ARGUMENTA OECONOMICA
No 3 · 1996
PL ISSN 1233-5835

Józef Dziechciarz, Marek Walesiak***

MODELLING THE COLLEGE STUDENT CHOICE PROCESS VIA CONJOINT ANALYSIS

In the article, the conjoint analysis method was applied to the situation in which School Management intends to start new, post high-school form of teaching opportunities. Each new school is described by five variables: duration, work intensity, organization, vocational diploma, and price per month. In this inquiry a sample of 165 high school students from Wrocław were asked to express their preference. The respondents were asked to evaluate the proposed schools by evaluating them on a 100-point scale. In order to estimate the part-worths, and relative importance of each characteristic in the choice process, collected data are analysed with conjoint measurement methodology.

1. INTRODUCTION

For consumer preference evaluation in marketing applications conjoint measurement is used. (The theoretical background for the conjoint analysis method is given in: Walesiak 1996, Dziechciarz and Walesiak 1995, Louviere 1988). As a base for that a set of products described by the vector of its characteristics' values is used. The conjoint impact of two or more product characteristics measured on the nominal scale (independent variable) on the dependent variable with the values measured on ordinal, interval or ratio scale is determined. The selection of measurement scale for the dependent variable determines which method of parameter estimation can be used (see for details: Vriens and Wittink 1994). The most common distinction is between metric and nonmetric estimation procedures. Metric procedures (e.g. ordinary least squares regression with dummy variables) can be applied to interval- or ratio-scaled measurement of dependent variable. Nonmetric estimation methods (e.g. monotonic analysis of variance) require ordinal data.

In order to estimate the part-worths (the term part-worths (utilities) is defined in section 2), and relative importance of each characteristic in the choice process, collected data are analysed with the conjoint measurement methodology. As the result of the analysis (Hair et al 1995; Anttila et al 1980):

1) the relative importance of each attribute to the overall evaluation of the object is shown,

* Department of Econometrics, Wrocław University of Economics.
** Department of Econometrics and Computer Science, Wrocław University of Economics.

2) the relative contribution of each attribute level to the overall evaluation of the object is shown,

3) the predicted market shares for the objects with different sets of features are given.

2. PROBLEM DESCRIPTION

School management intends to start new, post high-school form of teaching opportunities. Each new school is described by five variables:

Z_1 – Duration:	Z_2 – Work intensity:	Z_3 – Organization:
– one year,	– 20 hours per week,	– full time,
– two years.	– 30 hours per week.	– evening courses.

Z_4 – Vocational diploma:	Z_5 – Price per month:
– yes,	– 250 zł,
– no.	– 200 zł,
	– 150 zł.

In this inquiry a sample of 165 high school students from Wrocław were asked to express their preference. The data was collected in autumn 1995. Cattin and Wittink (1982) report that the sample size in commercial conjoint studies usually ranges from 100 to 1,000.

Instead of asking for the evaluation of all possible combinations of characteristics' values (i.e. 48 possibilities: $2 \times 2 \times 2 \times 2 \times 3$), the orthogonal array of 10 "schools" was constructed (Figure 1).

The respondents were asked to evaluate proposed schools by evaluating them on a 100-point scale (Table 1 shows selected respondent's rating of school descriptions).

3. METHODOLOGY

Ordinary least squares regression with dummy variables was used for the estimation of part-worths separately for each respondent. In regression analysis dependent variable is a value attached to each school (each product) by the respondent. In order to enable the measurement of the relative importance of each characteristic's value, dummy variables reflecting respondents' evaluation of given level of the independent variable are introduced into the model. Any nominal variable with k categories can be represented as $k-1$ dummy variables. In our example we should use six dummy variables in regression analysis.

