ERASMUS UNIVERSITY

ERASMUS SCHOOL OF ECONOMICS

Customers' ranking inabilities in optimizing customer journeys

Thesis in Business Analytics and Quantitative Marketing

Jovana Milanović
385619

Abstract

This paper is an introduction in how customer satisfaction and customer journeys can be optimized using ranked preferences of customers. Individuals are asked to rank similar alternatives from most preferred to least preferred. Due to the complexity of providing a complete ranking, one needs to deal with the ranking inability of the individuals. The latent-class rank-ordered logit model will be considered for estimating the preferences efficiently, keeping in mind the ranking inabilities of individuals. The model uses latent segments to recognize the ranking abilities of individuals. The latent-class rank-ordered logit model will be applied on two datasets about gaming platforms to obtain how this model can be implemented in a practical way for an electronic company and how this model can help improving the customer satisfaction and optimizing the customer journeys.

supervised by
Prof Dr. D. FOK
second supervisor
A. CASTELEIN

July 2, 2017

1 Introduction

Nowadays data play an enormous role in the decision making processes of companies. Having the knowledge of their customers' decisions, leads to improvement in marketing content [Kalyanaraman and Sundar, 2006], strategies and various other aspects of the company strategy [Wedel and Kamakura, 2012]. Looking at electronic companies, we see the various brands they are offering online and offline for a similar product type. When customers visit their online website and search for particular products, they get a list of various brands offering that specific product. Those products are directly displayed in a particular order, like from most purchased to least purchased. It is often possible to sort the products in a different way, for example from cheapest to most expensive. The way how the products are displayed to the customers at first sight is very important. They do not solely see the offered products and prices at first sight, but they already form an opinion if this site is useful and if the offered products are as wished. Displaying preferred products for the customers would be beneficial for the company [Kalyanaraman and Sundar, 2006]. The biggest obstacle for displaying the right products for the right customers is not having any information about the customers who are visiting the website. It is very complicated to predict what the customers are going to consider buying, because of the lack of information about them. Without that information, the website of the electronic company shows mostly a "most bought" sorting order as seen in figure 2. Nowadays companies can find out who the visiting customers are by their IP-address or because they are logged in on their personal shopping accounts of the website. This is also a way how the company keeps track of their customer journeys and knows personal information about the customers, like gender, age and place of residence. This personal information may be crucial to predict the buying behaviour of the customers. The company can also take out questionnaires if they want to know some specific information about the customers. All this information helps the company in optimizing the customer journey.

Taking out questionnaires can help the company acquire and increase their knowledge of the customers. It can also help the company acquiring knowledge about the preferences of the customers. Knowing the preferences the company can improve the marketing content and change the sorting order as seen in figure 2 from " most bought" to a more personalized sorting order, wherein the preferences of the customers can been seen. The preferences will be used to create a more efficient marketing content for the website page, but in which sorting order should those preferences be displayed online? This can be answered when considering the ranked preferences of the customers. The ranked preferences tell which products the customer ranks from most preferred to least preferred. This crucial information of ranked preferences should be used when considering a more personalized sorting strategy on the website. Combining the knowledge of personal information and ranked preferences could give higher conversion rates, better marketing strategies and in the end higher profits, but also happy and more loyal customers [Gibbert et al., 2002].

Striving to more happy customers and the remainder benefits as mentioned above could, inter alia, be accomplished by displaying a more personalized sorting order of the offered products. But do the customers actually know their ranked preferences if they do not own the products? Can they rank the products in an accurate way? It is important to find out for the company to what extent customers can rank products to provide the most personalized sorting order for the customers. Based on data of the preferences of customers and the right techniques and models to find this ranking ability and preferences of customers, the personalized sorting order and more efficient

marketing content will be found.
In this paper a more personalized sorting order and marketing content will be found by introducing a model which captures the ranking ability and preferences of customers. The respondents of the questionnaire will be asked to rank alternatives from most preferred to least preferred. If customers can only choose their most preferred alternative, the *Multinomial Logit model* (MNL) [McFadden et al., 1973] will be used. For a more personalized sorting order it is important to analyze the complete ranking of various alternatives. The *Rank-Ordered Logit* (ROL) model [Beggs et al., 1981] will consider a complete ranking of the alternatives. Analyzing the complete ranking, the preferences could be estimated more efficiently. Unfortunately in real life giving a complete ranking of often very similar alternatives could be too complex for respondents. Ranking the first two or three alternatives should be reachable, but more can be too difficult. Respondents may not know their complete ranking and provide only the first part of the ranking. The rest of the ranking is left out or they just copy the order of the remaining alternatives as given in the survey. This could lead to biased parameter estimates. If all respondents give a complete ranking or all the same size of incomplete ranking, this would not be a problem. In practice the ranking capabilities differ among respondents. This would lead to a loss in efficiency in this model. Overall if respondents report their least preferred alternatives in a random way, the ROL model will result in having biased parameter estimates.
Because the electronic company still wants to display all the alternatives online, it is important to know until which part of the ranking the alternatives are sorted in an accurate way by the customers. To deal with this ranking inability of the respondents, this paper will use a third model in which is tried to solve the ranking inability. All observed rankings will be taken into account keeping in mind that some rankings are not representing the true ranking preferences of the respondents. A ROL model using latent segments to endogenously identify ranking capabilities is now used. From now on the individual-specific ranking capabilities are allowed [Fok et al., 2012]. The *Latent-Class Rank-Ordered Logit* (LCROL) model will be suggested to use as the third model.
Using the LCROL model to capture the ranking inabilities of the customers will help to find an accurate personalized sorting order for the website. The following question will be answered for this;

To what extent can a more personalized sorting method be introduced?

After knowing to what extent customers can rank the alternatives and a more personalized sorting method can be introduced, it is important to find out how this knowledge of ranking inabilities can be used in a more practical way. The ranking abilities can be used in practice to find out what the preferences of the customers are and which needs the customers have. Using those preferences can help the company to find a way in which products can be promoted on the (most important) banners of the website. Knowing the preferences of the customers the following question should be investigated;

Keeping in mind the preferences and ranking ability of customers, what will be an effective way to improve marketing content to assess the needs of the customer?

The research will be done using two datasets containing information about the ranking of gaming platforms. A couple of years ago, before the smartphone was used as a

highly popular gaming platform among customers, a research was done for the ability of ranking different gaming platforms [Fok et al., 2012]. In the last ten years the popularity of having a smartphone as gaming platform has grown and nowadays it is normal that almost everyone has a smartphone. By introducing the smartphone as an gaming platform in a similar new survey, could give different insights. Using the information of preferences of the customers, the company could find new ways to promote and sort products in a more personalized way.
The paper is structured as follows. In the first part of the paper a more theoretical part about the models will be discussed, next to the discussion of the data and review of existing literature about customer satisfaction and journeys. After the theoretical part of the paper, the last part of the paper will be more focused on the practical use of the best performing model. In the practical part, the models will be applied on two datasets of which the results will lead to knowledge of the ranking abilities and preferences of customers. In the end a more managerial advice will be given for the electronic company using the insights of the results.

2 Literature

Before the theoretical part of the paper will be mentioned, it is important to find out what the link of optimizing customer journeys is with the ranking inabilities of customers. This link will show how the theoretical part of the paper can be implemented in a business environment.

2.1 Customer Satisfaction and Journey

Nowadays customer satisfaction takes high priority in companies. Customer satisfaction is important for customer retention and market share of the company. Customer satisfaction is, among other things, determined by the quality [Cronin Jr and Taylor, 1992] and offerings of the service, but also how the company deals with customer complaints. [Hansemark and Albinsson, 2004]. At marketing departments of companies, analyst do various research using the big amount of data that is available of customer satisfaction. The customer satisfaction data is determined by the customer journeys. The customer journey is the complete experience a customer goes through when interacting with the brand or company [Sorman, 2014]. Keeping the customer satisfaction high means that the customer will stay loyal to your brand or company, by intentions for a repurchase and the willingness to recommend the company [Van Der Wiele et al., 2002]. Improving customer satisfaction could lead to higher sales, profits, increase in market share [Jacobson and Aaker, 1987] and more efficient marketing strategies.

2.2 Customer Experience Strategy

Companies are continuously trying to improve the customer journey and act according to a specific *Customer Experience Strategy*. In figure 1 a *Customer Experience Strategy* is shown [Morgan, 2013].

Figure 1: Customer Experience Strategy

Looking at the top of the *Customer Journey Experience* in figure 1, the stages *"Measure and Develop"* and *"Assess needs and segment customers"* are shown. The way how the touchpoint website[1] will be developed by measuring the customers preferences ranked data and how it can help assessing the needs of customers will be investigated in this paper.

3 Data

Technology is improving every year, which means that electronic companies are very dynamic and continuously offering new products and get rid of older version of the same products. Using older datasets for improving customer satisfaction and optimizing customer journeys will therefore be useless. This paper will use two datasets, one from 2007 and introduce a newer one from 2017. The new dataset is developed, because already ten years passed from when the previous research [Fok et al., 2012] was done combined with a lot of technical improvements in those last ten years. The comparison of the results of the two datasets could lead to interesting findings and differences. The aim of the questionnaires is that one wants to learn something about ranking abilities and preferences of a group of respondents and apply these learnings to a more general knowledge of customers.
In this part of the paper, the datasets will be introduced and some insights will be given, found in the data.

[1]A touchpoint is any time a potential customer who comes in contact with your brand before, during, or after they purchase something from you. The touchpoint website is a touchpoint during the purchase.

