Empirical Evidence on the Determination of the Level of Education and Income of Consumers in the Preferential Purchase Criterion based on Product Differentiation

PhD. Julio Cesar Ceniceros Angulo
Professor
Department of Economic Sciences
Universidad De Occidente
Culiacán, México

Ernesto León Castro
PhD Candidate
Research assistant
Universidad De Occidente
Culiacán, México

M.Sc. . Emilio Garcia López.
Research Assistant
Instituto Tecnológico Regional de Culiacán
México

Abstract

***Introduction**: The purpose of working on the design of discrete choice models for business and market studies strongly concentrating the attention in product differentiation and demographic segmentation variables (income and education), is to identify how the purchasing decisions of consumers are strongly influenced by the level of education and income, in such way that when the consumer has higher income and education their purchasing decision is based on differentiation over price. **Method**: For the treatment and fulfillment of the objectives in this research is used the traditional econometric methodology, concretized through binary logistic regression model (RLB). For work processing and analysis of information SPSS V.19 software is used. **Results**: Those with an income in range of 3-6 minimum wages has 1.7 times more benefit of opting for the distinction criterion in compare to those who only receive up to 2 minimum wages. Similarly, those consumers who perceived over 10 have an advantage of almost 7 times over the category of comparison. In the case of education, the reference in this category is those who have primary education, in such a way that, for a person who has professional education they opt 2.5 times for differentiation as their buying criteria in compare to the reference category. Hence, it emphasizes the role of education of the buyers on the criterion of discrimination over the product attributes. **Discussion** or **Conclusion**: An explanation of consumer behavior around discrimination through differentiation is found in the central role of education level. This can be seen when we replace some combinations of the characteristics in the model. For example, in the case where a person is in the income range of up to two minimum wages and has preparatory education the probability of choose differentiation is 0.5723, however for the same range of income but with primary education they will opt for price instead of differentiation. The importance of the level of education in the consumer selection of a product is helpful, in case of a person who reported earning more than 10 minimum wages and has professional education the probability of belonging to the group of the differentiators is high (0.9317).*

Keywords: Educational level, product differentiation and binary logistic regression

Introduction

The efforts and trend to define the type of analytical models to use for forecasting and diagnostic the commercial direction, according to Padgett, cited in Kerlinger (2001), consist in a mixture of qualitative and quantitative elements that explain the phenomenon of interest, for example in the areas of administration, economy, marketing, among other social disciplines, the most common model is the one of discrete choice.

According to Ceniceros (2001), is clear the need to establish guidelines to properly channel the effort in achieving these goals; namely design models. However, this stage need the following steps: first, the inclusion of the business in an environment of international competition unseen before, which traces the problem of reengineering and is needed to take in account new sceneries for the implementation of new competitive strategies, in the rapidly changing markets where the competitive position is always at risk. However, to make this useful, is needed to segment or stratify the group of consumers, that is why is important to correlate and determine the contribution of every variable of segmentation that the person can take in account. That is why it is study the decision of attributes of the consumers in staple products over the price, and that is why it is important to understand this dilemma regarding consumer behavior. Thus, in this research seeks to explain the strategy of discrete choices, through a model that discriminates the dependent variable in binary form (price vs attributes). For the understanding of the consumer behavior, Churchill & Petter (2000) indicate that is formed for the feelings, actions and thoughts that can influence over the consumers and provoke changes in their decisions; Richers (1984) points that the consumer is characterized for the emotional activities in selecting and buying products that satisfy their needs and desires. The main factors that influence the purchase behavior, according Schiffman & Kanuk (2000) are the cultural, socials, personals and psychological ones.

There are variables that can influence the process of buying of the consumers, Bateson & Hofmann (2001) indicate that the perception of the quality of the products depend of the comparison of the expectative of the client, that is why if the product fail to accomplish these expectations, is perceived as a low quality, in the same sense, the consumer not only evaluates the product, the perception of quality takes in account the employees, the managers, and other aspects of the organization. The strategies used for the enterprises as differentiation of products and loyalty of the client are essential to survive the fierce competition in the markets, the excess of products and the standardization of these make hard for the marketing department to achieve their goals. One of the strategies used for the organization is the differentiation of products focused on quality. Dey, Lahiri & Zhang (2014) establish that some enterprises focus their attention to consumers that look price and others look quality, thus there are segments for each one. Quality differentiation provides an alternative to consumers who seek high quality products if they are not seduced by the low price Other companies used strategies focused in loyalty of the client; Kumar & Reinartz (2006) establish that you can make consumer lifetime value to the mark in long term that is fundamental to maximize the business success. Researchers have used various methods to calculate the individual value of each consumer and to make predictions or segment them into groups, like Jain and Singh (2002), Gupta & Lehmann (2006), Kim Jung Suh & Hwang (2006), Han, Lu & Leung (2012), among others.

