistent with those obtained in some applied works on time series (Moro and Canali, 1996).

From the study it emerges the relevance of some socio-demographic variables that influence the consumers choices together with the traditional economic variables (income and price) this is an important result for the segmentation of the market and definition of the target. Moreover, it is interesting to observe that the red wine demand curve is more rigid than that of white wine. This could be probably justified by the influence of historical-cultural factors on the consumption choices.

Here some other interesting question marks arise and already are in our works in progress. Mainly, our interest is oriented towards the recontruction of the empirical distribution of the elasticites, since the theorical one is not known and therefore nothing can be said about the significativity of the values.

References

Azoff, E.M. (1994). Neural Network time series forecasting of financial market. Wiley Financial Edition.

Deaton, A., Muellbauer, J. (1980a). An Almost Ideal Demand System. American Economic Review. Vol. 70, n.3, pp 312-326.

Giordano F., La Rocca M. (2002). Una introduzione alle reti neurali artificiali di tipo feed-forward. Summer School of the Italian Statistical Society "Modelli lineari e non lineari e algoritmi genetici per l'analisi e la previsione nell'ambito delle serie temporali, economiche, demografiche ed ambientali", unpublished.

Heckman, J.J. (1979). Sample selection Bias as a Specification Error. Econometria. Vol. 47, n.1, pp. 153-161.

Heien, D., Durham, C. (1991). A test for the habit formation hypothesis using households data. Review of Economics and Statistics. Vol. 73, pp. 189-199.

Heien, D., Wessels, C.R. (1990). Demand Systems Estimation with micro-data: a censored regression approach. Journal of Business and Economic Statistics, Vol. 8, n. 3, pp. 365-371.

Mauracher C. (2005). Le variabili sociodemografiche nel consumo di vino: un'analisi mediante Reti Neurali Artificiali. Ph.D. thesis, Padova University.

Subba Rao, P.V., Sikdar P.K., Krishna Rao K.V., Dhiringra S.L. (1998). Another insight into artificial neural networks through behavioural analysis of access mode choice. Computational Environmental and Urban Systems. Vol. 22, pp. 485-496.

Gruca, T.S., Klemz B.R. (1998). Using neural networks to identify competitive market structures from aggregate market reponse data. Omega International Journal of Management Science, Vol. 26, No. 1, pp. 49-62.