Доклади на Българската академия на науките
Comptes rendus de l'Académie bulgare des Sciences

Tome 66, No 3, 2013

SCIENCES ET INGENIERIE

Automatique et informatique

ECONOMETRIC VS ARTMAP PREDICTION OF ECONOMIC CHOICE

George Mengov, Nikolay Georgiev*

(*Submitted by Corresponding Member M. Hadjiski on November 26, 2012*)

Abstract

Forecasting economic behaviour is an important problem with practical implications for a number of scientific disciplines, including microeconomics, macroeconomics, marketing and economic psychology. The ability to predict the economic agent's choice is a coveted goal for both social scientists and market practitioners. In our time, such studies are conducted with field investigations or laboratory experiments. However, the traditional statistical techniques used to build explanatory models with predictive power are of limited capability and have inherent structural deficiencies. Here we show that an artificial neural network of the ARTMAP family forecasts far better than the state-of-the-art multinomial regression the economic decisions of the participants in a laboratory experiment resembling real markets. We found that when the number of options among which one must choose is four, and hence any systematic predictive success above 25% is valuable, Fuzzy ARTMAP achieved 42.28%, while the most popular logit regression model reached 37.87%. This result demonstrates the greater capability of the neural classifier to utilize correlated input factors, which remain underused by regression analysis. Yet, prediction rates such as the attained here are still very low, and could hardly be raised by more sophisticated statistical techniques, but should rather be improved by incorporating more in-depth psychological knowledge about the decision-maker.

Key words: multinomial logit model, logistic regression, ARTMAP neural network, economic choice, econometrics, prediction

Introduction. Finding accurate methods for predicting economic choice is a topic with implications for micro- and macroeconomics, marketing and other related fields. Currently, multinomial logit, probit and tobit models are the statistical tools most widely used to explain and predict qualitative response variables, and hence agents' choices [1]. However, all these methods have limitations, stemming from the assumptions upon which they rest. A related issue is the rapid decrease of predictive efficiency with the increase of categories to be chosen above two or three. In addition, in that case the number of model parameters grows exponentially and some of them may not be statistically significant and may bias the predictive accuracy, yet they cannot be discarded from models with even the minimal structure.

On the other hand, some relatively recent statistical classification techniques can offer much desired flexibility to econometrics, especially with regard to choice prediction. Such is the large group of artificial neural networks, including those from the ARTMAP family [2–4]. The latter stand somewhat apart from the mainstream of the field, but continue to attract attention from researchers in need of robust classification methods.

Here we present an economic experiment investigating the role of consumer satisfaction in choosing a supplier of a fictitious good. Our objective was to identify the limits of predictability of human decisions by the multinomial logit model and the Fuzzy ARTMAP neural classifier – two algorithms, representative each in its class. We could have used a more recent ARTMAP variant, but decided against it: the relatively small number of empirical observations at our disposal would not have allowed for minor algorithmic differences to make an impact.

The participants in the experiment had to choose one among four suppliers in twenty rounds of a utility-maximizing game. Thus the two classification methods competed in surpassing a prior probability of 0.25. In the next sections, we provide details on the experimental design and administration, and then discuss the classifying algorithms' performance.

The main goal of our experiment was to investigate the complex relationship between consumer satisfaction or dissatisfaction with a supplier of a good, and the decision to stay with or abandon that supplier in the future. A number of factors could play a role here. Apparently, the emotion of (dis)satisfaction – as provoked on the spot – is a powerful motivator of economic choice; in addition, it may be blended with past experiences with the same business partner, or influenced by attractive offers from other competitors. These factors formed the backbone of our hypothesis, later tested with the two algorithms.

Experiment. The experimental design accounted well for the factors described above. It put the participant in a situation to choose one offer among four and then bear the consequences, which were either unfulfilled promises on behalf of the suppliers, or exceeded expectations in terms of delivered good. That good was called omnium bonum ("good for everybody", in Latin). It had to be fictitious to avoid mental associations with real goods or services that could skew each participant's motivation. No transaction costs were involved in abandoning one supplier for another.