Method

For the treatment and fulfillment of the objectives this research is used the traditional econometric methodology, concretized through binary logistic regression model (RLB).

Research Context

This research develops in Culiacan, Sinaloa, Mexico, taking the sample information in this city. To do this, the area stratification was performed dividing the city into four regions (North, South, Northeast and Northwest) covering almost all of the malls (hypermarkets), where consumers buy staple food.

Sampling Procedure

A pilot test was performed during the summer of 2013 to determine the size of sample applied to buyers, to contrast it with the one held in June 2008 in the town hypermarkets in order to determine the new values (proportion of buyers favor the price when buying commodities) and through personal interview method interception in malls. It is important to remember that the work done during the year 2008, 310 questionnaires were applied. In this study a total of 349 questionnaires were applied. With the updated values of p and q.

Research Design

The present study corresponds to a research design non experimental and transversal where relate and explain through a discrete function type binary dependent variable (product differentiation versus prices), through covariates (segmentation).

Analytical Model

In the present treaty consider a first approximation models to validate empirically the problem of differentiation of basic foodstuffs; then it is part of the equation (1) with the addition of a stochastic error term (ε):

(1) $$k = \omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + \cdots + \omega_n x_n + \varepsilon.$$

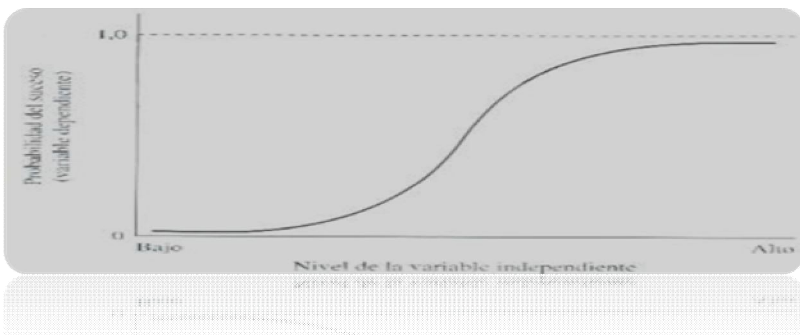
Therefore, the above expression, we can identify the following structural elements, vid., Gujarati (2000) that will go in subsequent reviewing and evaluating the feasibility of the models proposed in achieving the objectives in research:

1.) κ, theoretical value (response variable).
2.) ω_η, parameters or coefficients of the equation.
3.) x_n, independent variables.

4.) ε , residual or stochastic error term.

Consider, then, the model of logistic regression (LR), in the light of its functional structure. Firstly, concerning the k values, we need to Hair et Tatham and Black (1999), in a first version of this model considered as a dichotomous variable (binary), that is, referred to a variable response two groups, unlike Multiple Regression (MR) that predict the probability of occurrence of the phenomenon to be analyzed. So the response values are bounded between 0 and 1. To model the functional relationship between k and Xn, Hair, et, al, (1999) present us with the following sigmoid representation:

Figure 1: Representation of the Logistic Sigmoid Function



Source: Taken from Hair, et, al, (1999, P. 281)

Specifying the functional widespread part of in its operational form, according Gujarati (2000), we have:
If, P_i = probability of success of an event. One way to model a problem with dichotomous dependent variable can be:

(2) $$p_i = \frac{1}{1 + e^{-z_i}}$$ Logistics Distribution Function.

