Modelling heterogeneity in rank-ordered data

Bachelor Thesis Econometrics and Operations Research

ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Christian Vijfvinkel
414895

Supervisor: Prof. Dr. D. Fok
Second Assessor: A. Castelein

Abstract

We consider the modelling of rank-ordered data. This data results from questions in surveys which ask to rank alternatives and contains more information than data with only the most-preferred alternative of individuals. Standard models for modelling this kind of data make assumptions which are undesirable, especially when heterogeneity across individuals is present. These models are then less useful. We consider unobserved heterogeneity in ranking capabilities across individuals and preference heterogeneity in this paper. We use different models and evaluate which model performs best under different circumstances. We conclude that the latent-class rank-ordered logit model performs best in case of unobserved heterogeneity in ranking capabilities and in case of moderate preference heterogeneity, otherwise the mixed rank-ordered logit model performs best. We also find that one needs to be careful to conclude which kind of heterogeneity is present in the data, because the used models lead to contrary conclusions in both a simulation study and an empirical application.

July 2, 2017

1 Introduction

Obtaining and using consumer data to get to know people better is an important business. In the field of marketing, when you know people's preferences better it is easier to make more appropriate product recommendations, which increases customer satisfaction and sales. Not only in the field of marketing knowing the preferences of people can make great difference, think for example of population studies such as the modelling of voter preferences.

When investigating the preferences of individuals, a widely used tool for obtaining data is a survey. In a survey, individuals are often asked to give their most preferred item out of a set of different alternatives. However, when you ask someone to give a complete ranking of alternatives a lot more information can be obtained. We call this data rank-ordered data. A commonly used model to model this data is the multinomial logit (MNL) model introduced by McFadden (1973). However, this model only takes the most preferred choice into account. The rank-ordered logit (ROL) model introduced by Beggs, Cardell, & Hausman (1981) is a widely used model to estimate the preferences of individuals when a complete ranking of alternatives is given. However, these models make assumptions which could lead to less informative and less reliable estimates. It is therefore of great importance to see what happens if we relax some of these assumptions.

A downside of the use of the standard rank-ordered logit model is the implicit assumption that a person is able to rank all alternatives. When there are a lot of alternatives to rank or there are alternatives unknown to an individual, this assumption is not likely. Therefore in Fok, Paap, & Van Dijk (2012), the rank-ordered logit model is extended with unobserved heterogeneity in ranking capabilities called the latent-class rank-ordered logit (LCROL) model.
However, it is also likely that different individuals react different to changes in explanatory variables called preference heterogeneity, which is not modelled by the multinomial, rank-ordered or latent-class rank-ordered logit models. Current research concerning rank-ordered data mainly focuses on modelling different kinds of heterogeneity across individuals one at a time. For example, in Fok et al. (2012) the main purpose is to model heterogeneity in ranking capabilities, while in Calfee, Winston, & Stempski (2001) the preference heterogeneity is modelled. In this paper, we investigate what the differences in parameter estimates are when there is unobserved heterogeneity in ranking capabilities and preference heterogeneity across individuals using different models. We also investigate which model performs best when there is heterogeneity in ranking capabilities or preference heterogeneity across individuals present.

We organize the remainder of this paper as follows. In section 2 we give an overview of the methods for modelling rank-ordered data. Then in section 3 we explain the models we use in detail. Afterwards, in section 4 we explain which data we use for this research and in section 5 we report the results of the different models. Finally, in section 6 we give our conclusion and suggestions for future research.

2 Literature review

In the literature there are several methods available for modelling rank-ordered data. The most used methods are the multinomial logit model (McFadden, 1973) and the rank-ordered logit model (Beggs et al., 1981). Where the multinomial logit models only uses the most preferred choice from the set of alternatives as dependent variable, the rank-ordered logit model takes the full ranking of the alternatives into account. This results in more efficient parameter estimates, because the full ranking contains more information than only the most preferred choice. However, when the ranking is partially random because someone has too little knowledge of all the alternatives, the parameter estimates contain biases as shown in Hausman & Ruud (1987). Fok et al. (2012) solve this problem by modelling the unobserved heterogeneity in ranking capabilities across individuals using latent classes. They conclude that the use of the standard rank-ordered logit model leads to a bias in parameters estimates when the respondents are not able to give a full ranking. The proposed latent-classes enhanced rank-ordered logit model solves this problem and gives more efficient estimates than the multinomial logit and standard rank-ordered logit models. Other ways to deal with the different ranking abilities between respondents is the use of only a fixed number of rankings or to give a weight to each ranking in the rank-ordered logit model as is done in Hausman & Ruud (1987).

However, all these models have one downside in common, namely the parameters are assumed to be constant across all individuals. In these models the independence of irrelevant alternatives (IIA) property also holds. To relax this assumption one could construct a rank-ordered probit model as is done in Hajivassiliou & Ruud (1994). However, the use of a rank-ordered probit models involves the use of multiple fold integrals, which could be time consuming especially in the case of a lot of ranking choices (Beggs et al., 1981). A different way to model rank-ordered data without the IIA assumption and with different parameters for each individual is to construct a mixed rank-ordered logit model as is done in Calfee, Winston, & Stempski (2001). This model is based on the mixed logit model by Revelt & Train (1998) which involves less multiple-fold integrals than the rank-ordered probit model.

Other ways of modelling a rank-ordered dependent variable is for example the use of cluster analysis as is done by Busse, Orbanz, & Buhmann (2007) and Murphy & Martin (2003). In cluster analysis respondents with a common preference behaviour are grouped together. Then a probability model is constructed to represent the groups. A problem when using these kinds of methods to model a rank-ordered dependent variable is again the heterogeneity in ranking capabilities across individuals. The clusters then contain rankings of different lengths. In Murphy & Martin (2003) the constructed model is not able to incorporate this kind of heterogeneity. However, the model proposed in Busse et al. (2007) is able to model data consisting of respondents with different ranking abilities and they conclude that parameter estimates improve when also partially rankings are used.