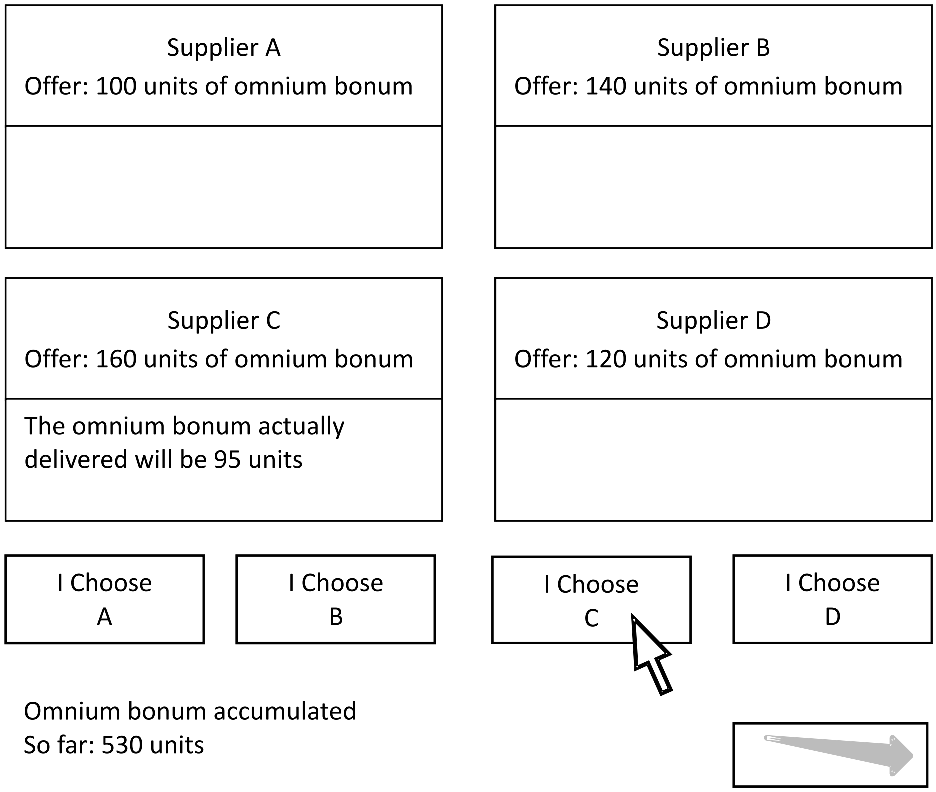

Fig. 1. Experimental screen of the software application. Initially, four offers were available and then the participant chose Supplier C with a mouse click. The actual quantity was delivered and then added to the total quantity accumulated so far

Figure 1 shows the computer screen in front of the participant. Initially, only the offers are on display, but immediately after a choice with a mouse click, the actual 'delivery' takes place. The new omnium bonum units are added to the total sum in the bottom-left corner, and then the arrow button must be clicked to proceed. A new screen (not shown here) appears and asks for the participant's self-assessed disappointment or satisfaction with the transaction in a Lickert scale [5]. After an answer, the round is over and the next one begins.

As it is shown in Figure 2, each supplier differs from the other three by the units of omnium bonum it offers and delivers, whereby the riskiest one (C) is also the most rewarding. All quantities are chosen so that all four suppliers remain competitive, and each of them is likely to form a distinct image in the eyes of the participant. The experiment was conducted in May 2010 with 34 students from the Faculty (School) of Economics and Business Administration at Sofia University, who gained additional credit points for a decision making-course. More details about the experimental design, setting and procedure are provided in [6].