Where $z_i = \beta_1 + \beta_2 x_i$

$\therefore$ The probability of occurrence of the event can be set as:

$$1 - p_i = \frac{1}{1 + e^{-z_i}}$$ $\Rightarrow$ The response variable can be expressed as follows odds ratio:

(3) $$\frac{p_i}{1 - p_i} = \frac{1 + e^{z_i}}{1 + e^{-z_i}}$$

However, this model can also be presented as follows in relation to its variable response, so to Pyndyck and Rubinfeld (2001), the model is based on the following expression logistic cumulative probability:

(4) $p_i = f(z_i) = f(\alpha + \beta x_i) = \dfrac{1}{1+e^{-z_i}} = \dfrac{1}{1+e^{-(\alpha+\beta x_i)}}$

Where e, base of natural logarithms ≈ 2.718, the author takes (4) and multiply both sides of the equation by $1+e^{-z_i}$ and is obtained $(1+e^{-z_i})p_i$=1, then divide between p_i and subtracting 1, we have:

$e^{-z_i} = \dfrac{1}{p_i} - 1 = \dfrac{1-p_i}{p_i}$, as $e^{-z_i} = \dfrac{1}{e^{z_i}}$ $\therefore$ $e^{z_i} = \dfrac{P_i}{1-p_i}$ applying the natural logarithm on both sides, we have:

$$Z_i = \log \frac{p_i}{1-p_i}$$

Therefore, considering (3), can finally express the response variable as:

(5) $\log \dfrac{p_i}{1-p_i} = Z_i = \alpha + \beta x_i$

Main Authors in Logistic Regression are Hosmer and Lemeshow (1989), in his classic work Applied Logistic Regression, reasoning as follows in relation to the expected value of the response variable in a linear function as:

$$E(y/x) = \beta_0 + \beta_1 x$$

In which states that moves in ranges $-\propto$ and $\propto$. But with dichotomous response variables type ranges are set to $[0 \le E(y/x) \le 1]$ If, $\pi(x) = E(Y/x)$, Therefore, the logistic model is specified as:

(6) $\pi(x) = \dfrac{e^{\beta_0 + \beta_1 x}}{1+e^{\beta_0 + \beta_1 x}}$

So it is necessary to re-apply the anti transform log, the signs of the coefficients, then be interpreted as follows:

(7) $\dfrac{prob_{(evento)}}{prob_{(noevento)}} = e^{\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n}$

We cannot omit here that as a structural part of the model have the error term, residual or random disturbance (ε). So Pyndyck & Rubinfeld (2001) this term relate the probability of success with the explanatory variables, assuming that the average residual is zero. Then, because:

E (ε_i)= (1-$\alpha - \beta x_I$) $\rho_I + (-\alpha - \beta x_i)$ (1- p_i)= 0, so in terms of $p_i = \alpha + \beta x_i$,

$\therefore$ $1 - p_i = 1 - \alpha - \beta x_i$

Table 1: Distribution of Probabilities ε_i

y_i	ε_i	probability
1	1-$\alpha - \beta x_i$	p_i
0	-$\alpha - \beta x_i$	**1-** p_i

Source: Taken from Pyndyck and Rubinfeld (2001, P. 314)

Another way of formulating the above we have in Hosmer and Lemeshow (1989), when it sets the output value as: $y = \pi(x) + \varepsilon$, where π, is the probability of success of the event in question, $\therefore$ If, $y = 1$, then, $\varepsilon = 1 - \pi(x)$ likely $\pi(x)$ and the case further, $y = 0$, then $\varepsilon = -\pi(x)$, likely 1- $\pi(x)$, so the residual is distributed according a ~ (0, $\pi(x)$ [1-$\pi(x)$]). Where the mean of a binomial distribution is obtained from $\mu = n\pi$, however variance, is obtained from $\sigma^2 = n\pi(1-\pi)$. According to Lind, Marchal and Wathen (2005). In conclusion, Gujarati (2000) states that the error distribution (u_i), when the number of cases is high (N), follows a normal distribution (N) as: (8) $u_i \sim N[0, \frac{1}{N_i p_i (1-p_i)}]$

Results

Let us turn now to consideration how the educational level of respondents with the approach used in the purchase of edible oils are related. The exact observation of the structure of this crossing of the variables can be represented by the following table.