The main difference between the cluster models and the logit based models like the rank-ordered logit model and the mixed logit model is the assumption of the distribution to capture for example preference heterogeneity. In cluster models, this distribution is assumed to be discrete, while in logit models this distribution is assumed to be continuous. Using discrete distributions, one can divide the individuals in clusters.

So, in the literature there is little attention for the comparison in model performance when different kinds of heterogeneity are present. Therefore, we focus on the model performance when unobserved heterogeneity in ranking capabilities across individuals is present and when preference heterogeneity is present.

3 Model specification

In this this research we use the multinomial logit model, the rank-ordered logit model, the latent-class rank-ordered logit model and the mixed rank-ordered logit model. We explain each of these models in detail in this section.

3.1 Multinomial logit model

The multinomial logit model is a widely used model for modelling discrete choices when there are more than two choices available. To explain the multinomial logit model as well as the models that follow in this paper, we make use of the unobserved random utility concept. The unobserved random utility which individual i gets from alternative j, denoted as U_{ij}, is given by:

$$U_{ij} = x_i'\beta_j + z_{ij}'\gamma + \epsilon_{ij} \tag{1}$$

Where x_i is a vector containing characteristics of individual i, β_j is a vector containing the parameters specific to alternative j, z_{ij} is a vector of alternative specific variables, γ is a vector containing the parameters, which are fixed across alternatives, for each alternative specific variable and ϵ_{ij} denotes the error-term. Every individual i has a set of unobserved utilities which contains the utility of each alternative j for individual i, $U_{i1}, U_{i2}, ..., U_{iJ}$. We use the multinomial logit model to model the most preferred option out of a set of J alternatives. This implies that the utility obtained by individual i from the most preferred alternative, denoted as alternative x, is larger than or equal to the maximum utility obtained by any of the other alternatives:

$$U_{ix} \geq \max(U_{i1}, U_{i2}, ..., U_{iJ}) \tag{2}$$

When the error terms are identical and independently distributed according to a type-I extreme value distribution, we talk about the multinomial logit (MNL) model. When we take Y_i as the most preferred choice of individual i out of a set of alternatives, the choice probability for the most preferred alternative x equals:

$$P(Y_i = x|\beta) = P(U_{ix} \geq \max(U_{i1}, U_{i2}, ..., U_{iJ})) = \frac{\exp(x_i'\beta_x + z_{ix}'\gamma)}{\sum\limits_{j=1}^{J} \exp(x_i'\beta_j + z_{ij}'\gamma)} \tag{3}$$

Where x, β, z and γ have the same meaning as in equation (1).

We estimate the parameters in the multinomial logit model using maximum likelihood. Therefore, we need to maximize the (log-) likelihood function over the β parameters. The likelihood and log-likelihood functions for the multinomial logit model are given by:

$$\begin{aligned}
\mathcal{L}(\beta) &= \prod_{i=1}^{N}\Bigg(\prod_{j=1}^{J} P(Y_i = j)^{I(Y_i=j)}\Bigg) = \prod_{i=1}^{N}\Bigg(\prod_{j=1}^{J} \frac{\exp(x_i'\beta_j + z_{ij}'\gamma)}{\sum\limits_{k=1}^{J}\exp(x_i'\beta_k + z_{ik}'\gamma)}^{I(Y_i=j)}\Bigg), \\
\log\mathcal{L}(\beta) &= \sum_{i=1}^{N}\Bigg(\sum_{j=1}^{J} I(Y_i = j)\log\Bigg\{\frac{\exp(x_i'\beta_j + z_{ij}'\gamma)}{\sum\limits_{m=1}^{J}\exp(x_i'\beta_m + z_{im}'\gamma)}\Bigg\}\Bigg)
\end{aligned} \tag{4}$$

where N is the number of individuals, J is the number of alternatives and $I(Y_i = j)$ is an indicator function which equals 1 if $Y_i = j$ and 0 otherwise. β_J is put equal to zero for identification purposes. Maximization of the log-likelihood function above as well as the log-likelihoods functions in the remaining of this paper can be done using any numerical optimization algorithm like the BFGS or Newton-Raphson algorithm.

So the multinomial logit model only models the most preferred alternative out of a set of different alternatives. However, when you ask someone to give a complete ranking of alternatives we can obtain a lot more information. For example, you then also know what an individual absolutely dislikes. To model the full ranking of an individual we use the rank-ordered logit model. We explain this model in more detail in the next section.

3.2 Rank-ordered logit model

In the rank-ordered logit model, we again make use of the random utility concept defined in equation (1). As told before, the rank-ordered logit model takes the full ranking of an individual into account. To further explain this model, we follow the notation used in Fok et al. (2012). Therefore, we denote the ranking given by individual i as $y_i = (y_{i1}, y_{i2}, ..., y_{iJ})'$, where y_{ij} is the rank given by individual i to alternative j. Also, we denote the ranking seen from the item perspective as $r_i = (r_{i1}, r_{i2}, ..., r_{iJ})'$ where r_{ij} denotes the alternative number that was ranked j^{th} by individual i. Using this notation, the probability of observing a ranking given by individual i is according to the rank-ordered logit model equal to:

$$P(r_i|\beta) = P(U_{ir_{i1}} > U_{ir_{i2}} > ... > U_{ir_{iJ}}) = \prod_{j=1}^{J-1} \frac{\exp(x_i'\beta_{r_{ij}} + z_{ir_{ij}}'\gamma)}{\sum_{l=j}^{J} \exp(x_i'\beta_{r_{il}} + z_{ir_{il}}'\gamma)} \tag{5}$$

The difference in the choice probabilities between the multinomial logit and rank-ordered logit model can best be explained by a small example. When there are, for example, three alternatives to choose from and individual i prefers the first alternative the most, then the third alternative and lastly the second alternative. The probability for this ranking according to the rank-ordered logit model equals, when again using the random utility concept, $P(U_{i1} > U_{i3} > U_{i2})$. However, the multinomial logit model models this as follows $P(U_{i1} > U_{i2} \cap U_{i1} > U_{i3})$. The rank-ordered logit model can be seen as a series of multinomial logit models, where first a multinomial logit model is created for the most preferred alternative, then one for the second most preferred alternative without using the most preferred alternative and so forth. This is a special property of the logit models.