3.1 Gaming data 2007

In previous research [Fok et al., 2012] on the ROL model with unobserved heterogeneity in ranking capabilities, there is a application on gaming platforms. 91 Dutch students were asked to rank gaming platforms from most preferred to least preferred. They could choose from six alternatives: X-box, PlayStation, Gamecube, PlayStation Portable, Gameboy or a regular personal computer (PC). Next to the ranking, the ownership of the different gaming platforms, average hours of playing per week and the gender and age of the students is provided in the dataset. The percentage of respondents owning the platforms can be found in table 1

Table 1: Percentage of respondents owning the platform 2007

	Xbox	PlayStation	PSPortable	GameCube	GameBoy	PC
% respondents	13.2%	31.9%	9.9%	8.8%	13.2%	87.9%

By analyzing the data of 2007, it is found that the PC was the most preferred platform (by 43% of the students) for playing games, followed by the PlayStation (by 21% of the students) and Xbox (by 20% of the students). The students gamed on average 3.9 hours per week. 25% of the students were females and the average gaming hours per week of those female students was equal to 1.3 hours.
In 2007 the respondents thought that importance of graphical capabilities are the most important items for a gaming platform, followed by availability of games. Importance of portability was the least important item.

3.2 Gaming data 2017

In 2017 122 respondents filled in the new survey on gaming platforms[2]. The main question for the respondents is still to give a complete ranking of the gaming platforms from most to least preferred. The remainder questions correspond with the old survey, but there are some differences.
First of all the use of the smartphone has extremely grown in the previous years and is therefore introduced as a new gaming platform [Nielson Company, 2014]. Expecting the smartphone to be an important introduction in gaming platforms, could change some of the results. Next to the introduction of the smartphone as a gaming platform, the GameCube alternative has been taken away from the alternatives. This is done due to the discontinuity of the production and retail availability of it [Parfitt, 2007]. The PlayStation Portable and GameBoy are replaced for a more general Portable Gaming Platform and a Wii. I did not made a distinction in different generations of the same platform. Next to that I asked the same questions about ownership, importance of specific features, age, gender and hours of gaming as in the previous research. I added one more question about having a Android or IOS device on the smartphone. The percentage of respondents for the ownership of the platforms can be seen in table 2

Table 2: Percentage of respondents owning the platform 2017

	Xbox	PlayStation	Portable Gameboy	Wii	smartphone	PC
% respondents	11.5%	27.0%	5.7%	19.7%	100%	41.0%

[2]The survey can be found in the appendix.

From the respondents who *only* own a smartphone, 45.5% most prefer the smartphone as gaming control. Of these 45.5% of respondents, 90% are female respondents.
In 2017 most people (35%) ranked the smartphone as most preferred platform, followed by the PlayStation (25%) and PC (15%). The decrease of ownership of the PC can be explained due to my question if respondents own a PC as gaming console. Next to the formulation of the question, nowadays the PC market is not really accessible for gaming anymore. Hardware is nowadays most of the time produced without a CD drive, what means that using PC games (which are still in the form of a CD) is only possible with external CD drives. Technological improvements made it difficult to keep on playing games on a PC what could lead in decreasing ownership of a PC as gaming platform.
The respondents game on average 4.9 hours in 2017. 59.8% of the respondents are females and their average gaming hours per week is equal to 2.3 hours.
Respondents of the survey consider the price of soft- and hardware together with the exchangeability of games as the most important features of gaming consoles. Nowadays importance of portability is still seen as least important item for gaming platforms.

4 Methodology

In this section the focus lies on the LCROL model which captures the unobserved heterogeneity in ranking abilities best [Fok et al., 2012]. It is assumed that respondents can choose from a fixed set of alternatives and rank their most preferred item as 1 until their least preferred as the number of alternatives. I will start off with a simple model where the respondents need to choose their most preferred alternative only. After that, I will extend this model to a model where a complete ranking needs to be provided. In the end the LCROL model will be explained.

4.1 Utility

Individuals are asked to rank the alternatives. Based on the utility functions of the alternatives, individual i will choose his most preferred alternative. The alternatives will be indicated by $j = 1, ..., J$ and the individuals by $i = 1, ..., N$. The models are based on the principles of utility maximization. The utilities for the individuals combined with the alternatives are given by $U_{i1}, ..., U_{iJ}$. Individual i preferring alternative j most will be indicated by $y_{ij} = 1$. If this holds, it will imply that the utility of individual i choosing alternative j will have the highest value comparing with the other utilities. This is denoted by $U_{ij} \geq \max\{U_{i1}, ..., U_{iJ}\}$. (V_{ij}) is the deterministic part of utility depending on individual- and alternative-specific factors of the individuals. The random term (ϵ_{ij}) is the stochastic component of the utility [Chapman and Staelin, 1982]. The deterministic part will be determined by individual characteristics

$$V_{ij} = x_i'\beta_j + w_{ij}'\gamma_m, \tag{1}$$

where x_i is a vector containing the characteristics of individual i, $w_{i,j}$ are the alternative-individual specific characteristics of individual i for every alternative j. The parameters are β_j, vector containing the parameters for alternative j and γ_m containing the parameter for the explanatory variable m. The γ_m parameter is the same for every alternative and individual.
The utilities will be used in every model to estimate the parameters of that particular model.

4.2 Multinomial Logit Model

The electronic company would first of all want to know what the most preferred alternative of the visiting customer is. Providing the company of the most preferred alternative and preferences of almost every customer, could lead to, among other things, more efficient marketing strategies. It will also give the company an insight of the first alternative they need to show in the more personalized way of displaying the alternatives. For those reasons, I will first take a look at the *Multinomial Logit Model* (MNL) [McFadden et al., 1973]. Assuming that ϵ_{ij} are independent and type-I extreme value distributed and using the deterministic part (V_{ij}), will help to find the probability that alternative j is most preferred by individual i. This will be defined as following

$$Pr[y_{ij} = 1; \beta] = Pr[U_{ij} \geq \max\{U_{i1}, ..., U_{iJ}\}] = \frac{\exp(V_{ij})}{\sum_{l=1}^{J} \exp(V_{il})}. \tag{2}$$

To avoid identification problems, β_J will be set equal to zero, when estimating the parameters of the MNL model.

4.3 Rank-Ordered Logit Model

Because there is not only interest in the most preferred alternative of the respondents but in the complete ranking of the alternatives, the *Rank-Ordered Logit* (ROL) [Ahn et al., 2006, Beggs et al., 1981] model will be applied after the MNL model. In the ROL model it is assumed that respondents can rank all the alternatives. The ROL model is an extension of the MNL model, as the ROL model can be seen as a series of MNL models. The ROL model gives more information for estimating the parameters compared to the MNL model. Instead of only looking at the probability that an alternative is chosen, the ROL model also takes into account the rank of that alternative and calculates the probability that alternative j, that received rank k by individual i, is chosen.

Every respondent will be asked to provide a complete ranking of the alternatives. The response of individual i will from now on be denoted now by the vector $y_i = (y_{i1}, ..., y_{iJ})$. Introducing the complete ranking, gives us the ROL model

$$Pr[r_i; \beta] = Pr[U_{ir_{i1}} > U_{ir_{i2}} > ... > U_{ir_{iJ}}] = \prod_{j=1}^{J-1} \frac{\exp(V_{ir_{ij}})}{\sum_{l=1}^{J} \exp(V_{ir_{il}})}, \tag{3}$$

where r_{ij} denotes the alternative that received rank j by individual i. $V_{ir_{ij}}$ will now show in every column j the utility of the j-th ranked alternative. It is important to take r_{ij} into account now, because the estimated parameters by the ROL model, will show the influence of providing a particular ranking.

Because of the fact that the standard ROL model assumes that each respondent can rank all the alternatives, the parameter estimates can be biased. In the real world most of the time respondents cannot rank all the alternatives. We assume that individuals are able to rank the top k items. Equation 3 will now not sum up all the $J - 1$ ranks but only sum up to k. The rest of the $J - k$ alternatives could be ordered randomly, so the probability of observing one of those random orders should be included in the model. Including this ranking inability will lead to the following model

$$
\begin{aligned}
Pr[y_i|k;\beta] &= Pr[U_{ir_{i1}} > U_{ir_{i2}} > ... > U_{ir_{ik}} > \max\{U_{ir_{ik+1}}, ..., U_{ir_{iJ}}\}] \\
&= \left(\prod_{j=1}^{k} \frac{\exp(V_{ir_{ij}})}{\sum_{l=j}^{J} \exp(V_{ir_{il}})}\right) \frac{1}{(J-k)!}. \qquad (4)
\end{aligned}
$$

This model deals with the ranking inability of the respondents. The ROL model assumes that k is equal to the number of alternatives, therefore the respondents can rank all the alternatives. For the ROL model, the last term of equation 4 will not be important, but for the next model that will not be the case.

4.4 Latent-Class Rank-Ordered Logit Model

When the number of alternatives becomes large, it is often difficult for respondents to provide a complete ranking according to their preferences in a stated preference setting [Ophem et al., 1999]. This will lead to low level of informative value of the ROL model. First of all in equation 4 it is assumed that k is already known on beforehand. Most of the time this is not the case. Also when the k is introduced, the most simple case is to let k be equal for all individuals. This will lead to letting k be equal to the individual i who has the lowest ranking ability. To get better results from the data, it is important to let the model allow for heterogeneity. A model that deals with this ranking inability is the *Latent-Class Rank-Ordered* (LCROL) model [Fok et al., 2012].
In the LCROL model J latent classes ($k = 0, 1, .., J-1$) are introduced, where J is the total number of alternatives. Class k tells us that k alternatives can be ranked by respondent i. The LCROL model is described as follows

$$
Pr[y_i; \beta, p] = \sum_{k=0}^{J-1} p_k Pr[y_i|k;\beta]. \qquad (5)
$$

In the LCROL model the probability of belonging to class k (p_k) will also play an important role, when calculating the probability of the choice of an alternative by a respondent. p_k will tell us what the probability is that a respondent will belong to class k and how many alternatives he is able to rank. Because p_k is a probability, it will be restricted to be at least 0 and at most 1 and $\sum_{k=0}^{J-1} p_k = 1$. Next to this restriction, it is also important that $p_0 \neq 1$, because otherwise we will have a dataset including respondents who cannot rank at all.
As mentioned earlier, the last term of equation 4 is now important. For comparing the different classes, it is important to know if the complete ranking of the respondent is taken into account or just a part of it.