Table 2: Contingency Table Edu * Selection Criteria

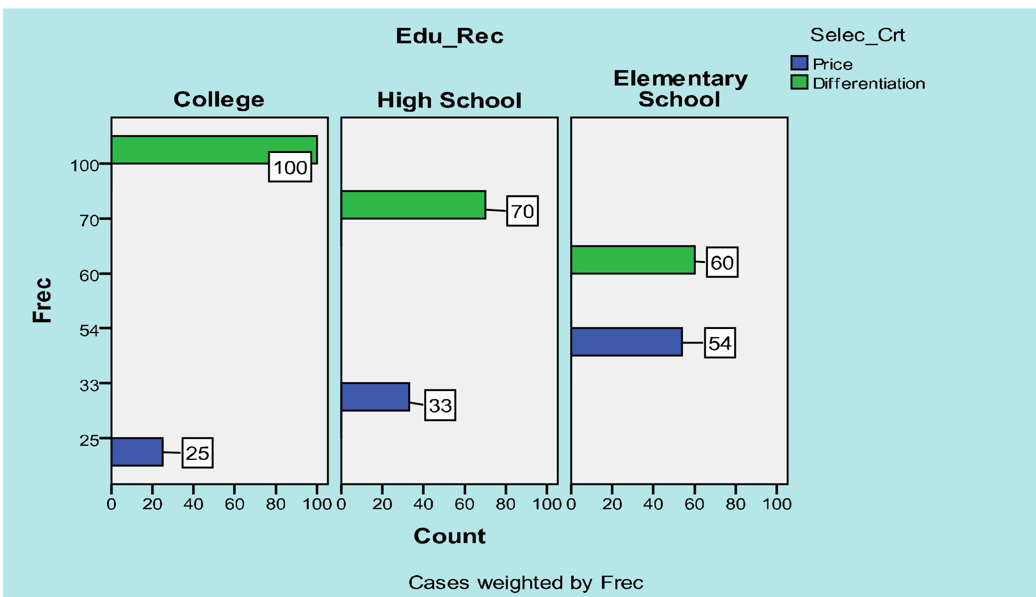
			Selection criteria.		**Total**
			Price	Differentiation	
Edu	unfinished primary	count	9	11	**20**
		% of total	2.6%	3.2%	**5.8%**
	Elementary School	count	21	18	**39**
		% of total	6.1%	5.3%	**11.4%**
	Middle School	count	24	31	**55**
		% of total	7.0%	9.1%	**16.1%**
	High School	count	33	70	**103**
		% of total	9.6%	20.5%	**30.1%**
	College	count	22	91	**113**
		% of total	6.4%	26.6%	**33.0%**
	Graduate School	count	3	9	**12**
		% of total	.9%	2.6%	**3.5%**
Total		**count**	**112**	**230**	**342**
		% of total	**32.7%**	**67.3%**	**100.0%**

From the information above, it is clear that higher levels of study (undergraduate and graduate), the trend clearly favors the purchase criterion based on comparison and assessment of product attributes. Generally, the higher level of education increased earned income and course of studies at those levels greater social recognition is sought. So consider that important factors are; first, on discrimination of stimuli related to learning consumer behavior, second, class being the same attitudes, lifestyles and behavior similar purchases are shared. Kerin, Hartley and Redelius (2009). By the above argument, it seems interesting to reclassify luck variables such that reconfigure a cluster consisting of those people who only have the level of education through high school and another made up what ranging from high school to post-graduate level. To clarify this proposal, consider the following information (reclassified).

Table 3. edu_rec * Selection Criteria

			Selection Criteria.		Total
			Price	Differentiation	
Edu_rec	Elementary School	count	54	60	**114**
		% of total	15.8%	17.5%	**33.3%**
	High School	count	33	70	**103**
		% of total	9.6%	20.5%	**30.1%**
	College	count	25	100	**125**
		% of total	7.3%	29.2%	**36.5%**
Total		count	**112**	**230**	**342**
		% of total	**32.7%**	**67.3%**	**100.0%**

Figure 2: Educational Level of Respondents



Source: own.

Now, we proceed to formalize the significance test through the following hypothesis first.

H_0: "The variable educated buyers have no connection with the purchase criteria used either the price or differentiation by attributes."

Table 4: Chi-Square Tests

	Valor	degrees of freedom	Sig. asintótica (bilateral)
Chi-square of Pearson	20.312[a]	2	.000
Likelihood ratio	20.533	2	.000

As tested, the test turns out to be highly significant. So the null hypothesis is rejected and thus proceeds to select important demographic segmentation variable in shaping the logistic regression model. Is this appropriate candidate variable in the model? Let's analyze this situation.

Hypothesis Testing with Wald Statistic and Bootstrap Procedure for Education

Before testing the hypotheses of individual significance of the variable educated, proceed to encoding the same as shown in Table 5.