We estimate the parameters in the rank-ordered logit model again using maximum likelihood. The likelihood and log-likelihood functions for the rank-ordered logit model equal:

$$\begin{aligned} \mathcal{L}(\beta) &= \prod_{i=1}^{N} \left(\prod_{k=1}^{J-1} \frac{\exp(x_i'\beta_{r_{ik}} + z_{ir_{ik}}'\gamma)}{\sum_{m=k}^{J} \exp(x_i'\beta_{r_{im}} + z_{ir_{im}}'\gamma)} \right), \\ \log \mathcal{L}(\beta) &= \sum_{i=1}^{N} \left(\sum_{k=1}^{J} (x_i'\beta_{r_{ik}} + z_{ir_{ik}}'\gamma) - \log \left\{ \sum_{m=k}^{J} \exp(x_i'\beta_{r_{im}} + z_{ir_{im}}'\gamma) \right\} \right) \end{aligned} \tag{6}$$

where N is again the number of individuals, J is the number of alternatives and β_J is put equal to zero for identification purposes.

So, the rank-ordered logit model uses more information compared to the multinomial logit model, which generally results in more efficient estimates. However, the rank-ordered logit model also has some disadvantages. One of them is that is the heterogeneity in ranking capabilities of individuals is not taken into account. It is unlikely that all individuals can give a complete ranking, because most individuals are likely not completely familiar with all alternatives presented. This results in partially unreliable rankings. The rank-ordered logit model assumes that an individual can give a complete unbiased ranking of alternatives. What we mean by this is that an individual is assumed to give a complete ranking according to the underlying utility function of that individual. When this is not the case, the rank-ordered logit model gives unreliable results. Therefore, to make a more realistic model, one has to incorporate the heterogeneity in ranking capabilities across individuals into the model. One way to do this is to construct the so-called latent-class rank-ordered logit model, introduced by Fok et al. (2012). We explain this model in the next section.

3.3 Latent-class rank-ordered logit model

When individuals are not able to rank all the alternatives according to their real utility levels, the alternatives which are not known by the individual are likely ranked randomly. When one wants to estimate the parameters of a rank-ordered logit model it seems obvious to only use the alternatives that are ranked correctly. We denote the number of correctly ranked alternatives by k. So, the alternatives corresponding to a rank higher than k are not used in the estimation of the model parameters, but we know that the utility obtained from each of the first k items is larger than the utility obtained by any of the other alternatives which are not ranked correctly. The probability of observing a ranking of individual i where the number of correctly ranked alternatives equals k is:

$$\begin{aligned} P(y_i|k,\beta) &= P(U_{ir_{i1}} > U_{ir_{i2}} > ... > U_{ir_{ik}} > \max(U_{ir_{ik+1}}, ..., U_{ir_{iJ}})) \\ &= \left(\prod_{j=1}^{k} \frac{\exp(x_i'\beta_{r_{ij}} + z_{ir_{ij}}'\gamma)}{\sum_{l=j}^{J} \exp(x_i'\beta_{r_{il}} + z_{ir_{il}}'\gamma)} \right) \frac{1}{(J-k)!} \end{aligned} \tag{7}$$

As already told, an individual can only rank k alternatives correctly. So, $J-k$ items are ranked randomly and the probability of observing an ordering of the last $J-k$ items equals $\frac{1}{(J-k)!}$, which is the last part of equation (7).

To use the expression in equation (7), we have to determine which value(s) of k we use. One could assume that k is fixed at a certain amount of alternatives for all individuals, like is done in Hausman & Ruud (1987). However, it is unlikely that all individuals have the exact same ranking ability. Also, when assuming a fixed number of ranking ability a lot of valuable information is lost. Think, for example, at the case where only 1 % of the individuals can rank only two out of six items correctly, while the other 99 % can rank all alternatives completely. If you assume a fixed ranking ability across individuals, you can only use the first two alternatives ranked of each individual to avoid estimation bias. In this case you do not use a lot of valuable information.
To make more accurate estimation results for the model parameters, we make use of the latent-class rank-ordered logit model (Fok et al., 2012). Using this model, we assume that k varies across individuals and does not need to be fixed.

The latent-class rank-ordered logit model uses latent-classes to incorporate the unobserved heterogeneity in ranking capabilities across individuals. Therefore, we divide the individuals into J latent classes, where each class k equals the number of alternatives that an individual ranks correctly, with $k = 0, 1, ..., J-1$. Where k is not larger than $J-1$, because when someone can rank $J-1$ alternatives correctly it automatically follows that he or she can rank all the J alternatives correctly.

In the latent-class rank-ordered logit model, the probability of observing a ranking of individual i equals:

$$P(y_i|\beta, p) = \sum_{k=0}^{J-1} p_k P(y_i|k, \beta) \tag{8}$$

where p_k is the probability that an individual i belongs to class k with $0 \leq p_k \leq 1$ and $\sum_{k=0}^{J-1} p_k = 1$ and where $P(y_i|k, \beta)$ is the probability of observing ranking y_i when only k alternatives are ranked correctly, see equation (7). With a correct ranking, we mean that the alternatives are ranked according to the latent utility function of an individual.