4.5 Parameter estimation

For estimation of the parameters, *Maximum Likelihood* (ML) estimation in combination with the BFGS method will be used. The following considered likelihood (6)

$$
\mathcal{L}(\beta, \theta) = \prod_{i=1}^{N} \sum_{k=0}^{J-1} p_k Pr[y_i|k;\beta] = \sum_{k=0}^{J-1} \frac{p_k}{(J-k)!} \Big[\prod_{l=1}^{k} \frac{\exp(V_{ir_{il}})}{\sum_{m=l}^{J} \exp(V_{ir_{im}})}\Big], \qquad (6)
$$

will lead to the log-likelihood (7)

$$\log(\mathcal{L}(\beta,\theta)) = \sum_{i=1}^{N} \log\Big\{\sum_{k=0}^{J-1} p_k \exp\Big[-\log((J-k)!) + \sum_{l=1}^{k}\Big(V_{ir_{il}}) - \log\sum_{m=l}^{J} e^{V_{ir_{im}}}\Big)\Big]\Big\}, \tag{7}$$

where p_k is equal to

$$p_k = \frac{\exp^{\theta_j}}{\sum_{l=0}^{J-1}\exp^{\theta_l}}. \tag{8}$$

The parameter p_k is a transformation of θ-parameters, because for maximizing the log-likelihood function the BFGS method will be considered. In this optimization method a way to circumvent the restriction $\sum_{k=0}^{J-1} p_k = 1$ and $0 \leq p_k \leq 1$, which means restricted optimization algorithms would have to be used, is to transform the θ parameters. By taking the exponent and dividing by the sum of all the transformed parameters, this restriction is imposed while maintaining unrestricted optimization. For identification p_{J-1} will be set equal to zero.

4.6 Comparison

To find out what the differences between the results of the two datasets are, I will estimate two MNL, ROL and LCROL models. One for the older dataset and one for the newer one. The explanatory variables will consist of platform intercepts, ownership and hours of gaming. The ownerships will be indicated by dummy variables with an 1 if a respondent owns a platform and a 0 otherwise. To be consistent, I will use the PC as the base alternative. Comparing the performances of the three different models for each dataset will lead to an answer to what extent a more personalized sorting method can be introduced. This will be done by investigating the probabilities of belonging to a class but also which parameters lead to better ranking abilities.
To find out what an effective way is to display and promote the alternatives, it is important to interpret the parameter estimates of both datasets and compare their results in the end. This will give a clear picture of the changes over the past few years.

5 Individual class probabilities

Next to the extension of using a more recent dataset it could be also interesting to find out if the ranking ability, represented by the model, is accurate enough when looking at the respondents' given ranking. Respondents who are familiar with the alternatives, can easily and accurately provide a complete, or at least a part of the ranking. If the respondent places an alternative they are familiar with at a specific rank, every alternative placed at a higher rank should probably be placed following his real preferences and not randomly. This familiarity of gaming platforms can be explained by ownership of the platforms but also by gaming hours. Higher gaming hours probably indicate that the individual is a heavy gamer and has knowledge about platforms and gaming in general.
To explain this idea better, you should try to imagine that a respondent owns for example a PlayStation. Giving the PlayStation a rank of four will mean that every gaming

platform placed at a rank below four should be some knowledge about. The alternatives placed rank five and six could be ranked randomly or not.
To find out if this way of thinking matches the class where the respondent will be placed in by the LCROL model, I will estimate the probabilities of belonging to a class for every respondent separately based on the respondents characteristics. Where θ is equal to a random number in equation 8, the new θ will is based on explanatory variables and estimated for every respondent separately. The probability will in this case be based on the individual characteristic of average playing hours per week and a constant. The probability will now be equal to

$$p_{ik} = \frac{\exp^{\theta_{ik}}}{\sum_{l=0}^{J-1} \exp^{\theta_{il}}}, \tag{9}$$

where $\theta_{ik} = x_i'\phi_k$ and $k = 0, 1, ..., J-1$. For identification p_{iJ-1} will be set equal to zero. The log-likelihood function that will be maximized from now on is $\log(\mathcal{L}(\beta, \phi))$. The LCROL model will be estimated and the $\hat{\phi}$ will be used to find the individual probabilities of belonging to a class.
Next to estimating the individual probabilities separately for each individual, it is important to see to what extent this individual probabilities match with the familiarity of the platforms. The precision of the individual probabilities will be checked by calculating the hit rates of predicted class[3] and the lowest[4] given rank of an owned platform. Obtaining a high hit rate will indicate that individual i's ranking ability is formed by his individual characteristics and familiarity with the platforms.

6 Results

In the methodology part, the LCROL model shows it can estimate to what extent respondents can rank the alternatives accurately. The ranking ability and preferences of the respondents are measured by the LCROL model and the results will be used to develop a more personalized sorting order and to improve the marketing content on the website page. The ranking ability will help determining how the customers can be segmented and the preferences of the customers will help in explaining how their needs can be assessed. These findings need to be implemented in practice. Using the ranking abilities and preferences of the customers the company could segment the customers into different groups, sort products in a more personalized way and find new ways to promote products, when displaying them to the customers.
In this section, the results of the models based on the gaming data will be shown. The LCROL model is applied in a more practical way for electronic companies. This means that the results of the LCROL model will be used mainly to explain. Next to analyzing the parameter estimates and the individual ranking probabilities, a more managerial advice will be given in the end.

[3] The predicted class is determined by using the highest individual probability of belonging to a class among the different classes.
[4] Lowest means that rank 5 can be seen as the lowest rank and rank 1 by highest rank.

6.1 Model comparison

6.1.1 Gaming data 2007

Looking at the MNL and ROL models of table 3 we see a clear difference in the parameter estimates of the two models. Due to the differences in the parameter estimates, the LCROL model will to be estimated. Including the classes of the ranking abilities of respondents in the model, gives us the LCROL model as shown in the fourth column of table 3. The results match the results presented in [Fok et al., 2012] almost always up to two decimal places. The LCROL model shows that 23% of the respondents cannot rank the alternatives at all. Performing an *Likelihood-Ratio* (LR) test for $p_1 = 0$ equals 15.03. The class containing respondents who cannot rank at all, cannot be neglected. This would mean that the MNL and ROL model would perform worse when comparing it with the LCROL model. The MNL and ROL say that $p_0 = 0$. Comparing the MNL ($p_1 = 1$) and ROL model ($p_5 = 1$) with the LCROL model, the restriction that those two models perform better than the LCROL model can be rejected. Knowing that p_0 needs to be included in the model and p_2,p_3 and p_4 are relatively small classes, an new restriction can be applied on the original model, namely $p_2 = p_3 = p_4 = 0$. The last column of table 3 represents the restricted model. Performing a LR test gives a LR-statistic of 1.85 and hence the restriction cannot be rejected.

Table 3: Parameter estimates gaming dataset 2007

Variable	MNL	ROL	LCROL	LCROL^a	LCROL^b
intercept					
XBox	0.91	1.39	1.53	1.41	1.48
	(0.51)	(0.29)	(0.51)	(0.43)	(0.52)
PlayStation	0.58	0.94	1.11	0.99	1.03
	(0.46)	(0.27)	(0.47)	(0.41)	(0.46)
PlayStation Portable	-0.03	0.80	0.44	0.70	0.50
	(0.63)	(0.29)	(0.52)	(0.56)	(0.54)
GameCube	0.51	0.05	-3.50	-0.75	-2.27
	(0.61)	(0.30)	(1.61)	(0.55)	(1.08)
GameBoy	-1.47	0.09	-2.71	-0.66	-1.66
	(0.92)	(0.29)	(1.41)	(0.80)	(0.93)
Hours spent on gaming					
XBox	-0.10	-0.17	-0.14	-0.13	-0.13
	(0.07)	(0.05)	(0.06)	(0.06)	(0.06)
PlayStation	-0.11	-0.13	-0.11	-0.11	-0.11
	(0.07)	(0.04)	(0.06)	(0.06)	(0.06)
PlayStation Portable	-0.10	-0.23	-0.36	-0.33	-0.39
	(0.12)	(0.05)	(0.12)	(0.14)	(0.13)
GameCube	-0.39	-0.19	-0.01	-0.21	-0.14
	(0.24)	(0.05)	(0.15)	(0.11)	(0.16)
GameBoy	-0.05	-0.24	-0.23	-0.33	-0.32
	(0.17)	(0.05)	(0.15)	(0.11)	(0.15)
Platform Ownership	1.78	0.96	1.72	1.47	1.72
	(0.39)	(0.19)	(0.37)	(0.29)	(0.35)
p_0			0.23		0.21
p_1	1		0.21	0.37	0.27
p_2			0.06	0.00	
p_3			0.07	0.06	
p_4			0.00	0.06	
p_5		1	0.43	0.51	0.52
LR-statistic^c	109.60	35.38	-	15.03	1.54

This table contains the results of the MNL, ROL and LCROL model using the dataset of 2007. The estimates of the variables and the probabilities are presented in the table. In the last row the LR-statistic is showed.

^{a} Latent-Class Rank-Ordered Logit model with p_0=0

^{b} Latent-Class Rank-Ordered Logit model with p_2=p_3=p_4=0

^{c} = LR-statistic to test against the LCROL model in the fourth column.

6.1.2 Gaming data 2017

For the gaming data of 2017 three models are estimated. Looking at the MNL and ROL model of table 4, the parameter estimates differ a lot. This leads to the LCROL model in the fourth column of table 4. The most remarkable of the results is probably the probabilities of ranking abilities. 32% of the respondents cannot rank at all and 30% of the respondents can rank four alternatives out of six. The remarkable is here the big difference in ranking ability by the respondents, either someone has a lot of knowledge of the platforms and can almost rank everything, or he is completely unfamiliar and gives a random ranking. Performing a LR-test for $p_1 = 1$ (MNL model) and $p_5 = 1$ (ROL model) gives LR-statistics equal to 86.55 and 64.58 respectively. The LCROL model performs better comparing to both the MNL and ROL model and both p_1 and p_5 cannot be neglected.