5: Encodings Categorical Variables

		frequency	Encoding parameters	
			(1)	(2)
edu_rec	Elementary School	114	.000	.000
	High School	103	1.000	.000
	College	125	.000	1.000

As can be seen from the above table, basic education category is the category of reference or comparison. ¿Why this category was selected as a reference? We tried to investigate the role of education in the criterion of discrimination and is part of supposing that those with more education tend to attribute discrimination on the criterion of purchase. Now ¿Is coefficient associated with the covariate level equal to zero in the study population? Put another way, are significantly related the independent and dependent variables? Formally express the null hypothesis.

H_0: "The education variable is not associated in explaining the way how consumers decide to use binary criteria purchase."

For having more elements of reliability and enhance the accuracy of the results of hypothesis testing is performed using the bootstrap procedure with a total of 1,000 samples generated. The results are presented below.

Table 6: Bootstrap Specifications

Sampling Method	single
Number of samples	1000
Level of confidence interval	95.0%
Type confidence interval	percentile

Table 7: Variable and Categories in the Equation

		B	E.T.	Wald	degrees of freedom	Sig.	Exp(B)
Paso 1[a]	Educareca			19.473	2	.000	
	Educareca(1)	.647	.282	5.241	1	.022	1.909
	Educareca(2)	1.281	.292	19.261	1	.000	3.600
	constant	.105	.188	.315	1	.574	1.111

Table 8: Bootstrap for Variables in the Equation

		B	Bootstrap				
			Bias	Tip. Error	Sig. (bilateral)	Confidence interval 95%	
						lower	upper
Paso 1	Educareca(1)	.647	.012	.291	.023	.077	1.251
	Educareca(2)	1.281	.012	.295	.001	.735	1.903
	constant	.105	.005	.186	.555	-.247	.484

If you look carefully you will notice that the coefficients pair the Educareca (1) and Educareca (2) variables fall within the respective confidence intervals, so likewise the H_0 is rejected coming to the same conclusion. As we shall return later as significant covariate in the construction of discrete choice model.

Testing Hypothesis of Independence between the Variables Income Level and Purchasing Criteria

To analyze the significance or otherwise of this strategic variable from the economic and marketing theory within the block of demographic segmentation, proceed to the use of the bootstrap technique.

Table 9: Contingency table Ing* Selection Criteria

			Selection Criteria.		Total
			Price	Differentiation	
Ing	Up to 2 minimum wages	count	45	46	91
		% of total	13.0%	13.3%	26.4%
	3-6 minimum wages	count	53	119	172
		% of total	15.4%	34.5%	49.9%
	7-10 minimum wages	count	14	44	58
		% of total	4.1%	12.8%	16.8%
	More than 10 minimum wages	count	2	22	24
		% of total	.6%	6.4%	7.0%
Total			114	231	345
		% total	33.0%	67.0%	100.0%

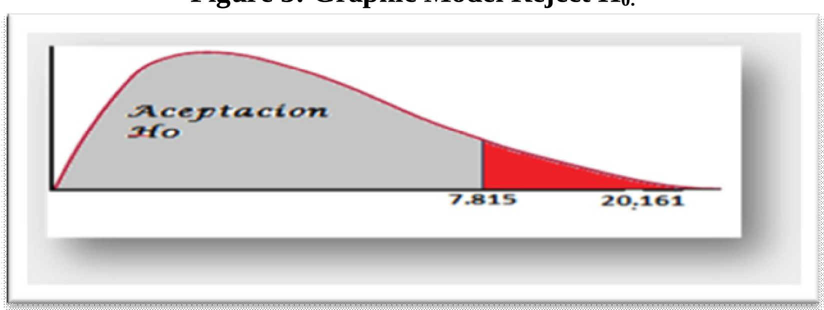
As we can see in the table below the test statistic is highly significant with a calculated Pearson Chi-square value of 20,161 and a very low probability value (.000)

Table 10: Chi-Square Tests

	value	degrees of freedom	Sig. asymptotic (bilateral)
Chi-square test	**20.161**	**3**	**.000**
Likelihood ratio	21.321	3	.000

For additional information is incorporated in Figure 3, the theoretical value for 3 degrees of freedom and level of α = 0.05 significance.

Figure 3: Graphic Model Reject H_0.



In short, we now have our second variable within the block of demographic segmentation incorporating the proposed model. Therefore, the multiple logistic regression model is developed as follows:

Table 11: Summary of Case Processing

Unweighted cases		N	percentage
Selected cases included in the analysis	Selected cases included in the analysis	339	96.6
	lost cases	12	3.4
	Total	351	100.0
Unselected cases		0	.0
Total		351	100.0

Where it is necessary to encode the price variable with zero and differentiation variable with number 1.