To estimate the parameters in the latent-class rank-ordered logit model, we again make use of maximum likelihood estimation. The likelihood and log-likelihood functions equal:

$$\mathcal{L}(\beta, p) = \prod_{i=1}^{N} \sum_{k=0}^{J-1} \frac{p_k}{(J-k)!} \left[\prod_{l=1}^{k} \frac{\exp(x_i'\beta_{r_{il}} + z_{ir_{il}}'\gamma)}{\sum\limits_{m=l}^{J} \exp(x_i'\beta_{r_{im}} + z_{ir_{im}}'\gamma)} \right], \tag{9}$$

$$\log \mathcal{L}(\beta, p) = \sum_{i=1}^{N} \log\Big(\sum_{k=0}^{J-1} p_k \exp\Big[-\log((J-k)!) + \sum_{l=1}^{k} \Big\{ (x_i'\beta_{r_{il}} + z_{ir_{il}}'\gamma) - \log\Big(\sum_{m=l}^{J} \exp(x_i'\beta_{r_{im}} + z_{ir_{im}}'\gamma) \Big) \Big\} \Big] \Big) \tag{10}$$

We again use the BFGS algorithm to maximize the log-likelihood function. To incorporate the restrictions which p_k has to obey, we optimize over J additional parameters called θ_j, where $j = 1, ..., J$. We transform these parameters into the p_k probabilities as follows: $p_k = \frac{\exp(\theta_k)}{\sum_{i=1}^{J} \exp(\theta_i)}$. Here, θ_J is set equal to zero for identification purposes.

So, the latent-class rank-ordered logit model incorporates the heterogeneity in ranking capabilities between individuals which results in a more realistic model. However, the latent-class rank-ordered logit model as well as the standard rank-ordered logit and multinomial logit models make some assumptions which could be undesirable. The independence of irrelevant alternatives (IIA) assumption holds and the parameters are assumed to be constant across individuals. This IIA property states that the ratio of preferring one alternative over another does not depend on other "irrelevant" alternatives.
It is likely that different individuals react different to changes in explanatory variables, the so-called preference heterogeneity. Therefore, as an extension to the paper by Fok et al. (2012) we do not assume the IIA property and we model preference heterogeneity. To model this, we use the mixed rank-ordered logit (Calfee et al., 2001) model. We compare the latent-class rank-ordered logit model and the mixed rank-ordered logit model and investigate whether these models are appropriate for each kind of heterogeneity or that there cannot be made clear conclusions about which model models which kind of heterogeneity

the best.

3.4 Mixed rank-ordered logit model

The mixed rank-ordered logit model is a special case of the general mixed logit model defined in Revelt & Train (1998). This model solves the IIA property and incorporates the possibility of different parameters across individuals, which makes it possible to model preference heterogeneity across individuals. This makes the mixed logit model a more realistic model than for example the multinomial or standard rank-ordered logit model.

To model the preference heterogeneity across individuals, the mixed logit model assumes a probability distribution, the so-called mixing distribution $f(\beta)$, for the parameters β which are assumed to vary across individuals. The choice probabilities in the mixed logit model are a weighted average of the standard choice probabilities with weights given by the mixing distribution $f(\beta)$. When estimating the parameters in a mixed logit model, we estimate the parameters that characterize the mixing distribution of a parameter which varies across individuals. For example, when we assume that a certain parameter is distributed according to the normal distribution, then we estimate the mean and standard deviation of this distribution. We can check the presence of preference heterogeneity by testing whether the standard deviation is significantly different from zero.

When we estimate the mixed logit parameters, these parameters are of course not known. The estimation process therefore involves the computation of integrals, which do not have a closed-form solution. The number of integrals is equal to the number of parameters which are assumed to vary across individuals. When you assume that more than one parameter varies across individuals, then you could also incorporate correlation between these parameters.
We also have to assume a probability distribution for the parameters which one assumes to vary across individuals. Often, the normal distribution is chosen, as is done in Revelt & Train (1998). However, every probability distribution is possible to use as mixing distribution. To determine which probability distribution to use one can, for example, examine whether the parameters have to be positive.

We can use the mixed logit model in the rank-ordered data framework, by using the choice probabilities of the rank-ordered logit model as stated in equation (5). We then talk about the mixed rank-ordered logit model (Calfee et al., 2001). In the mixed rank-ordered logit model, the probability to observe a ranking equals:

$$P[U_{ir_{i1}} > U_{ir_{i2}} > ... > U_{ir_{iJ}}] = \int \prod_{j=1}^{J-1} \frac{\exp(x_i'\beta_{r_{ij}} + z_{ir_{ij}}'\gamma)}{\sum\limits_{l=j}^{J} \exp(x_i'\beta_{r_{il}} + z_{ir_{il}}'\gamma)} f(\beta_i|\theta_i) d\beta_i \qquad (11)$$

Where $f(\beta_i|\theta_i)$ is the mixing distribution where the real parameters equal the vector θ_i. This model allows for correlation across the stochastic terms of the utility function of individuals, therefore this model does not impose the IIA property. In this way, the signs of the off diagonal elements in the covariance matrix of the error terms denote positive or negative association between the choices.