Table 4: Parameter estimates gaming dataset 2017

Variable	MNL	ROL	LCROL
intercept			
Wii	0.55	-0.24	0.10
	(0.49)	(0.20)	(0.53)
PlayStation	1.12	0.31	0.54
	(0.42)	(0.19)	(0.44)
Xbox	1.03	-0.24	-0.83
	(0.56)	(0.20)	(0.53)
smartphone	1.05	-0.41	0.49
	(0.39)	(0.20)	(0.40)
Portable GameBoy	1.33	-0.20	-0.50
	(0.60)	(0.20)	(0.53)
Hours spent on gaming			
Wii	-0.08	-0.02	-0.29
	(0.07)	(0.02)	(0.16)
PlayStation	-0.02	0.01	0.03
	(0.04)	(0.02)	(0.05)
Xbox	-0.12	-0.01	0.05
	(0.09)	(0.02)	(0.05)
smartphone	-0.16	-0.05	-0.18
	(0.06)	(0.02)	(0.07)
Portable GameBoy	-0.34	-0.02	-0.03
	(0.17)	(0.02)	(0.06)
Platform Ownership	1.87	1.17	4.02
	(0.29)	(0.14)	(1.14)
p_0			0.32
p_1	1		0.17
p_2			0.16
p_3			0.03
p_4			0.30
p_5		1	0.02
LR-statistic^a	86.55	64.58	-

This table contains the results of the MNL, ROL and LCROL model using the dataset of 2017. The estimates of the variables and the probabilities are presented in the table. In the last row the LR-statistic is showed.
LR-statistic^a = LR statistic to test against the LCROL model in fourth column.

6.1.3 Comparison data

For both datasets it is notable that the LCROL models performs better than the MNL and ROL, because there are still a lot of respondents who cannot rank at all but still rank more than one and less than five alternatives. The differences of the two datasets can be found in to what extent the LCROL model performs better than the MNL and ROL model. In both datasets is the LCROL model used as the unrestricted model for a LR-test. In the dataset of 2007 the MNL performs worse relative to the LCROL model when comparing it with how the MNL model performs relative to the LCROL model in the dataset of 2017. The difference between the MNL and LCROL model is

bigger in 2007 than in 2017. On the other hand the difference between the ROL and LCROL model in the dataset of 2007 is smaller than in 2017. This could be due to the estimated probabilities of the LCROL model in both datasets. In 2007 the biggest part of the respondents could rank all the alternatives. In 2017 most of the respondents could not rank at all, but if they could rank, they were able to rank at most four alternatives alternatives.
The ranking inability of the of the respondents in 2017 can maybe be explained by the familiarity with the gaming platforms. The ownership of products in 2007 can be seen in table 1. In table 2 it can be seen that less respondents owned the gaming platforms except for the smartphone and Wii (successor of the GameCube) in 2017.

6.2 Parameter interpretation

Next to comparing the different models by likelihoods to find out to what extent respondents can rank the gaming platforms, the estimation results of the parameters will also be interesting for the electronic company. The parameter estimates will help the company understand their customers better and will be usable for improving the marketing content and promotional advertisements.

6.2.1 Gaming data 2007

Looking at the LCROL models in table 3, the following insights are found. The ownership of a platform is the first insight, because it has a positive effect on the preference of that platform. The positive effect could also be due to reverse causality. Not only does ownership influence the rank of an alternative but also if someone likes the alternative leads to ownership [Fok et al., 2012]. Also the compatibility of games could lead to the positive effect of ownership on the preference of a platform. The positive effect could also be due to the backward compatibility of games. When someone owns a lot of PlayStation 2 games he would probably buy the PlayStation 3 instead of a Xbox so he can still play his games on the new PlayStation. Because there was no distinction made in different version of gaming platforms this influence could not be estimated or observed. Secondly, the more hours someone play games, the more he prefers a PC. The last insight is that in 2007 the PC was the most preferred gaming platform by on average 43% of the respondents.

6.2.2 Gaming data 2017

The LCROL model in table 4 shows the results of the parameter estimates based on the dataset of 2017. The ownership of a platform is the first noticeable insight, because it has a huge positive effect on the preference of that platform. Owning a product gives a high preference for that product.
Secondly, the more hours someone play games, the more he prefers a PC, Xbox or PlayStation. Finally in 2017 the smartphone was the most preferred gaming platform by on average 35% of the respondents, followed by the PlayStation and PC respectively.

6.2.3 Comparison data

Comparing the LCROL model parameter estimates of the two different datasets, gives an insight that over the years the effect of gaming hours on the preference of the PC has changed a little bit. The results show us that the influence of the introduction of the smartphone is huge. The smartphone is now seen as a gaming platform and is the most preferred gaming platform in the survey of 2017, followed by the PlayStation. This indicates that the popularity of the PC has dropped in combination with a growth of the PlayStation and smartphone.

6.3 Individual class estimates

In this part of the results section, the results of how accurate the individual probabilities p_{ik} are estimated will be showed. This is important to find out to what extent this individual probabilities match with the familiarity of the platforms.
The individual probabilities of respondents for the dataset from 2007 and 2017 are shown in the appendix in table 8 and table 9 respectively. Next to the individual probabilities the influence of the average of gaming hours per week on the individual probabilities is showed in table 7 for 2007 and 2017.

6.3.1 Gaming Data 2007

When looking at table 8 and finding for every respondent the maximum probability indicates to which class every respondent belongs. The LCROL model predicts that the respondents are separated into the first class, p_0, and the last class, p_5 when considering individual probabilities. So individuals can either rank all the alternatives (69.2% of the respondents) or none (30.8% of the respondents). When considering for every individual the probabilities of belonging to a class, the average value of the largest probability for belonging to a segment is 0.43[5]. If this probability would be high, the model would have a more certain view of which group the individual belongs to. In this case 0.44 is not a very high probability, what indicates that the way how the individual probabilities are created is not accurate enough. The precision of the individual probabilities is also checked by computing the hit rate as explained in section 5. The inaccuracy of the individual probabilities can also be seen when computing the hitrate. The hitrate equals 17.6% which is very low. I can conclude that the individual probabilities are not measured accurate enough and the familiarity of platforms cannot match my thinking of ranking abilities, as explained in the example of section 5.
Even though the individual probabilities do not seem to be accurate, it is still possible to get some insights of the characteristics of individuals belonging to segment 0 or segment 5. The individuals are separated into two groups in the dataset and the characteristics showed in the first column of table 5 are investigated. From table 5, I can assume that respondents belonging to segment 5 are heavy gamers, due to the overall average of 5.6 hours. This conclusion agrees with the thoughts of heavy gamers can rank more easily all the alternatives. Also considering ownership the results are as I expected. Next to gaming hours, ownership leads to being more able to rank the alternatives too and this is clearly visible if you compare the ownership per alternative for the two segments.
Table 5 also shows some interesting results on the female share in the segments and what the type of games are played in both segments.

[5] Sum of all maximum probabilities for every individual divided by the total number of individuals

Table 5: Characteristics of the segments for Gaming data 2007

	Segment 0	Segment 5
Hours spent on gaming	0.04	5.6
Ownership Xbox	0.00%	19.00%
Ownership PlayStation	0.04%	44.44%
Ownership GameCube	0.00%	12.70%
Ownership PSPortable	0.04%	12.70%
Ownership Gameboy	0.04%	17.46%
Ownership PC	75.00%	93.65%
Female	53.57%	12.70%
Action Games	46.43%	85.71%
Strategy Games	64.29%	65.08%
Puzzle Games	32.14%	17.46%
Sport Games	64.29%	69.84%

This table contains characteristics of the two segments wherein the individuals are separated based on the maximum individual probabilities over the segments. The results, except for hours spent on gaming, are showed as a percentage of the individuals belonging to that segment. The last four rows about games, show the percentage of respondents in that particular segment playing that kind of games. The hours spent on gaming is equal to the average gaming hours of the individuals belonging to that particular segment.

6.3.2 Gaming Data 2017

In the dataset of 2017, the LCROL model predicts that the respondents are separated into the first class, p_0, the second class p_1 and the fifth class, p_4. So individuals can either rank none (18.0% of the respondents), one (32.8% of the respondents) or four of the alternatives (48.4% of the respondents). Corresponding to the estimated individual probabilities, there was one person who could rank all the alternatives. He is not included in table 6. Like in the dataset of 2007 the individual probabilities are not accurate enough. The average value of the largest probability for every individual belonging to a segment is 0.40.

In table 6 my expectations are not met. I assumed that respondents belonging to a higher segment game more hours on average and that the proportion. It seems that respondents who belong to segment 0 and 1 are more heavier gamers than respondents belonging to segment 4. Next to my assumptions about gaming hours, I also expected that the share of women will decrease when the segment rises, because of their low average of gaming hours as showed in section 3.2. The last outstanding insight is the one of ownership. I would expect that when someone owns a platform would belong to a higher segment. Looking at the three segments of table 6 this is not the case. The prediction of belonging to the segment based on an owned product, that has the lowest rank is not accurate. The hitrate equals 10.7% which is very low. This indicates again that the way the probabilities are formed are inaccurate.

Table 6: Characteristics of the segments for Gaming data 2017

	Segment 0	Segment 1	Segment 4
Hours spent on gaming	4.1	7.0	3.8
Ownership Xbox	13.64%	10%	16.95%
Ownership PlayStation	27.27%	35%	33.90%
Ownership Wii	27.27%	20%	27.12%
Ownership Portable Gameboy	13.64%	7.5%	11.86%
Ownership smartphone	100%	100%	100%
Ownership PC	50%	47.5%	67.80%
Female	77.27%	52.5%	57.63%
IOS	50%	70%	57.63%
Action Games	22.72%	25%	40.68%
Strategy Games	18.18%	27.5%	15.25%
Puzzle Games	36.36%	22.5%	22.03%
Sport Games	22.72%	25%	22.03%

This table contains characteristics of the three segments wherein the individuals are separated based on the maximum individual probabilities over the segments. The results, except for hours spent on gaming, are showed as a percentage of the individuals belonging to that segment. The last four rows about games, show the percentage of respondents in that particular segment playing that kind of games. The hours spent on gaming is equal to the average gaming hours of the individuals belonging to that particular segment.