Card 1	
Duration	one year
Work intensity (per week)	30 hours
Organization	full time
Vocational diploma	no
Price per month	150 zł
Preference	

Card 2	
Duration	two years
Work intensity (per week)	20 hours
Organization	full time
Vocational diploma	no
Price per month	200 zł
Preference	

Card 3	
Duration	one year
Work intensity (per week)	30 hours
Organization	evening courses
Vocational diploma	yes
Price per month	200 zł
Preference	

Card 4	
Duration	two years
Work intensity (per week)	30 hours
Organization	evening courses
Vocational diploma	no
Price per month	150 zł
Preference	

Card 5	
Duration	two years
Work intensity (per week)	20 hours
Organization	evening courses
Vocational diploma	yes
Price per month	150 zł
Preference	

Card 6	
Duration	one year
Work intensity (per week)	20 hours
Organization	evening courses
Vocational diploma	no
Price per month	250 zł
Preference	

Card 7	
Duration	two years
Work intensity (per week)	30 hours
Organization	full time
Vocational diploma	yes
Price per month	250 zł
Preference	

Card 8	
Duration	one year
Work intensity (per week)	20 hours
Organization	full time
Vocational diploma	yes
Price per month	150 zł
Preference	

Card 9	
Duration	two years
Work intensity (per week)	30 hours
Organization	evening courses
Vocational diploma	yes
Price per month	200 zł
Preference	

Card 10	
Duration	two years
Work intensity (per week)	30 hours
Organization	full time
Vocational diploma	yes
Price per month	200 zł
Preference	

Fig. 1. Cards – 10 school descriptions

Table 1
Respondent's rating of school descriptions for conjoint analysis (selected respondents)

Respondent	Number of school variant									
	1	2	3	4	5	6	7	8	9	10
001	40	30	99	20	90	10	70	60	80	50
002	10	01	90	01	10	01	5	99	45	30
003	65	20	60	60	50	15	10	70	45	30
–	–	–	–	–	–	–	–	–	–	–
096	50	10	60	30	80	40	99	20	70	90
097	05	01	85	01	45	03	60	80	50	75
–	–	–	–	–	–	–	–	–	–	–
163	70	70	95	60	90	50	99	99	90	99
164	30	10	99	50	20	40	80	70	60	90
165	60	80	40	20	30	10	30	70	50	99

Multiple regression model with six dummy variables for s-th respondent is following:

$$\hat{Y}_s = b_{0s}+b_{1s}X_{1s}+b_{2s}X_{2s}+b_{3s}X_{3s}+b_{4s}X_{4s}+b_{5s}X_{5s}+b_{6s}X_{6s}, \qquad (1)$$

where: $b_1, \ldots, b_6$ – regression parameters; b_0 – constant; $X_1, \ldots, X_6$ – dummy variables defined as follows (termed effects coding):

Variable Z_1	X_1
Level I	1
Level II	−1

Variable Z_2	X_2
Level I	1
Level II	−1

Variable Z_3	X_3
Level I	1
Level II	−1

Variable Z_4	X_4
Level I	1
Level II	−1

Variable Z_5	X_1	X_6
Level I	1	0
Level II	0	1
Level III	−1	−1

The estimates of part-worths (utilities) for s-th respondent are the following (see Walesiak 1996):

a) for variable with two levels

Variable Z_j	Dummy variable X_p	Part-worths (utilities)
Level I	1	$U^s_{j1} = b_{ps}$
Level II	−1	$U^s_{j2} = -b_{ps}$

b) for variable with three levels

Variable Z_1	Dummy variable X_p	Dummy variable X_q	Part-worths (utilities)
Level II	0	0	$U^s_{j1} = b_{ps}$
Level II	0	1	$U^s_{j2} = b_{qs}$
Level III	−1	−1	$U^s_{j3} = -(b_{ps}+b_{qs})$

where: $U^s_{jl_j}$ – part-worth of l-th level of j-th variable for s-th respondent, j – number of variable ($j = 1, \ldots, 5$), p, q – numbers of dummy variables ($p, q = 1, \ldots, 6$), l_j – number of level for variable j ($l_1 = l_3 = l_3 = l_4 = 1, 2$; $l_5 = 1, 2, 3$), s – number of the respondent ($s = 1, \ldots, 165$).

Next we calculate the relative importance of each attribute in the school choice process. Empirical results are presented in Table 2 and graphic representations are shown in Figure 2.