The likelihood function of the mixed rank-ordered logit model equals (Calfee et al., 2001):

$$L(\theta) = \sum_{i=1}^{N} \log \left[\int \prod_{j=1}^{J-1} \frac{\exp(x_i'\beta_{r_{ij}} + z_{ir_{ij}}'\gamma)}{\sum_{l=j}^{J} \exp(x_i'\beta_{r_{il}} + z_{ir_{il}}'\gamma)} f(\beta_i|\theta) d\beta_i \right] \tag{12}$$

To estimate the parameters, we simulate the log likelihood function as well as the probability of observing a certain ranking, because the choice probability functions now involve integrals which do not have an analytical solution. Therefore, we use maximum simulated likelihood (Revelt & Train, 1998) to estimate the parameters.
First, we have to take R draws from the chosen mixing distribution to make a good representation of the mixing distribution. Using a small amount of draws can speed up the computation, which is essential when running times of several hours are no exception. Therefore we use Halton draws to generate draws from the mixing distribution. The use of Halton draws is far more efficient than using random numbers as found in Train (2000), because the Halton sequence provides more evenly coverage of the unit interval than random numbers. A Halton sequence is based on a given number, usually a prime. When we take for example the Halton sequence of the number 2, the unit interval is first divided in half, then in fourths, then in eights and so on. We then end up with the sequence $\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, ...$
These Halton draws are first transformed to the specified mixing distribution. We label these R draws β^r, with $r = 1, 2, ..., R$. Then for each β^r, we calculate the rank-ordered logit probability like in equation (5), while the coefficients who are not assumed to be vary across individuals remain constant. We call this probability of a ranking given by individual i: $S_i(\beta_i^r)$. The simulated choice probabilities of a ranking given by individual i, $SP_i(\beta)$, are then equal to the average of these rank-ordered logit probabilities:

$$SP_i(\beta) = \frac{1}{R} \sum_{r=1}^{R} S_i(\beta_i^r) \tag{13}$$

The variance of this simulated probability decreases when you take more draws, however taking more draws also increases the computation time. So, one have to make a trade-off between variance reduction and computation time. We choose 500 Halton draws to get accurate estimates in a reasonable amount of time. This simulated probability gives an unbiased estimate of the probability in equation (11).

The simulated log-likelihood function is now defined as $SLL(\theta) = \sum_{i=1}^{N} \log SP_i(\beta)$. We maximize this simulated log-likelihood function using a numerical optimization algorithm. The estimated parameters resulting from the optimization of the simulated log-likelihood function are biased, because the logarithmic transformation of the simulated choice probabilities is a biased estimate of the logarithm of the real choice probabilities (Train, 2000). However, this bias decrease when the number of draws, R, increases.
The modelling of different parameters for each individual is equal to the modelling of preference heterogeneity across individuals. The latent-class rank-ordered logit model, as already told before, models the unobserved heterogeneity in ranking capabilities. To investigate the difference in parameter estimates between these two types of heterogeneity modelling we compare the mixed rank-ordered logit model with the latent-class rank-ordered logit model.

4 Data

To compare the different models in terms of parameter estimates when different kinds of heterogeneity are present, we first perform a simulation study. This also allows us to control which kind of heterogeneity is present in the data and also allows us to control the degree of this heterogeneity. Therefore, we simulate latent utilities according to the following data generating process: $U_{ij} = \beta_{0j} + x_{i1}\beta_{1j} + x_{i2}\beta_{2j} + w_{ij}\beta_3 + \epsilon_{ij}$. Here, x_{i1} is an individual specific variable which we draw from the standard normal distribution, x_{i2} is also an individual specific variable which can only take on the values 0 or 1 with equal probability and w_{ij} is an alternative specific variable which we draw from a normal distribution with mean equal to respectively 3.0, 4.0, 5.0 and 2.0 for the four different alternatives and standard deviations respectively equal to 1.0, 2.0, 3.0 and 0.5. In this way, the alternatives are pretty different from each other. We assume that there are four alternatives to limit the number of parameters we have to estimate and we take the fourth alternative as base alternative. We choose the β parameters in such a way that there are no parameters which are a lot larger than others and we choose about an equal distribution in positive and negative parameters. The values for these parameters are displayed in tables 1 and 2 in section 5.

We perform this simulation two times. In the first simulation, we generate the utilities with fixed parameters across individuals, so no preference heterogeneity is present, but we incorporate heterogeneity in ranking capabilities into the rankings. We do this by first assigning a ranking ability to each simulated individual, which denotes the number of items which can be ranked according to the underlying utility function, and we then randomly shuffle the number of lowest ranked items according to this ranking ability. When the ranking ability of an individual is, for example, one, we randomly shuffle the three lowest ranked items. In this simulation we perform 1000[a] simulations, where we generate in each simulation 1000 individuals. There are four alternatives to choose from, therefore there are four latent-classes where the probability of the latent-classes is denoted as p_0, p_1, p_2 and p_3. So p_0 denotes the probability that an individual cannot rank any of the items correctly, p_1 equals the probability that an individual can rank only one alternative correctly and so on. In the first simulation we generate the data with p_0 equal to 0, p_1 and p_2 equal to 0.30 and p_3 equal to 0.40. So, there are no individuals which cannot rank any of the alternatives correctly. This allows us to get reliable estimates when using the multinomial logit model, because this model only takes the most preferred alternative into account. To get parameter estimates using the mixed rank-ordered logit model we have to assume a probability distribution for the parameter which is considered to vary across individuals. We assume that all individuals react differently on changes in the alternative-specific variable in the mixed rank-ordered logit model, so in our case we assume that the β_3 parameter follows a probability distribution. We assume a normal distribution for this variable.

For the second simulation, we simulate data without heterogeneity in ranking capabilities, but with preference heterogeneity. We do this by randomly draw the β_3 parameter from a normal distribution with mean 0.75 and standard deviation equal to 1.0. In this way, the alternative-specific variable has no fixed coefficient, which means that every individual has a different reaction on changes of this variable relative to the population mean.