6.4 Managerial advice

The managerial advice for the company will be based on the question;

Keeping in mind the preferences and ranking ability of customers, what will be an effective way to improve marketing content to assess the needs of the customer?

To answer the question, I looked at two website pages and one part of a website page of an electronic company to find if some improvements could be made. The website pages can be found in the appendix as figure 2, 3 and 4. Looking at figure 3, I see the company promotes a lot for PlayStation, so there is some space for improvement in the content marketing.
For some insights I will recommend how the pages shown in 2 and figure 3 need to be organized based on the banners and layout of each page. Next to that I will give a general advice about the ordering of the aspects in the filter box in figure 4 and the sorting of products on each page, based on the ranking ability of the customer obtained by the results of the LCROL model and individual probabilities for the dataset of 2017.

6.4.1 Collecting data

First of all it is important for the company to collect the right data to use the LCROL model. For the LCROL model, the ranking of alternatives is needed. It is hard to ask every customer to provide a ranking of every offered product by the company. Next to the lack of information on rankings of products, there are two types of customers. The "known" and "unknown" customers. The known customers can be recognized by their IP-addresses or because they are logged in on their personal shopping accounts of the website. The "unknown" customers are hard but not impossible to recognize. You

need them to let them selves identify on your website, by letting them accept cookies or let them register. The company could manage to let customers register by offering them discounts, free delivery, more benefits if they register. Thus, if the company has some information of the customer, it could send some of them a questionnaire about ranking similar products of the company or try to predict a ranking of a customer by analyzing its clicks on various products.
Still there will be customers visiting the website which stay unknown. For those customers more general sorting orders, like "most bought" or "best rated" should be displayed, until more information of them is available.

6.4.2 Insight 1: *"Heavy gamers prefer the PC, PlayStation and Xbox"*

According to the fourth column of table 4, Respondents who spend more time on gaming are more likely to prefer the PC, PlayStation and Xbox. Knowing that your customer is a heavy gamer could lead to conclusions to display more games for those platforms on the first displayed page. When heavy gamers go to the *Games* page, it will be efficient to show games of those three platforms on the first *Games* page. According to LCROL model, the preferences of the respondents could help improving the marketing content of the page showed in figure 3. The top banner, showed on this page, could be more efficiently used. Instead of referring to the entertainment of the month, this could be the banner that refers to the games that are selected for the customer. Knowing the customer is a heavy gamer and which platform he owns, it is useless to show the consoles underneath, he will probably already own a couple of the consoles. It would be more useful to show gaming accessories. The two small banners next to the big one, could be filled with new releases of games or maybe another gaming console. Continuing to the games page as seen in figure 2, immediately all kind games of the PlayStation platform are showed. Knowing your customer is a heavy gamer, it would be more efficient to show more PC games combined with a couple of PlayStation games, combined with the knowledge what type of games the customer likes.
This insight is more focused on the preferences of the customers than on the ranking ability, however it is still possible to include the ranking ability. Knowing you are dealing with a heavy gamer, who belongs to a specific segment and what kind of platforms he owns, it would be good not to show at first sight more types of alternatives than he can rank and only show games for the platforms he owns.

6.4.3 Insight 2: *"Females play more puzzle games on their smartphone"*

In the dataset of 2017 59.8% of the respondents are woman and 40.2% are men. 49.3% of the women prefers the smartphone most followed by 10.0% preferring the playstation most. This compared with 14% of the men preferring smartphone most and 44.9% preferring the PlayStation most.
Given an average of 2.3 hours of gaming per week, means they play on average 20 minutes per day. A session of a game on for example a PlayStation or Xbox last on average longer than 20 minutes per game [Tarng et al., 2008]. Due to their average gaming hours, I can conclude they probably game on their smartphone in particular. Knowing your customer is a woman with little interest for gaming, it is probably better not to show a PlayStation game in the banner but a new puzzle game from the Applestore or Android-Store. For this the company can try to make an agreement with the game providers.
Because women are not heavy gamers, maybe promoting more "female games" like

dance and sing games, next to puzzles, including the required accessories will be more effective. This could lead to more enthusiasms for gaming and let the female consider to buy games including console, accessories etc. So when knowing your customer is a not-heavy female gamer, focus on more "female games" and adapt your banners and content to the females preferences.

6.4.4 Insight 3: *"Ownership of a platform is a very important factor for the next purchase of a product for that type of platform"*

Knowing which platform the customers owns, the website page needs to be adjusted to that platform. Still it is important to promote somewhere other gaming platforms, but keep that promotion on a low level. When someone owns a platform, he will not buy it twice, so do not promote that console. But when there is in the near future a new release coming of that platform, make sure you promote it huge and showy. The website of Bol.com does this very well, as shown in figure 5. It let you already pre-order upcoming games. The website used as example in the figures 2, 3 and 4 also let you pre-order, but does not promote this visibly. This could be an efficient improvement in the marketing content if you information of the ownership of platforms for that specific customer.

6.4.5 Insight 4: *"The PlayStation is the most preferred gaming platform by men"*

As found in the dataset of 2017 and mentioned in insight 2, men mostly prefer the PlayStation as gaming platform. Looking back at figure 2 and 3, this would almost be perfect displayed pages for a man. Men prefer the PlayStation most. Knowing this, the big banner should be focused on a PlayStation game or selection of games. Adjust your content to the man by promoting PlayStation products by the most noticeable parts of the page, like the banner. The rest of the content of the page on figure 3 could be in the same way organized as for heavy gamers, only the focus is now on the products of the PlayStation. The page of figure 2 is well organized, only the content could be more adapted to the preferred type of games of the customers.

6.4.6 Insight 5: *"Include the ranking capabilities of customers in the filter box and in the first displayed sorting"*

For this insight the company should base a personalized sorting order on the results of the LCROL model. The LCROL model can tell to what extent the customers can rank the alternatives. If it is known that an individual can only rank his top three alternatives, the company should visibly display the first three alternatives in the sorting order on their website page. Underneath those alternatives, there could be a banner placed or some other important marketing content.
The ranking ability of customers could also be implemented in advertisements of the company on other website pages. As seen in figure 6, due to cookies I get my search history displayed from the company website in an advertisement on another website. It would be more efficient, if there is knowledge of my ranking ability, to display the amount of products I can rank. In figure 6 I see six products combined with their prices. My ranking ability, in case of those products, would probably be equal to two. Because many products are displayed in the advertisement, the advertisement does not attract me. Implementing the ranking abilities in the advertisements on other websites pages, would be more beneficial for the company.

For the filter box in figure 4, it is important to sort the options in the way the customers would rank them from the most preferred to the least preferred. Heavy gamers and customers owning a platform could provide a more complete ranking of the platforms. The company should at least sort the filter box in the way of showing the most preferred alternative on top. This is also what should be done in the first displayed sorting order of products on various website pages. From the questions of how important specific features of the gaming platform are, price of soft- and hardware shared together the first place. It could be a good idea to show immediately after the console part in the filter box the price range.

7 Conclusion

Using the LCROL model to capture the ranking inabilities of the customers helps in finding an accurate personalized sorting order of products for the website. The comparison of the models shows clearly that when one needs to deal with ranked data with unobserved heterogeneity in ranking capabilities, the LCROL model performs better than the ROL model. As questioned in the introduction to what extent a more personalized sorting method can be introduced, the LCROL model gives the answer. Based on the results of the LCROL model, the company can predict to which segment the customers belong, and fit the sorting order of products shown as first on the website page to the ranking ability of the customers. To get accurate predictions, the data must be of good quality. I think this is the hardest part of observing the ranking abilities. The dataset from 2017 was biased towards segment 0. Using those results, would answer the question that it is very hard to introduce a more personalized sorting method based on ranking abilities. Getting more accurate knowledge on the ranking of various products, could solve this problem.

Obtaining good quality of data, will improve the results of the LCROL model on ranking abilities, but also on the results of preferences of customers. Those results can, next to a more personalized sorting order, be implemented in displayed advertisements of the company on other websites due to cookies. Based on the preferences of customers, the second question of the introduction can be answered. Keeping in mind this ranking ability of customers, an effective way to promote products and have an useful marketing content is to adapt it to the knowledge of the preferences of the customers. Knowledge of the customers will let the company access the needs of them. It is showed by the LCROL model in table 4 that heavy gamers prefer a PC, PlayStation and Xbox. Adapt the content of the marketing for heavy gamers to more PC-, PlayStation- and Xbox-orientated products. Besides that ownership of a platform has a positive effect on the ranking ability but will also be an important factor for the next purchase of a product usable for the same type of owned platform. Let the promotion of products of the same type of platform be the most important ones. But also letting customers know that they can pre-order products of those platforms would be beneficial. In addition, let man be the target for PlayStation marketing and females for puzzle type games for the smartphone. Try to get a collaboration with Apple and Android for offering products from their stores. Finally adjust the filter box and the sorting of the first displayed page of products to the ranking ability of the customer, combined with the knowledge as stated above.

Technology is improving every year, so try to keep up as a company with the developments of new products and be the first one offering them. For example, *Virtual Reality* (VR) and *Augmented Reality* (AR) are very upcoming now. A lot of companies already

use it, but more and more games are implementing VR and AR. For example VR is already available for PlayStation. Considering AR games, Pokemon GO was one of the most played games worldwide in 2016 [Leswing, 2017].
Keeping up with the developments in technology, obtaining good quality of data will let the LCROL model provide an accurate view on the ranking ability and preferences of customers. Displaying a more personalized sorting order on the website page will be beneficial for the company. Next to this, the personalized sorting order will lead to improvements in customer satisfaction and optimize the customer journeys in the end.