Table 12: Encoding for Categorical Variables

		frequency	Encoding parameters		
			(1)	(2)	(3)
Ing	Up to 2 minimum wages	89	.000	.000	.000
	of 3-6 minimum wages	169	1.000	.000	.000
	From 7-10 minimum wages	57	.000	1.000	.000
	More than 10 minimum wages	24	.000	.000	1.000
edu_rec	Elementary School	114	.000	.000	
	High School	100	1.000	.000	
	College	125	.000	1.000	

In the case of income level reference variable is the category consisting of up to 2 minimum wages. With regard to education the reference category is basic education.

Table 13: Variables that are not in the Equation

			Rating Rao.	degrees of freedom	Sig.
Step 0	Variables	monthly income	19.686	3	.000
		Monthly_ Income (1)	.784	1	.376
		Monthly_ Income (2)	2.226	1	.136
		Monthly _Income (3)	7.126	1	.008
		Educareca	20.187	2	.000
		Educareca(1)	.000	1	.992
		Educareca(2)	15.216	1	.000
	Statistical global		28.939	5	.000

Before estimating the model, Rao score statistic indicates the individual contribution (weight) of each variable in the improvement of the model fit. As shown Educreca has a greater contribution if monthly monetary income is similar. Finally, the overall statistical is considering all the variables together is also highly significant As for the omnibus test, allows us to test the null hypothesis that the incorporation of all covariates taken together do not significantly improve the fit of the overall model. Is performed through a Chi square test. (x^2).

Table 14: Testing Omnibus on the Model Coefficients

	Chi square	df	Sig.
Step 1 Step	30.460	5	.000
block	30.460	5	.000
model	30.460	5	.000

That is, H_0: monthly income, Edureca = 0. With the results of the above table the null hypothesis is clearly rejected. Regarding the Hosmer and Lemeshow test, the result indicates that H_0 should not be rejected and thus the model predictor.

Table 15: Hosmer and Lemeshow

step	Chi square	gl	Sig.
1	4.208	7	.755

With regard to the leader board or confusion matrix is likewise a test of the predictive ability of the model because through this we would like to know the percentage of cases correctly classified by the main diagonal. In this case without adjusting the cut in the overall percentage .5, corresponds to 68.1%.

Table 16: Ranking Table [A]

Observed			predicted		
			Selection criteria.		percent correct
			Price	Differentiation	
Step 1	Selection criteria.	Price	26	86	23.2
		Differentiation	22	205	90.3
	overall percentage				**68.1**

a. The cut is .500

Finally, in the construction of discrete choice model considering covariates monthly income and education level of consumers within the block of demographic segmentation, we have according to Table 17, the following function multiple logistic regression.

Table 17: Variables in the Equation

		B	E.T.	Wald	df	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								lower	Upper
Paso 1[a]	monthly income			8.568	3	.036			
	Monthly_ Income (1)	.586	.281	4.340	1	.037	1.797	1.035	3.118
	Monthly_ Income (2)	.687	.408	2.845	1	.092	1.989	.895	4.421
	Monthly _Income (3)	1.931	.790	5.978	1	.014	6.897	1.467	32.431
	Educareca			9.526	2	.009			
	Educareca(1)	.560	.288	3.774	1	.052	1.750	.995	3.079
	Educareca(2)	.951	.319	8.908	1	.003	2.589	1.386	4.835
	Constant	-.268	.247	1.186	1	.276	.765		

a. Variable (s) entered (s) in step 1: Monthly_ Income, Educareca.

Therefore, the interpretation by example in terms of the odds ratio (odds ratio) is that whoever has an income between 3-6 minimum wages is 1.7 times more advantage of opting for the distinction criterion that only gets up to 2 minimum wages. Notice how the perceiver over 10 wages has an advantage of almost 7 times more on the comparison category. In the case of education as the reference category is that who have primary education, in such a way that for a subject having professional instruction the advantage of opting for differentiation as buying criteria is 2.5 times the category reference. Having brought these thoughts to the meaning of the modified coefficients (exp), we express in linear terms and consider the function: I, is monthly_ income and Edu's is education level then we have.