[a] Due to processor power limitations we are only able to make 1000 simulations in a reasonable amount of time.

We choose a reasonably high standard deviation compared to the mean to incorporate a large difference in preference for the alternative-specific variable. We again assume that there are four alternatives to choose from with the fourth alternative as base category and perform again 1000 simulations with 1000 individuals in each simulation.
In the appendix (table 4 and table 5) we report the distribution of the ranks of the four alternatives averaged over the 1000 simulations for both simulations. Here one can see how often each alternative is ranked $1^{st}, 2^{nd}, 3^{rd}$ and 4^{th}. We see that in the first simulation alternative 4 is the most popular alternative, while in the second simulation all alternatives have about the same popularity. We evaluate all the estimated parameters using the root mean squared error (RMSE) measure.

Next to the simulation study, we also use a dataset consisting of real-world data to investigate the performance of our models and check which kind of heterogeneity is present. The data available for this research is based on a survey among 91 students of the Erasmus University Rotterdam concerning the ranking of gaming platforms [b]. The 91 students ranked the following game platforms: *Xbox, Playstation, Playstation Portable (PSP), GameCube, GameBoy* and *PC*. Further, the dataset also contains which game platform the students own, the importance of several characteristics of the gaming platforms (on a 1- to 5-scale), the gender and age of the respondents and the weekly number of hours spent on gaming by the respondents. For the parameter estimation we only use 90 observations, because there is one individual who gave two alternatives the same rank. Therefore, we ignore this observation. In table 6 in the appendix, the distribution of the ranks of the different alternatives is displayed. So, one can see how often each alternative is ranked $1^{st}, 2^{nd}$, etcetera. We see that the PC is the most popular alternative.

5 Results

In this section, we present the results of the simulation study and the results of the application on real-life data. We first analyse the result of the simulation where only heterogeneity in ranking capabilities across individuals is incorporated into the data, and then analyse the simulation where only preference heterogeneity is incorporated into the data. Finally, we report the results of the different models when we use the gaming platform preference data.

5.1 Unobserved heterogeneity in ranking capabilities simulation

In the first simulation we generate data with unobserved heterogeneity in ranking capabilities, but without preference heterogeneity. We use the multinomial logit, rank-ordered logit, latent-class rank-ordered logit and mixed rank-ordered logit models to estimate the parameters. The true parameter values, as well as the parameter estimates are displayed in table 1.

Starting with the results of the multinomial logit model, we see that these parameter estimates are approximately unbiased. This is completely as expected because there are no individuals which cannot rank any of the alternatives correctly. The results of the rank-ordered logit model, however, are biased. This is also as expected, because 60% of the individuals are not able to give a completely correct ranking, while the rank-ordered logit model assumes that all individuals can give a complete ranking. When using the

[b] Source: `http://qed.econ.queensu.ca/jae/2012-v27.5/fok-paap-van_dijk/`

Table 1: Estimated mean and RMSE (in brackets) of the parameters over 1000 simulations where each simulation consists of N = 1000 individuals, when only heterogeneity in ranking capabilities across individuals is present

Parameter	True	MNL[a]	ROL[b]	LCROL[c]	LCROL[d]	Mixed ROL[e]
β_{01}	1.00	1.01	-0.70	1.02	1.01	-0.66
		(0.18)	(1.71)	(0.16)	(0.16)	(1.67)
β_{11}	0.75	0.76	0.56	0.76	0.76	0.62
		(0.10)	(0.21)	(0.08)	(0.08)	(0.15)
β_{21}	-0.30	-0.30	-0.15	-0.30	-0.30	-0.17
		(0.18)	(0.20)	(0.16)	(0.16)	(0.20)
β_{02}	0.25	0.25	-1.08	0.26	0.26	-1.10
		(0.21)	(1.34)	(0.17)	(0.17)	(1.36)
β_{12}	-0.50	-0.50	-0.06	-0.50	-0.50	-0.09
		(0.12)	(0.44)	(0.10)	(0.10)	(0.42)
β_{22}	0.45	0.45	0.25	0.45	0.45	0.29
		(0.24)	(0.25)	(0.19)	(0.19)	(0.22)
β_{03}	-0.25	-0.26	-1.15	-0.25	-0.25	-1.27
		(0.24)	(0.91)	(0.19)	(0.18)	(1.02)
β_{13}	1.00	1.02	0.53	1.02	1.02	0.60
		(0.15)	(0.48)	(0.12)	(0.12)	(0.41)
β_{23}	0.80	0.82	0.34	0.82	0.82	0.41
		(0.29)	(0.48)	(0.23)	(0.23)	(0.42)
β_3	-0.75	-0.76	-0.32	-0.76	-0.76	-0.46
		(0.05)	(0.43)	(0.04)	(0.04)	(0.29)
p_0	0.00	-	-	0.01	-	-
				(0.01)		
p_1	0.30	1.00	-	0.29	0.30	-
				(0.04)	(0.03)	
p_2	0.30	-	-	0.30	0.30	-
				(0.04)	(0.04)	
p_3	0.40	-	1.00	0.40	0.40	1.00
				(0.04)	(0.04)	