References

[Ahn et al., 2006] Ahn, J., Lee, J., Lee, J.-D., and Kim, T.-Y. (2006). An analysis of consumer preferences among wireless lan and mobile internet services. ETRI journal, 28(2):205–215.

[Beggs et al., 1981] Beggs, S., Cardell, S., and Hausman, J. (1981). Assessing the potential demand for electric cars. Journal of econometrics, 17(1):1–19.

[Chapman and Staelin, 1982] Chapman, R. G. and Staelin, R. (1982). Exploiting rank ordered choice set data within the stochastic utility model. Journal of marketing research, pages 288–301.

[Cronin Jr and Taylor, 1992] Cronin Jr, J. J. and Taylor, S. A. (1992). Measuring service quality: a reexamination and extension. The journal of marketing, pages 55–68.

[Fok et al., 2012] Fok, D., Paap, R., and Van Dijk, B. (2012). A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. Journal of applied econometrics, 27(5):831–846.

[Gibbert et al., 2002] Gibbert, M., Leibold, M., and Probst, G. (2002). Five styles of customer knowledge management, and how smart companies use them to create value. European Management Journal, 20(5):459–469.

[Hansemark and Albinsson, 2004] Hansemark, O. C. and Albinsson, M. (2004). Customer satisfaction and retention: the experiences of individual employees. Managing Service Quality: An International Journal, 14(1):40–57.

[Jacobson and Aaker, 1987] Jacobson, R. and Aaker, D. A. (1987). The strategic role of product quality. The Journal of Marketing, pages 31–44.

[Kalyanaraman and Sundar, 2006] Kalyanaraman, S. and Sundar, S. S. (2006). The psychological appeal of personalized content in web portals: Does customization affect attitudes and behavior? Journal of Communication, 56:110–132.

[Leswing, 2017] Leswing, K. (2017). Pokmon go was the most downloaded iphone app worldwide in 2016, apple says.

[McFadden et al., 1973] McFadden, D. et al. (1973). *Conditional logit analysis of qualitative choice behavior*.

[Morgan, 2013] Morgan, C.-A. (2013). Managing the experience of your customers.

[Nielson Company, 2014] Nielson Company (2014). Multi-platform gaming: For the win!

[Ophem et al., 1999] Ophem, H. V., Stam, P., and Praag, B. V. (1999). Multichoice logit: Modeling incomplete preference rankings of classical concerts. Journal of Business & Economic Statistics, 17(1):117–128.

[Parfitt, 2007] Parfitt, B. (2007). Nintendo ends gamecube support.

[Sorman, 2014] Sorman, A. (2014). The best way to map the customer journey: Take a walk in their shoes.

[Tarng et al., 2008] Tarng, P.-Y., Chen, K.-T., and Huang, P. (2008). An analysis of wow players' game hours. In Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games, pages 47–52. ACM.

[Van Der Wiele et al., 2002] Van Der Wiele, T., Boselie, P., and Hesselink, M. (2002). Empirical evidence for the relationship between customer satisfaction and business performance. Managing Service Quality: An International Journal, 12(3):184–193.

[Wedel and Kamakura, 2012] Wedel, M. and Kamakura, W. A. (2012). Market segmentation: Conceptual and methodological foundations, volume 8. Springer Science & Business Media.

8 Appendix

Table 7: Parameter estimates Gaming dataset 2017

Variable	2007	2017
intercept		
p_0	0.20	0.76
p_1	-0.93	15.00
p_2	0.09	2.00
p_3	-1.74	4.33
p_4	-2.87	-3.13
Hours of gaming		
p_0	-0.24	-0.01
p_1	0.01	-0.69
p_2	-0.88	-0.26
p_3	-0.56	-0.16
p_4	-0.06	0.07

This table contains the results of LCROL for estimating the individual probabilities using the dataset of 2007 and 2017. The estimates of the variables which are used tho calculate the individual probabilities are presented in the table.

Table 8: Individual probabilities gaming dataset 2007

	p_0	p_1	p_2	p_3	p_4	p_5
1	0.31	0.16	0.08	0.02	0.02	0.41
2	0.27	0.19	0.03	0.01	0.02	0.48
3	0.24	0.21	0.02	0.01	0.02	0.51
4	0.16	0.23	0	0	0.02	0.57
5	0.31	0.16	0.08	0.02	0.02	0.41
6	0.02	0.29	0	0	0.02	0.67
7	0.13	0.25	0	0	0.02	0.59
8	0.31	0.1	0.28	0.04	0.01	0.25
9	0.31	0.1	0.28	0.04	0.01	0.25
10	0.11	0.25	0	0	0.02	0.61
11	0.11	0.25	0	0	0.02	0.61
12	0.28	0.19	0.04	0.02	0.02	0.46
13	0.31	0.1	0.28	0.04	0.01	0.25
14	0.24	0.21	0.02	0.01	0.02	0.51
15	0.24	0.21	0.02	0.01	0.02	0.51
16	0.2	0.22	0.01	0.01	0.02	0.54
17	0.31	0.1	0.28	0.04	0.01	0.25
18	0.31	0.1	0.28	0.04	0.01	0.25
19	0.31	0.1	0.28	0.04	0.01	0.25
20	0.31	0.1	0.28	0.04	0.01	0.25
21	0.31	0.1	0.28	0.04	0.01	0.25
22	0.28	0.19	0.04	0.02	0.02	0.46
23	0.32	0.13	0.15	0.03	0.02	0.34
24	0.32	0.13	0.15	0.03	0.02	0.34
25	0.31	0.16	0.08	0.02	0.02	0.41
26	0.31	0.1	0.28	0.04	0.01	0.25
27	0.31	0.16	0.08	0.02	0.02	0.41
28	0.02	0.29	0	0	0.02	0.67
29	0.31	0.16	0.08	0.02	0.02	0.41
30	0.07	0.27	0	0	0.02	0.64
31	0.32	0.13	0.15	0.03	0.02	0.34
32	0.31	0.1	0.28	0.04	0.01	0.25
33	0.32	0.13	0.15	0.03	0.02	0.34
34	0.31	0.1	0.28	0.04	0.01	0.25
35	0.31	0.1	0.28	0.04	0.01	0.25
36	0.31	0.16	0.08	0.02	0.02	0.41
37	0.07	0.27	0	0	0.02	0.64
38	0.13	0.25	0	0	0.02	0.59
39	0.31	0.16	0.08	0.02	0.02	0.41
40	0.31	0.16	0.08	0.02	0.02	0.41
41	0.31	0.1	0.28	0.04	0.01	0.25
42	0.31	0.1	0.28	0.04	0.01	0.25
43	0.32	0.11	0.24	0.04	0.02	0.27
44	0.13	0.25	0	0	0.02	0.59
45	0.2	0.22	0.01	0.01	0.02	0.54
46	0.32	0.13	0.15	0.03	0.02	0.34
47	0.32	0.13	0.15	0.03	0.02	0.34
48	0.24	0.21	0.02	0.01	0.02	0.51
49	0.32	0.13	0.15	0.03	0.02	0.34
50	0.13	0.25	0	0	0.02	0.59
51	0.31	0.1	0.28	0.04	0.01	0.25
52	0.31	0.1	0.28	0.04	0.01	0.25
53	0.13	0.25	0	0	0.02	0.59
54	0.2	0.22	0.01	0.01	0.02	0.54
55	0.02	0.29	0	0	0.02	0.67
56	0.2	0.22	0.01	0.01	0.02	0.54
57	0.31	0.1	0.28	0.04	0.01	0.25
58	0.31	0.16	0.08	0.02	0.02	0.41
59	0.31	0.1	0.28	0.04	0.01	0.25
60	0.31	0.1	0.28	0.04	0.01	0.25
61	0.32	0.13	0.15	0.03	0.02	0.34
62	0.32	0.12	0.21	0.04	0.02	0.3
63	0.32	0.13	0.15	0.03	0.02	0.34
64	0.11	0.25	0	0	0.02	0.61
65	0.02	0.29	0	0	0.02	0.67
66	0.31	0.1	0.28	0.04	0.01	0.25
67	0.31	0.1	0.28	0.04	0.01	0.25
68	0.2	0.22	0.01	0.01	0.02	0.54
69	0.32	0.13	0.15	0.03	0.02	0.34
70	0.32	0.12	0.21	0.04	0.02	0.3
71	0.11	0.25	0	0	0.02	0.61
72	0.32	0.13	0.15	0.03	0.02	0.34
73	0.31	0.1	0.28	0.04	0.01	0.25
74	0.02	0.29	0	0	0.02	0.67
75	0.31	0.1	0.28	0.04	0.01	0.25
76	0.24	0.21	0.02	0.01	0.02	0.51
77	0.32	0.13	0.15	0.03	0.02	0.34
78	0	0.32	0	0	0.01	0.67
79	0.32	0.13	0.15	0.03	0.02	0.34
80	0.02	0.29	0	0	0.02	0.67
81	0.07	0.27	0	0	0.02	0.64
82	0.32	0.13	0.15	0.03	0.02	0.34
83	0.31	0.16	0.08	0.02	0.02	0.41
84	0.31	0.16	0.08	0.02	0.02	0.41
85	0.07	0.27	0	0	0.02	0.64
86	0.11	0.25	0	0	0.02	0.61
87	0.07	0.27	0	0	0.02	0.64
88	0.07	0.27	0	0	0.02	0.64
89	0.31	0.1	0.28	0.04	0.01	0.25
90	0.31	0.1	0.28	0.04	0.01	0.25
91	0.280	0.19	0.04	0.02	0.02	0.46

This table contains the probabilities of the segments for every respondent separately. In the first column, the respondent number is shown. The highest probability for every individual indicates to which segment the individual belongs.