The formula (2) is used for calculating the relative importance W^s_j of each attribute for respondent s (Hair et al 1995, p. 608):

$$W^s_j = \frac{\max\limits_{l_j}\{U^s_{jl_j}\} - \min\limits_{l_j}\{U^s_{jl_j}\}}{\sum\limits_{j=1}^{m}\left(\max\limits_{l_j}\{U^s_{jl_j}\} - \min\limits_{l_j}\{U^s_{jl_j}\}\right)}. \tag{2}$$

Furthermore total utility for i-th school and s-th respondent is given by the expression (Walesiak 1996):

$$U_{is} = \sum_{j=1}^{m} U^s_{jl^i_j} + a_s, \tag{3}$$

where: l^i_j – number of level for j-th variable in school i, $i = 1, \ldots, 10$ – number of school variant, a_s – constant for s-th respondent.

For example, the total utility of school 1 for respondent 3 can be calculated from values in Table 2:

$$U_{1,3} = 8.75 + 5.0 + (-2.5) + (-3.75) + 23.33 + 37.92 = 68.75.$$

Using formula 4, the total utility for i-th school can be estimated (Walesiak 1996):

$$U_i = \frac{1}{S}\sum_{s=1}^{S}\left(\sum_{j=1}^{m} U^s_{jl^i_j} + a_s\right). \tag{4}$$

4. EMPIRICAL RESULTS

Table 2

Conjoint analysis empirical results for the overall sample and selected respondents

	Part-worth estimates										AV
	Number of respondent										
	1	2	3	–	96	97	–	163	164	165	
1. Duration											
a) one year	−0.13	22.88	8.75	–	−6.13	8.25	–	0.63	9.88	−5.0	3.83
b) two years	0.13	−22.88	−8.75	–	6.13	−8.25	–	−0.63	−9.88	5.0	−3.83
2. Work intensity (per week)											
a) 20 hour	−4.88	0.63	−5.0	–	−11.13	−2.75	–	−0.63	−14.88	−2.5	−0.12
b) 30 hours	4.88	−0.63	5.0	–	11.13	2.75	–	0.63	14.88	2.5	0.12
3. Organization											
a) full time	−2.38	1.63	−2.5	–	−3.88	1.5	–	6.63	−2.38	25.0	1.10
b) evening courses	2.38	−1.63	2.5	–	3.88	−1.5	–	−6.63	2.38	−25.0	−1.10
4. Vocational diploma											
a) yes	27.38	23.88	3.75	–	16.13	32.5	–	17.88	17.38	7.5	18.39
b) no	−27.38	−23.88	−3.75	–	−16.13	−32.5	–	−17.88	−17.38	−7.5	−18.39
5. Price per month											
a) 250 zł	−12.33	−23.17	−25.42	–	19.67	−4.25	–	−3.58	7.67	−1.67	−6.30
b) 200 zł	12.17	19.33	2.08	–	−14.83	7.25	–	4.42	2.17	8.33	−0.29
c) 150 zł	0.17	3.83	23.33	–	−4.83	−3.0	–	−0.83	−9.83	−6.67	5.59
6. Constant	52.33	26.17	37.92	–	49.83	35.75	–	70.08	52.33	51.67	47.98
7. Relative importance of each attribute (%)											
a) duration	0.27	32.56	19.72	–	11.24	16.26	–	2.10	18.54	10.53	13.31
b) work intensity	10.37	0.89	11.27	–	20.41	5.42	–	2.10	27.93	5.26	9.14
c) organization	5.05	2.31	5.63	–	7.11	2.96	–	22.27	4.46	52.63	14.87
d) vocational diploma	58.24	33.99	8.45	–	29.59	64.04	–	60.08	32.63	15.79	39.25
e) price per month	26.06	30.25	54.93	–	31.65	11.33	–	13.45	16.43	15.79	23.42
8. Pearson's *R*	0.957	0.932	0.993	–	0.866	0.987	–	0.995	0.904	1.000	0.998

AV – average value; *R* – multiple correlction between the observed and estimated preferences (this statistic displayed how well the model fits the data).

Source: The Categories option of SPSS v. 6.1 for Windows is used in analysis of this example.