[a] Multinomial logit model
[b] Rank-ordered logit model
[c] Latent-class rank-ordered logit model
[d] Latent-class rank-ordered logit model with $p_0 = 0$
[e] Mixed rank-ordered logit model with estimated standard deviation of β_3 equal to 0.38 with a mean standard error of 0.04 and RMSE of 0.38

latent-class rank-ordered logit models, we see that these models give approximately unbiased estimates with a lower RMSE than the multinomial logit model. When we remove the latent-class where no individual could rank any of the alternatives correctly there is also a slight improvement in parameter estimates and RMSE of the parameters, so removing redundant latent-classes improves the estimates. From these results one can see that the latent-class rank-ordered logit models gives unbiased parameter estimates and a lower RMSE for the parameters than the multinomial logit model, which makes it a more accurate model than the standard models for modelling rank-ordered data. When we look at the results of the mixed rank-ordered logit model, we see that the parameter estimates resulting from this model are biased. This is expected because, like the standard rank-ordered logit model, this model assumes that every individual can provide a completely correct ranking. However, in this data 60% of the individuals cannot do this. The standard deviation of the β_3 variable equals 0.38 and has a mean standard error over the 1000 simulations of 0.04. This standard deviation is significantly different from zero at the 5% significance level. So, the mixed rank-ordered logit model estimates a standard deviation of parameter β_3 significantly different from zero, which could indicate preference heterogeneity for the alternative specific variable w_{ij}. However, the data is generated with only fixed parameters, so there is no preference heterogeneity present. Therefore, this indicates that the heterogeneity in ranking capabilities is difficult to differentiate from the preference heterogeneity. To further investigate the difference in modeling heterogeneity in ranking capabilities and the preference heterogeneity, we perform a second simulation.

Before performing the second simulation with only preference heterogeneity in the data, we first have a further look into the performance of the models when heterogeneity in ranking capabilities is present. Therefore, we investigate the performance of the multinomial logit, rank-ordered logit, latent-class rank-ordered logit and mixed rank-ordered logit models when the degree of heterogeneity in ranking capabilities across individuals differs. We now evaluate the performance of these models by computing the mean RMSE only over the β parameters, because these are the only parameters that all models estimate. In case of the mixed rank-ordered logit model, we only use the estimated mean of the β_3 parameter. The data is generated according to same data generating process mentioned earlier, but now we vary the degree of heterogeneity in ranking capabilities. We do this by varying the probabilities that an individual can rank only a certain amount of alternatives correctly. In this case, when there are four alternatives present, we vary p_1, p_2 and p_3. We, again, assume that the probability that nobody can rank any of the alternatives correctly is equal to zero. To visualize the performance of the models with varying heterogeneity in ranking capabilities we construct a contour plot of the mean RMSE over the β parameters for each model. The contour plots of the rank-ordered logit, latent-class rank-ordered logit and mixed rank-ordered logit models are displayed in figure 1.

In the contour plots, the x-axis denotes the probability that an individual can rank only one alternative correctly, p_1 and the y-axis denotes the probability that an individual can rank all alternatives correctly, p_3. We can also implicitly obtain the probability that an individual can rank two alternatives correctly, p_2, from the contour plot. To compute the mean RMSE over the β parameters for each model, we simulate ten times 1000 individuals for each feasible combination of p_1, p_2 and p_3. We calculate this mean RMSE for different values of p_1, p_2 and p_3 ranging from 0 to 1 with 0.1 increments (therefore you can see the sharp edges on the border of each graph). Each contour plot is a below triangular plot, because the probabilities p_1, p_2 and p_3 have to sum up to one. We can interpret the contour plots as follows. When we, for example, look at the point $(p_1, p_3) = (0.2, 0.2)$ in the graph of the rank-ordered logit model, we see that the mean RMSE over the β parameters equals about 0.38 and $p_2 = 1 - p_1 - p_3 = 1 - 0.2 - 0.2 = 0.6$. Each point in the contour plot can be interpreted in this same way. We do not show the contour plot of the multinomial logit model, because this contour plot gives the same mean RMSE for all possible values of p_1, p_2 and p_3. This is because the multinomial logit model only takes the most preferred alternative into account which is always ranked correctly in our setting. However, we do compute the mean RMSE over the β parameters for the multinomial logit model to compare this model with the other models. This mean RMSE equals 0.18.