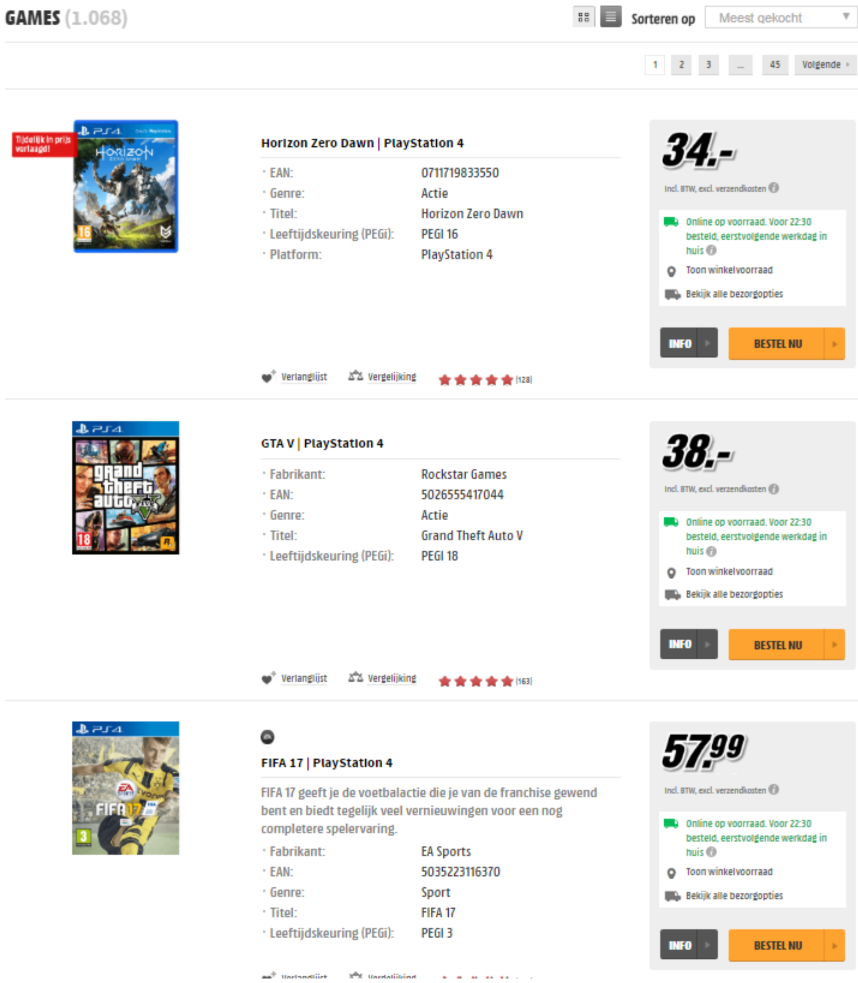
Table 9: Individual probabilities gaming dataset 2017

	p_0	p_1	p_2	p_3	p_4	p_5
1	0.3	0.13	0.16	0	0.37	0.03
2	0.25	0.22	0.13	0	0.36	0.02
3	0.3	0.13	0.16	0	0.37	0.03
4	0.37	0.04	0.2	0	0.34	0.04
5	0.48	0	0.27	0	0.17	0.08
6	0.3	0.13	0.16	0	0.37	0.03
7	0.3	0.13	0.16	0	0.37	0.03
8	0.3	0.13	0.16	0	0.37	0.03
9	0.25	0.22	0.13	0	0.36	0.02
10	0.14	0.49	0.07	0	0.28	0.01
11	0.34	0.08	0.18	0	0.36	0.04
12	0.25	0.22	0.13	0	0.36	0.02
13	0.2	0.35	0.1	0	0.33	0.02
14	0.34	0.08	0.18	0	0.36	0.04
15	0.14	0.49	0.07	0	0.28	0.01
16	0.42	0.01	0.23	0	0.28	0.05
17	0.42	0.01	0.23	0	0.28	0.05
18	0.3	0.13	0.16	0	0.37	0.03
19	0.09	0.64	0.05	0	0.21	0.01
20	0.03	0.85	0.02	0	0.1	0
21	0.3	0.13	0.16	0	0.37	0.03
22	0.3	0.13	0.16	0	0.37	0.03
23	0.2	0.35	0.1	0	0.33	0.02
24	0.34	0.08	0.18	0	0.36	0.04
25	0.14	0.49	0.07	0	0.28	0.01
26	0.42	0.01	0.23	0	0.28	0.05
27	0.34	0.08	0.18	0	0.36	0.04
28	0.2	0.35	0.1	0	0.33	0.02
29	0.09	0.64	0.05	0	0.21	0.01
30	0.25	0.22	0.13	0	0.36	0.02
31	0.3	0.13	0.16	0	0.37	0.03
32	0.4	0	0.25	0	0	0.35
33	0.44	0.01	0.24	0	0.25	0.06
34	0.37	0.04	0.2	0	0.34	0.04
35	0.2	0.35	0.1	0	0.33	0.02
36	0.06	0.76	0.03	0	0.15	0
37	0.25	0.22	0.13	0	0.36	0.02
38	0.4	0.02	0.22	0	0.31	0.05
39	0.34	0.08	0.18	0	0.36	0.04
40	0.3	0.13	0.16	0	0.37	0.03
41	0.3	0.13	0.16	0	0.37	0.03
42	0.25	0.22	0.13	0	0.36	0.02
43	0.42	0	0.26	0	0.01	0.32
44	0.25	0.22	0.13	0	0.36	0.02
45	0.14	0.49	0.07	0	0.28	0.01
46	0.25	0.22	0.13	0	0.36	0.02
47	0.34	0.08	0.18	0	0.36	0.04
48	0.3	0.13	0.16	0	0.37	0.03
49	0.44	0	0.27	0	0.01	0.28
50	0.2	0.35	0.1	0	0.33	0.02
51	0.14	0.49	0.07	0	0.28	0.01
52	0.34	0	0.22	0	0	0.43
53	0.34	0.08	0.18	0	0.36	0.04
54	0.2	0.35	0.1	0	0.33	0.02
55	0.47	0	0.26	0	0.2	0.08
56	0.25	0.22	0.13	0	0.36	0.02
57	0.3	0.13	0.16	0	0.37	0.03
58	0.25	0.22	0.13	0	0.36	0.02
59	0.34	0.08	0.18	0	0.36	0.04
60	0.3	0.13	0.16	0	0.37	0.03
61	0.2	0.35	0.1	0	0.33	0.02
62	0.01	0.95	0	0	0.04	0
63	0.48	0	0.27	0	0.17	0.08
64	0.25	0.22	0.13	0	0.36	0.02
65	0.25	0.22	0.13	0	0.36	0.02
66	0.2	0.35	0.1	0	0.33	0.02
67	0.3	0.13	0.16	0	0.37	0.03
68	0.2	0.35	0.1	0	0.33	0.02
69	0.25	0.22	0.13	0	0.36	0.02
70	0.2	0.35	0.1	0	0.33	0.02
71	0.3	0.13	0.16	0	0.37	0.03
72	0.25	0.22	0.13	0	0.36	0.02
73	0.2	0.35	0.1	0	0.33	0.02
74	0.25	0.22	0.13	0	0.36	0.02
75	0.46	0	0.25	0	0.22	0.07
76	0.3	0.13	0.16	0	0.37	0.03
77	0.3	0.13	0.16	0	0.37	0.03
78	0.2	0.35	0.1	0	0.33	0.02
79	0.25	0.22	0.13	0	0.36	0.02
80	0.14	0.49	0.07	0	0.28	0.01
81	0.2	0.35	0.1	0	0.33	0.02
82	0.3	0.13	0.16	0	0.37	0.03

	p_0	p_1	p_2	p_3	p_4	p_5
83	0.34	0.08	0.18	0	0.36	0.04
84	0.25	0.22	0.13	0	0.36	0.02
85	0.2	0.35	0.1	0	0.33	0.02
86	0.2	0.35	0.1	0	0.33	0.02
87	0.25	0.22	0.13	0	0.36	0.02
88	0.2	0.35	0.1	0	0.33	0.02
89	0.25	0.22	0.13	0	0.36	0.02
90	0.42	0.01	0.23	0	0.28	0.05
91	0.25	0.22	0.13	0	0.36	0.02
92	0.2	0.35	0.1	0	0.33	0.02
93	0.2	0.35	0.1	0	0.33	0.02
94	0.37	0.04	0.2	0	0.34	0.04
95	0.2	0.35	0.1	0	0.33	0.02
96	0.14	0.49	0.07	0	0.28	0.01
97	0.09	0.64	0.05	0	0.21	0.01
98	0.4	0	0.25	0	0	0.35
99	0.3	0.13	0.16	0	0.37	0.03
100	0.2	0.35	0.1	0	0.33	0.02
101	0.3	0.13	0.16	0	0.37	0.03
102	0.3	0.13	0.16	0	0.37	0.03
103	0.49	0	0.28	0	0.13	0.1
104	0.25	0.22	0.13	0	0.36	0.02
105	0.34	0.08	0.18	0	0.36	0.04
106	0.42	0.01	0.23	0	0.28	0.05
107	0.14	0.49	0.07	0	0.28	0.01
108	0.47	0	0.26	0	0.2	0.08
109	0.14	0.49	0.07	0	0.28	0.01
110	0.2	0.35	0.1	0	0.33	0.02
111	0.06	0.76	0.03	0	0.15	0
112	0.2	0.35	0.1	0	0.33	0.02
113	0.2	0.35	0.1	0	0.33	0.02
114	0.46	0	0.25	0	0.22	0.07
115	0.25	0.22	0.13	0	0.36	0.02
116	0.25	0.22	0.13	0	0.36	0.02
117	0.3	0.13	0.16	0	0.37	0.03
118	0.25	0.22	0.13	0	0.36	0.02
119	0.42	0	0.26	0	0.01	0.32
120	0.2	0.35	0.1	0	0.33	0.02
121	0.25	0.22	0.13	0	0.36	0.02
122	0.34	0.08	0.18	0	0.36	0.04