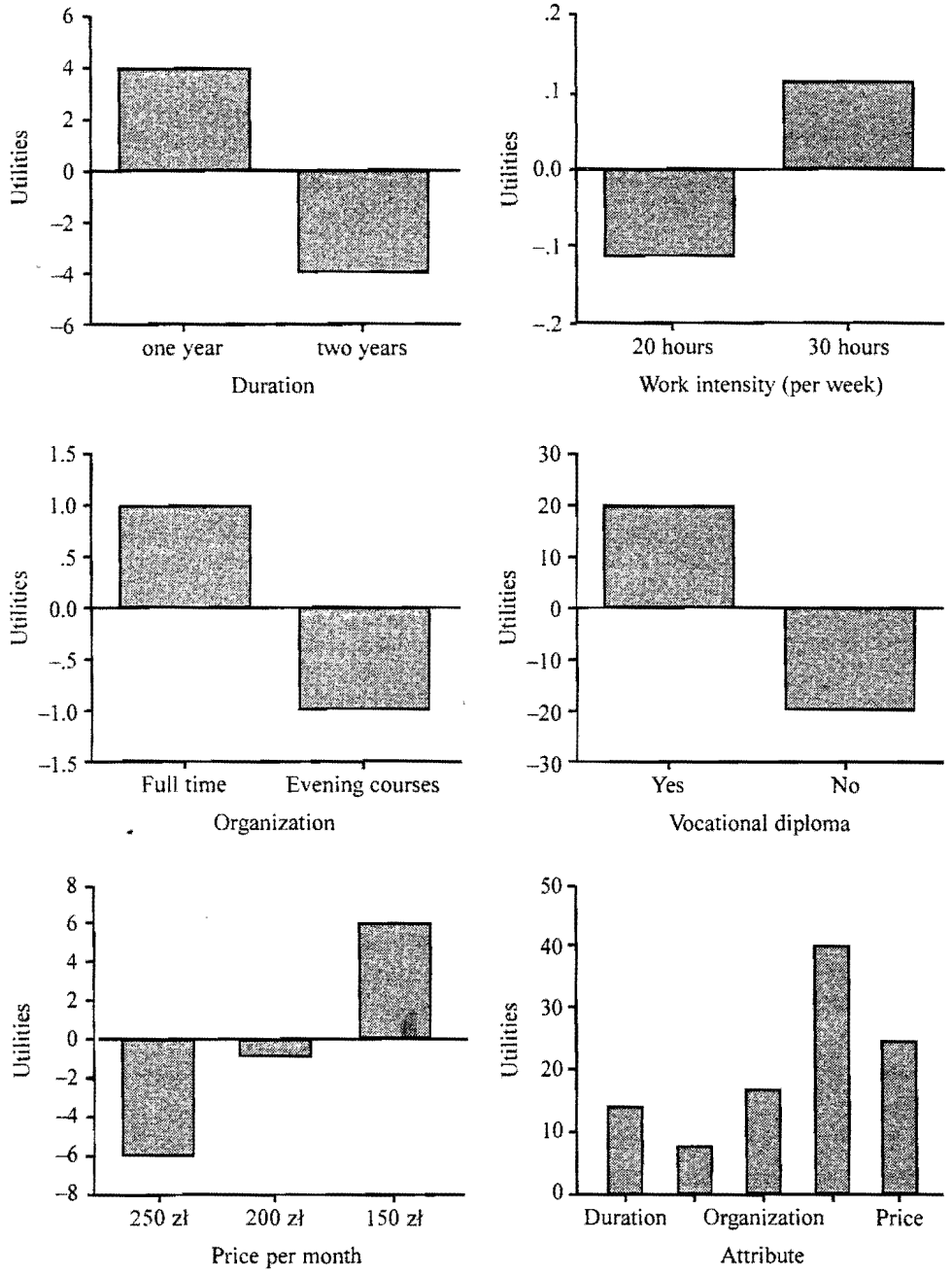

Fig. 2. Graphical results of part-worths estimates and relative importance of each attribute for aggregate results

5. ILLUSTRATIVE SIMULATION

In addition to understanding the aggregate and individual preference structures of the respondents, the results of the conjoint analysis may be used to simulate choice process. For the simulation six schools were chosen. Each card describes one school.