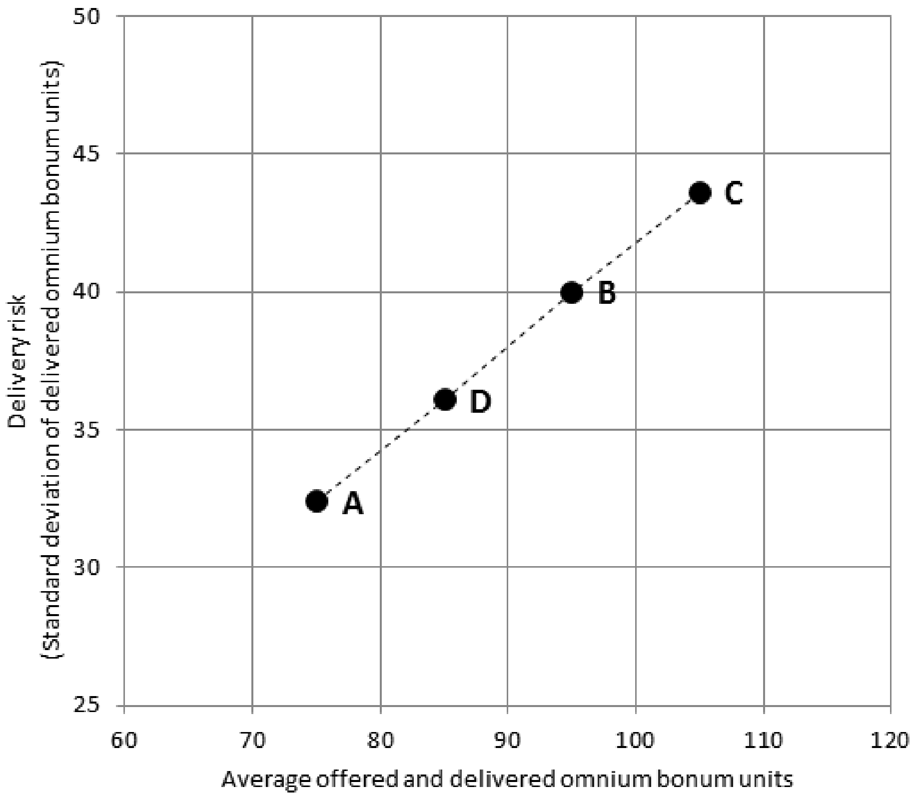

Fig. 2. An essential idea of the experimental design: the more omnium bonum is offered (and delivered) on average, the greater the risk of delivery, as measured by the standard deviation

Results and discussion. We applied multinomial logistic regression (multinomial logit models) to try to understand and predict our participants' choices. Initially, a set of potentially significant independent factors was identified based on general economic and psychological considerations. Then they were tested one at a time as possible predictors in logit models. Eventually, two candidate variables emerged for a final selection. Both involved analytical transformations of the variable DS (disappointment/satisfaction with the acquisition of omnium bonum in the current round), and both achieved 46.47% rate of success in predicting people's choices in a calibration sample (to be explained below). However, the two variables were highly correlated and therefore could not be a part of the same model. We solved the problem by applying the standard econometric technique to form a compound variable by summing them up. Thus, the right-hand side of the final logit model contained a single factor – a sum of two quantities, and was this

$$DS_s(\max(t(s=c))^{1,t-1}) + \text{avg}(DS_s)^{1,t-1}|\text{avg}(DS_s)^{1,t-1}|. \quad (1)$$

Table 1

Prediction rates of supplier choices

Prediction method	Number of independent variables	Number of estimated parameters	Prediction rate: Calibration sample: first 12 rounds ($n = 340$)	Prediction rate: Test sample: last 8 rounds ($n = 272$)
No method (Pure guessing)	n. a.	n. a.	0.25[a]	0.25[a]
Multinomial Logit Model – Augmented	26	81[b]	0.5618	0.3713
Multinomial Logit Model – Minimal	4	15[c]	0.4794	0.3787
Fuzzy ARTMAP neural network	36	$\sim 10^2$–10^3	0.9971	0.4228

[a]Theoretical estimate;

[b]Only 29 are significant at the 0.05 level;

[c]Only 7 are significant at the 0.05 level.

Here the subscript $s = A$, B, C, D designates the supplier of omnium bonum. Note that eq. (1) is in fact a vector of four elements because there are four suppliers. That model achieved 47.94% rate of success in predicting people's choices in a calibration sample – slightly better than the two that gave rise to it.

The expression in eq. (1) is not straightforward to interpret, but an intuitive understanding is still possible. The quantity $DS_s(\max(t(s = c))^{1,t-1})$ is the satisfaction (or disappointment) experienced by the participant the last time supplier s was chosen ($s = c$), no matter how far in the previous rounds this had happened. Notation $\max(t(\cdot))$ stands for 'most recent' dealing with s. Superscript $1, t-1$ indicates that the game is at moment t and all preceding choices from the first to the penultimate round are considered. In other words, the term accounts for the fact that people do remember how they felt about supplier s the last time they made a deal with it.