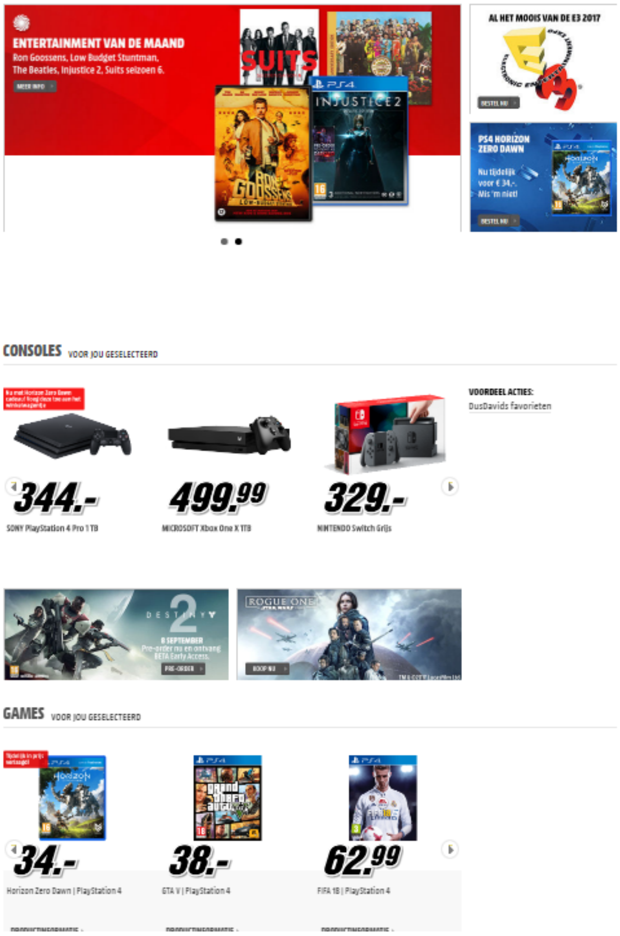
This table contains the probabilities of the segments for every respondent separately. In the first column, the respondent number is shown. The highest probability for every individual indicates to which segment the individual belongs.

Figure 2: Electronic company games page



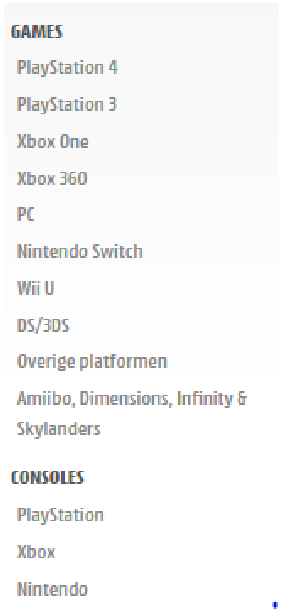
This figure shows the ordering of games based on a “most bought” sorting oder. All the games, showed in the figure, are meant for the PlayStation.

Figure 3: Electronic company gaming page



This figure shows the first displayed page when visiting the gaming page of the electronic company.

Figure 4: Electronic company filter box



The filter box gives the option to go immediately to the website page where you can find the wished products. It consists of a games filter and a console filter.

Figure 5: Pre-order upcoming games

This figure shows games from Bol.com, which are available for pre-order. This marketing content is showed visibly when visiting the gaming page of the website.

Figure 6: Advertisement of the electronic company at another website

This figure shows the the advertisement of the electronic company on another website. The content of the advertisement is based on search history on the original electronic company website.

```
data("Game")
G <- mlogit.data(Game, shape = "wide", choice = "ch",
    varying = 1:12,ranked = TRUE)
alternatives = data.frame(G$alt)
ncolG = ncol(G);
nrowG = nrow(G);
naltexp = 2;
nalt = 6;
nobs = nrow(Game);

expvar = cbind(Game$own.Xbox,Game$own.PlayStation ,Game$
    own.PSPortable ,Game$own.GameCube,Game$own.GameBoy,Game
    $own.PC,Game$hours)

#calculate utilities


#create Rmatrix with exact utilities of alternative in
    right columns, first 6 columns in Rmatrix and Ru are
    the ranks from 1 to 6
Rmatrix=matrix(0,91,6)
for(k in 1:6)
{
  colnames(Rmatrix)[k]=k
}

for(i in 1:nrow(Game))
{
  for(j in 1:6)
  {
    if(Game[i,j]==1)
    {

      Rmatrix[i, 1]=colnames(Game)[j]
    }
    if(Game[i,j]==2)
    {

      Rmatrix[i, 2]=colnames(Game)[j]
    }
    if(Game[i,j]==3)
    {

      Rmatrix[i, 3]=colnames(Game)[j]
    }
    if(Game[i,j]==4)
    {

```

```
      Rmatrix[i, 4]=colnames(Game)[j]
    }
    if(Game[i,j]==5)
    {

      Rmatrix[i, 5]=colnames(Game)[j]
    }
    if(Game[i,j]==6)
    {

      Rmatrix[i, 6]=colnames(Game)[j]
    }
  }
}

Rmatrix2=matrix(0,91,6)

for(i in 1:91)
{
  for(j in 1:6)
  {
    if(Rmatrix[i,j]=="ch.Xbox")
    {
      Rmatrix2[i,j]=1
    }
    if(Rmatrix[i,j]=="ch.PlayStation")
    {
      Rmatrix2[i,j]=2
    }
    if(Rmatrix[i,j]=="ch.PSPortable")
    {
      Rmatrix2[i,j]=3
    }
    if(Rmatrix[i,j]=="ch.GameCube")
    {
      Rmatrix2[i,j]=4
    }
    if(Rmatrix[i,j]=="ch.GameBoy")
    {
      Rmatrix2[i,j]=5
    }
    if(Rmatrix[i,j]=="ch.PC")
    {
      Rmatrix2[i,j]=6
    }
  }
}
```

```
par=c(0.2,0.2,0.2,0.2,0.2,1.5,1.1,
0.4,-3.5,-2.7,1.7,-0.1,-0.5,-1,0,-0.2)


#loglikelihood function
function_loglikelihood_LCROL=function(par){

  betas=c(par[6:10],0,par[11:16],0)

  X = matrix(nrow = nobs, ncol = nalt)
  colnames(X)=c("Xbox","Playstation","PSPortable","
      GameCube","Gameboy","PC")
  X[,1] = betas[1]+betas[7]*expvar[,1]+betas[8]*expvar
      [,7]
  X[,2] = betas[2]+betas[7]*expvar[,2]+betas[9]*expvar
      [,7]
  X[,3] = betas[3]+betas[7]*expvar[,3]+betas[10]*expvar
      [,7]
  X[,4] = betas[4]+betas[7]*expvar[,4]+betas[11]*expvar
      [,7]
  X[,5] = betas[5]+betas[7]*expvar[,5]+betas[12]*expvar
      [,7]
  X[,6] = betas[6]+betas[7]*expvar[,6]+betas[13]*expvar
      [,7]


  Ru=matrix(0,91,6)
  for(i in 1:91)
  {
    for(j in 1:6)
    {
      Ru[i,j]=X[i,Rmatrix2[i,j]]
    }
  }

 theta=c(par[1:5],0)
   p <- rep(0, 6)

   for(j in 1:(nalt))
   {
     p[j]=exp(theta[j])/sum(exp(theta))
   }

  LOGlik=matrix(0,91,1)

  for(i in 1:N)
  {
    tempPAlles=matrix(0,6,1)
    tempk0=matrix(0,1,1)
    tempkRest=matrix(0,5,1)
```

```
    tempBinnen=matrix(0,5,1)
    Ruu=Ru[i,]
    RuuOnder=matrix(0,6,6)
    RuuOnder[,1]=t(Ruu)
    RuuOnder[2:6,2]=t(Ruu[2:6])
    RuuOnder[3:6,3]=t(Ruu[3:6])
    RuuOnder[4:6,4]=t(Ruu[4:6])
    RuuOnder[5:6,5]=t(Ruu[5:6])
    RuuOnder[6,6]=t(Ruu[6])

    for(k in 1:(nalt-1))
    {
      for(l in 1:k)
      {
        tempBinnen[l]=Ruu[l]-log(sum(exp(RuuOnder[(l:nalt),l])))
      }
      tempkRest[k]=exp(sum(tempBinnen)-log(factorial(nalt-k)))
    }
    tempk0[1,1]=exp(-log(factorial(nalt)))
    tempPAlles=rbind(tempk0,tempkRest)

    for(j in 1:nrow(tempPAlles))
    {
      tempPAlles[j,1]=tempPAlles[j,1]*p[j]
    }
    LOGlik[i,1]=log(sum(tempPAlles))
  }
  LOGLIK30=sum(LOGlik)
  #print(LOGLIK30)
  return(LOGLIK30)
}

ml <- maxLik(function_loglikelihood_LCROL, start = par, method="BFGS")
summary(ml)

thetaa=ml$estimate[1:5]
thetaa=c(thetaa,0)

ptjes <- rep(0, 6)

for(j in 1:(nalt))
{
  ptjes[j]=exp(thetaa[j])/sum(exp(thetaa))
}
```

Questionnaire Gaming preferences

1. What is your gender?
   - Male
   - Female

2. What is your age?

3. How many hours on average do you play games per week (including games on your smartphone)?

4. Which platforms do you own as a gaming platform?
   - PC (as a gaming platform)
   - Xbox
   - Wii
   - Smartphone (as a gaming platform)
   - PlayStation
   - Portable Gameboy

5. Do you use an Android or IOS device on your smartphone?
   - Android
   - IOS
   - I do not have a smartphone

6. Rank the gaming platforms from most preferred as 1 to least preferred as 5.

   - ☐ Xbox
   - ☐ PlayStation
   - ☐ Wii
   - ☐ PC
   - ☐ smartphone
   - ☐ Portable Gameboy

7. How important are the following items for a gaming platform? (1 means not important at all, 5 means very important)

	1	2	3	4	5
importance of the price of the hardware					

importance of the price of software					
importance of availability of games					
importance of portability					
importance of multiplayer capabilities (online or offline)					
importance of other capabilities (e.g. playing DVD's)					
importance of exchangeability of games (with previous versions of the gaming platform)					

8. which genre of games do you like most?
    - Puzzle
    - Shooting/Action
    - Strategy
    - Sports/Race