(9) $Pr = -.268 + .586\,(I_1) + .687\,(I_2) + 1.931\,(I_3) + .560\,(Edu_1) + .951\,(Edu_2)$

But the intention is to predict and classify the linear function has to transform a logistic function as follows:

(10)
$$P = \frac{1}{1 + e^{-[-.268 + .586\,(I1) + .687\,(I2) + 1.931\,(I3) + .560\,(Edu1) + .951\,(Edu2)]}}$$

Discussion and Conclusions

An explanation of consumer behavior regarding discrimination through differentiation is found in the central role of education. To see some of the sample results in the model replace some combinations of features regarding both demographic covariates. As can be clearly seen in Table 18 to low levels of schooling (basic education and income up to 2 minimum wages), the probability (PRE_1) that a consumer chooses the criterion of differentiation is only 4333 and therefore, the model classifies the group of prices. (Column PGR_1). Now, consider the case where a person registered or has declared also receive up to 2 minimum wages nevertheless take preparatory instruction, then expresses a probability of 0.5723 to opt for differentiation and indeed the model classifies it as a differentiator. It is revealing the role because of consumer education. Note in contrast as persons who reported earning more than 10 minimum wages and have professional instruction the probability of belonging to the group of differentiators is high (.9317). For more detail see Table 23. Where different forecasts appear as the value assigned to demographic covariates.

Table 18: Forecasts and Membership Groups

Income	educa_recat	PRE_1	PGR_1
Up to 2 minimum wages	Elementary School	0.4333	Price
3-6 minimum wages	Elementary School	0.5787	Differentiation
3-6 minimum wages	College	0.7805	Differentiation
Up to 2 minimum wages	High School	0.5723	Differentiation
More than 10 minimum wages	College	0.9317	Differentiation
7-10 minimum wages	College	0.7974	Differentiation
Up to 2 minimum wages	College	0.6643	Differentiation

Source: Authors.

References

Alvarez, Rafael. (1995). Multivariate Statistics with SPSS Parametric not. Application to Health Sciences. Spain: Diaz de Santos.

Bateson; E. G.& Hofman D. (2001). Marketing de serviços. Bookman, Porto Alegre.

Ceniceros, Julio. (2001). Export forecast model Kabocha Pumpkin Japanese market. Unpublished master's thesis, University of the West, Culiacan, Sinaloa, Mexico.

Churchill, G. A. & Peter, J. P. (2000). Marketing: criando valor para o cliente. Saraiva, São Paulo.

Dey D., Lahiri, A. & Zhang, G. (2014).Quality Competition and Market Segmentation in the Security Software Market. MIS Quarterly. Vol. 38 No. 2, pp. 589-606/June 2014

Gujarati, Domadar. (2010). Econometrics. (Fifth Edition). Mexico: Mc Graw Hill.

Gupta, S., & Lehmann, D. (2006). Managing customers as investments. Upper Saddle River, NJ: Wharton School Publishing.

Hair, Joseph Anderson, Rolph, Tatham, Ronald. and Black, William. (1999). Multivariate Analysis. (5th Edition). Spain: Prentice Hall Iberia.

Han, S. H., Lu, S. X.,& Leung, S. C. H. (2012). Segmentation of telecom customers based on customer value by decision tree model. Expert Systems with Applications, 39, 3964–3973.

Hosmer, David and Lemeshow, Stanley. (1989). Applied Logistic Regression. United States of America: Wiley Interscience Publication.

Jain, D., & Singh, S. (2002). Customer lifetime value research in marketing: A review and future directions. Journal of Interactive Marketing, 16, 34–46.

Kerlinger, Fred.& Lee, Howard. (2001). Behavior Research. Mexico: Mc Graw Hill.

Kim, S.Y., Jung, T.S., Suh, E.H., & Hwang, H.S. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. Expert Systems with Applications, 31, 101–107.

Kumar, V., & Reinartz, W. (2006). Modeling Customer Lifetime Value. Journal of Service Research, Volume 9, No. 2, November 2006 139-155. DOI: 10.1177/1094670506293810. 2006 Sage Publications.

Lind, Douglas, Marchal, William & Wathen, Samuel. (2005). Statistics for Business and Economics. Mexico: Mc Graw Hill.

Pindyck, Robert. and Rubeinfeld, Daniel. (2001). Econometric Models and Forecasts. (4th Edition): Mc Graw Hill.

Richers, R. (1984). O enigmático mais indispensável consumidor: teoria e pratica. Revista de Administração, jul/set 19(3): 45-56.

Schiffman, L. G. & Kanuk, L. L. (2000). Comportamiento do Consumidor. Rio de Janeiro: Editora LTC.