The second term in eq. (1) is the average DS associated with the s-th supplier thus far in the game, multiplied by its absolute value. As an effect, this operation amounts to taking the square of the variable with preserving its algebraic sign, and therefore can account for both disappointment and satisfaction, whichever is relevant at the moment. Essentially, the two terms in eq. (1) tell us about the most recent and the average emotion a supplier has induced in a customer. This combination looks plausible in general, and relates well with Redelmeier

T a b l e 2

Variables and parameters in Multinomial Logit Model – Minimal

Variable	Supplier	Regression coefficient	p-value
Intercept A	A	−0.0268	0.9041
$DS_A(\max(t(A))^{1,t-1}) + \text{avg}(DS_A)^{1,t-1}\lvert\text{avg}(DS_A)^{1,t-1}\rvert$		0.0722	0.0274
$DS_B(\max(t(B))^{1,t-1}) + \text{avg}(DS_B)^{1,t-1}\lvert\text{avg}(DS_B)^{1,t-1}\rvert$		0.0162	0.6189
$DS_C(\max(t(C))^{1,t-1}) + \text{avg}(DS_C)^{1,t-1}\lvert\text{avg}(DS_C)^{1,t-1}\rvert$		0.0080	0.7361
$DS_D(\max(t(D))^{1,t-1}) + \text{avg}(DS_D)^{1,t-1}\lvert\text{avg}(DS_D)^{1,t-1}\rvert$		−0.0752	0.0064
Intercept B	B	0.4412	0.0286
$DS_A(\max(t(A))^{1,t-1}) + \text{avg}(DS_A)^{1,t-1}\lvert\text{avg}(DS_A)^{1,t-1}\rvert$		0.0045	0.8830
$DS_B(\max(t(B))^{1,t-1}) + \text{avg}(DS_B)^{1,t-1}\lvert\text{avg}(DS_B)^{1,t-1}\rvert$		0.0833	0.0019
$DS_C(\max(t(C))^{1,t-1}) + \text{avg}(DS_C)^{1,t-1}\lvert\text{avg}(DS_C)^{1,t-1}\rvert$		0.0033	0.8710
$DS_D(\max(t(D))^{1,t-1}) + \text{avg}(DS_D)^{1,t-1}\lvert\text{avg}(DS_D)^{1,t-1}\rvert$		−0.0338	0.1744
Intercept C	C	0.9879	0.0000
$DS_A(\max(t(A))^{1,t-1}) + \text{avg}(DS_A)^{1,t-1}\lvert\text{avg}(DS_A)^{1,t-1}\rvert$		0.0089	0.7510
$DS_B(\max(t(B))^{1,t-1}) + \text{avg}(DS_B)^{1,t-1}\lvert\text{avg}(DS_B)^{1,t-1}\rvert$		−0.0046	0.8635
$DS_C(\max(t(C))^{1,t-1}) + \text{avg}(DS_C)^{1,t-1}\lvert\text{avg}(DS_C)^{1,t-1}\rvert$		0.0563	0.0061
$DS_D(\max(t(D))^{1,t-1}) + \text{avg}(DS_D)^{1,t-1}\lvert\text{avg}(DS_D)^{1,t-1}\rvert$		−0.0528	0.0249

and KAHNEMAN's discovery [7] that humans remember easily their last emotion in an experience, and also form a relatively good average estimate of its intensity. Finally, the use of eq. (1) in the actual econometric model implies that the greater the satisfaction, the greater the probability for a choice to be repeated is. Other, more 'economic' variables, such as the difference between promised and delivered omnium bonum were highly correlated with DS, and so could not be used alongside the quantities from eq. (1).

We have put the compound variable from eq. (1) in two different models – one model with it as a single predictor, and another including additional independent variables possibly adding explanatory power. Both models were calibrated with a subsample of the participants' responses from the first 12 rounds (calibration sample), and then tested with the last eight rounds' data (test sample). As it is shown in Table 1, the two models performed similarly in testing, with the more parsimonious (Multinomial Logit Model – Minimal) predicting slightly better (37.87%). Obviously the large number of insignificant coefficients in the 'Augmented' model was the price it paid for better accounting for the calibration sample.