Card 11	
Duration	one year
Work intensity (per week)	30 hours
Organization	full time
Vocational diploma	yes
Price per month	250 zł

Card 12	
Duration	two years
Work intensity (per week)	20 hours
Organization	full time
Vocational diploma	yes
Price per month	200 zł

Card 13	
Duration	two years
Work intensity (per week)	20 hours
Organization	full time
Vocational diploma	yes
Price per month	150 zł

Card 14	
Duration	one year
Work intensity (per week)	20 hours
Organization	evening courses
Vocational diploma	yes
Price per month	200 zł

Card 15	
Duration	two years
Work intensity (per week)	20 hours
Organization	evening courses
Vocational diploma	yes
Price per month	200 zł

Card 16	
Duration	one year
Work intensity (per week)	20 hours
Organization	evening courses
Vocational diploma	yes
Price per month	250 zł

In the next phase of the project, six tested schools were divided into three groups. In the first of them (cards 11-13) a regular (full time) school is chosen, in the second group (cards 14-16) evening schools are evaluated, in the last one (cards 11-16) full time and evening courses are possible choices. Additionally, it is assumed that in simulation III one full time and one evening school is to be selected.

Predictions of the expected market shares were made with three models: the maximum utility model, BTL (Bradley-Terry-Luce) probabilistic model, logit probabilistic model (Hair et al 1995, p. 591). The maximum utility model is the probability of choosing a school as the most preferred. The BTL model computes the probability of choosing a school as the most preferred by dividing the school's utility by the sum of all the simulation total utilities. The logit model is very similar to BTL but uses the natural log of the utilities instead of the utilities.

Simulations' results are presented in Table 3. All three models indicated that in variant I simulation school 13 (in variant II simulation school 15 – full time, and in variant III simulation schools 13 – full time and 15 – evening school), would be more preferred (indicated in Table 3 by asterisk).

Table 3
Choice simulator results for the six schools formulations

School formulation	Predicted total utility (formula (4))	Market shares predictions		
		Maximum utility model (%)	Probabilistic models	
			BTL (%)	Logit (%)
Variant I				
11	65.1	40.40	32.95	39.80
12	63.2	18.28	31.80	18.57
13*	70.1	41.31	35.25	41.63
Variant II				
14	31.9	7.78	21.61	6.42
15*	61.0	60.20	41.24	63.99
16	55.0	32.02	37.15	29.59
Variant III				
11	65.1	31.01	18.89	29.26
12	63.2	8.89	18.10	9.37
13*	70.1	33.74	20.44	34.26
14	31.9	3.64	9.52	3.02
15*	61.0	17.27	17.52	18.14
16	55.0	5.45	15.53	5.95

Acknowledgements: The research presented in the paper was supported by the projekt KBN 1 H02B 016 08.
Source: author's computations.

6. CONCLUSIONS

The research shown in the article proves that the methodology under consideration (conjoint analysis, conjoint measurement) may be useful tool at least for:

- measurement of the relative importance of each individual characteristics of the market phenomena;
- constructing the product which answers to the consumers needs and expectations;
- forecasting the future market share.

REFERENCES

Anttila, M., van den Heuvel, R. R., Möller, K. (1980): *Conjoint Measurement for Marketing Management*, "European Journal of Marketing", vol. 14, No 7, pp. 397-408.

Cattin, P., Wittink, D. R. (1982): *Commercial Use of Conjoint Analysis: a Survey*, "Journal of Marketing" Summer, pp. 44-53.

Dziechciarz, J., Walesiak, M. (1995): *Pomiar łącznego oddziaływania zmiennych (conjoint measurement) w badaniach marketingowych*, [*Conjoint Measurement in the Market Research*], in: Zeliaś, A. (ed.) *Spatial Data Modelling and Forecasting in the Economy*. AE, Kraków, p. 149-158.

Hair, J. F., Anderson, R. E., Tatham, R. L., Black, W. C. (1995): *Multivariate Data Analysis with Readings*. Prentice Hall, Englewood Cliffs.

Louviere J. J. (1988): *Analyzing Decision Making. Metric Conjoint Analysis*. Sage, Beverly Hills.

SPSS Categories for Windows (1994): Version 6.1. SPSS Inc., Chicago.

Vriens, M., Wittink, D. R. (1994): *Conjoint Analysis in Marketing*. Manuscript. University of Groeningen.

Walesiak, M. (1996): *Metody analizy danych marketingowych* [*The Methods of Marketing Data Analysis*]. PWN, Warszawa.