Coming back to the 'Minimal' model, it is noteworthy that although it has the smallest possible number of predictive variables – just one, as given by eq. (1), it must spawn and multiply to account for the total number of choice categories in the experiment (Table 2, Column 1). In particular, with $m = 4$ options to be chosen, the minimal number of independent variables is four. The latter must enter $m(m-1)$ equations and, if intercepts are included, there are m^2-1 parameters to be estimated. That is how, with a choice of one among four, a logit model with a single predictive variable implies 15 parameters.

It turns out (Table 2) that about half of the parameters in the most parsimonious model are statistically insignificant but there is nothing to do about it. This example reveals an important deficiency of the econometric technique – the minimal structure of the equation is fixed, and if there are statistically insignificant parameters in it, they cannot be discarded. Adding further factors causes an outburst in their number, as the 'Augmented' variant in Table 1 shows.

On the other hand, neural networks in general and ARTMAP in particular contain far more parameters than most regression models, making the goal for statistical significance impossible and irrelevant. The network algorithm forms clusters of the input patterns as they arrive and links them to the correct output categories. Depending on the ART parameter values, dozens or hundreds of input clusters may be created to develop a refined reflection of the multidimensional input space. In the training process, they get connected to the output clusters which in our case represent the four suppliers.

As it is shown in Table 1, the neural network precision, as given by the rate of correct prediction in the calibration sample, approaches 100%. The real test of course is with a sufficiently large test sample. By that standard, neural models have historically fared better than their more rigid statistical 'cousins', and the present example is no exception. As it can be seen from Table 1, the neural network recognized properly 339 of all 340 records in the calibration sample, and then predicted correctly 42.28% of the 272 choices in the test sample.

Several circumstances have contributed to the Fuzzy ARTMAP advantage of about four and a half percentage points over the logit model (Table 1). First, there is no structural limitation for the number of independent variables – as no statistical significance is pursued, they can be many. We achieved our best result with 36, but variants with other sets performed similarly. Second and related, collinearity is not an issue as neural networks have shown to be tolerant to correlating factors influencing the dependent variable. Finally, the algorithm's flexibility allows for finding the optimal quality of input pattern recognition – neither too coarse, nor too refined.

Conclusions. We have shown that a modern classification tool like Fuzzy ARTMAP can raise the accuracy of choice prediction with four options from 0.25 to 0.42. Perhaps this is the limit of statistical classification and no further model sophistication would lead to improvement. But, to see our result in the opposite perspective, more than half of all guesses were still wrong. This shows that indeed, the task of forecasting people's decisions must be approached with due modesty. Maybe the right way forward would be to implement a more theoretical approach, and to account for more of the psychological mechanisms behind choice.

REFERENCES

[1] Gujarati D. Basic Econometrics, London, McGraw-Hill, 2011.
[2] Carpenter G., S. Grossberg, N. Markuzon, J. Reynolds, D. Rozen. IEEE Transactions on Neural Networks, **3**, 1992, No 5, 698–713.
[3] Carpenter G., S. Grossberg, J. Reynolds. Neural Networks, **4**, 1991, No 5, 565–588.
[4] Carpenter G. CAS/CNS Technical Report TR-2003-008, 2003.
[5] Mengov G., H. Egbert, S. Pulov, K. Georgiev. Neural Networks, **21**, 2008, No 9, 1213–1219.
[6] Georgiev N., G. Mengov. Annuaire Univ. Sofia, Fac. Econ. Gest., **10**, 2012 (in print, in Bulgarian).
[7] Redelmeier E. A., D. Kahneman. Pain, **66**, 1996, No 1, 3–8.

Department of Statistics and Econometrics
Faculty of Economics and Business Administration
St. Kliment Ohridski University of Sofia
125, Tsarigradsko chaussée Blvd, Bl. 3
1113 Sofia, Bulgaria
e-mail: g.mengov@feb.uni-sofia.bg

**BI Norwegian Business School*
Nydalsveien 37
0484 Oslo, Norway
e-mail: Nikolay.Georgiev@bi.no