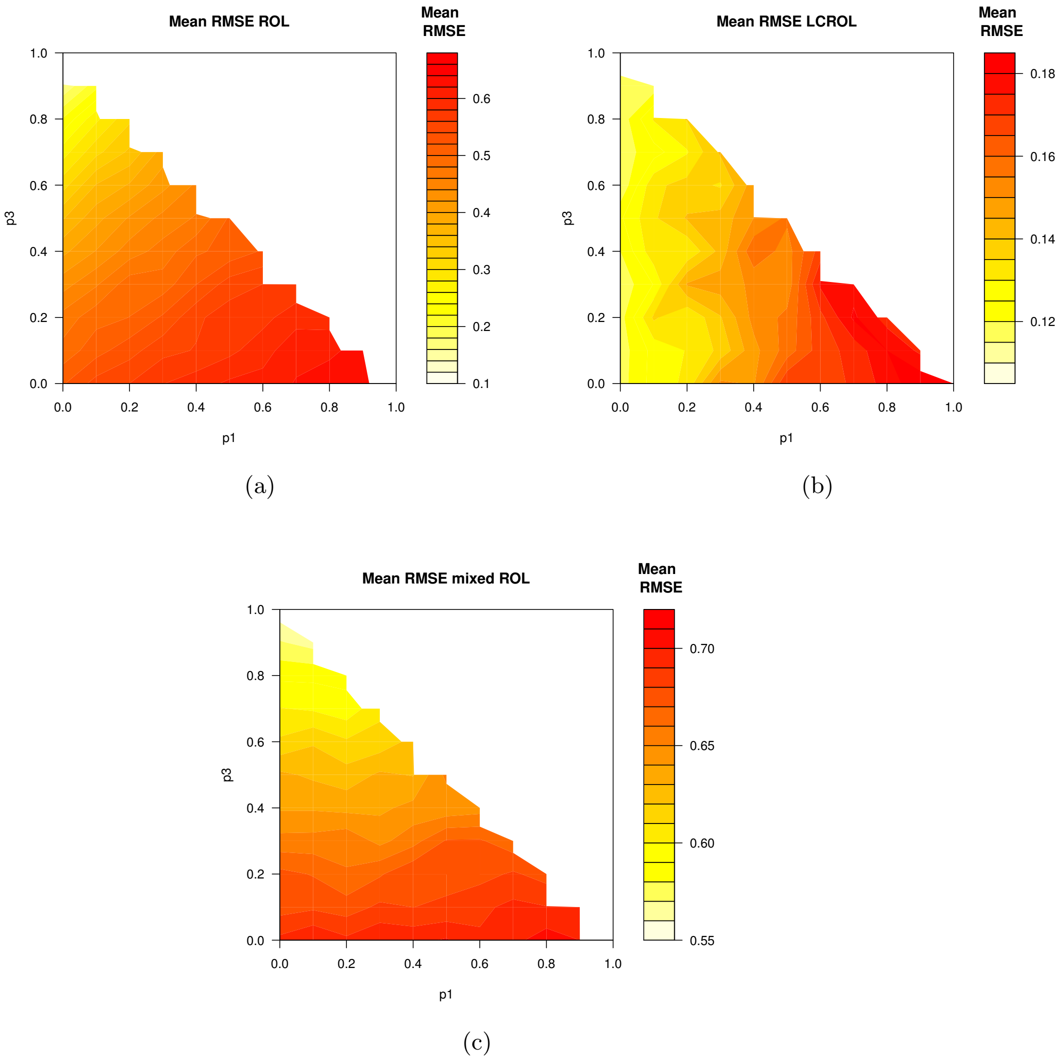

Figure 1: Contour plots of the ROL (a), LCROL (b) and mixed ROL (c) models with varying degree of unobserved heterogeneity in ranking capabilities.

When we look at the contour plot of the standard rank-ordered logit, figure 1a, we see that the mean RMSE decreases when the probability that one can rank all alternatives correctly increases. This is as we expect, because the standard rank-ordered logit model assumes that all individuals can rank all alternatives correctly. The same conclusion holds for the mixed rank-ordered logit model, the mean RMSE also decreases when p_3 increases. This is also as expected, because the mixed rank-ordered logit model also assumes that each individual can rank all alternatives correctly. However, we see that mean RMSE's resulting from the mixed rank-ordered logit model are higher than the ones from the multinomial logit model, standard rank-ordered logit model and from the latent-class rank-ordered logit model, which makes the mixed-rank ordered logit model a bad choice when there only is heterogeneity in ranking capabilities. We can also conclude that we only prefer the rank-ordered logit model above the multinomial logit model when there is a high probability that all individuals can rank all alternatives correctly.
When we look at the latent-class rank-ordered logit model's contour plot, we see that this model performs better compared to the standard and mixed rank-ordered logit models

especially when p_1 is low. Also, the latent-class rank-ordered logit model performs well when p_2 is high. Therefore, we conclude that in the latent-class rank-ordered logit model we do not need all individuals to be able to rank all alternatives correctly, as expected. What further draws our attention is the fact that the mean RMSE of the latent-class rank-ordered logit model seems smaller in all cases than the standard rank-ordered logit and mixed rank-ordered logit models. To further analyse this, we constructed another contour plot of the difference in mean RMSE of the standard rank-ordered logit model and the latent-class rank-ordered logit model, in figure 2.

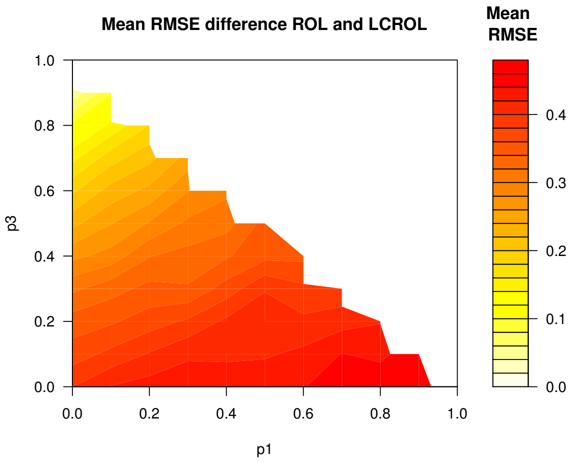

Figure 2: Contour plot of the difference in mean RMSE between the LCROL and ROL models when varying the degree of unobserved heterogeneity in ranking capabilities.

In figure 2 we see that the difference in mean RMSE between the standard rank-ordered logit model and the latent-class rank-ordered model is always positive. We do see that when almost everyone is able to rank all alternatives, the difference becomes small (upper-left corner in figure 2). This is as expected, because then the standard rank-ordered logit and latent-class rank-ordered models are then in fact the same model. However, everything taken into account, we prefer the use of the latent-class rank-ordered logit model above the standard rank-ordered logit model or the mixed rank-ordered logit model when only heterogeneity in ranking capabilities is present.

The contour plot of the latent-class rank-ordered model, figure 1b, also indicates that only when there is a very high probability that all individuals can rank only one alternative correctly, the multinomial logit and the latent-class rank-ordered logit model perform roughly the same. So, we can conclude that when there only is heterogeneity in ranking capabilities the best model one can use is the latent-class rank-ordered logit model.

5.2 Preference heterogeneity simulation

In the second simulation we generate data again according to the data generating process mentioned in section 4, but now with only preference heterogeneity instead of unobserved heterogeneity in ranking capabilities. The results are displayed in table 2.

When we look at the estimates of the multinomial logit model, we see that these estimates are biased. Especially the intercepts show a large bias. The standard rank-ordered logit

Table 2: Estimated mean and RMSE (in brackets) of the parameters over 1000 simulations where each simulation consists of N = 1000 individuals, when only preference heterogeneity across individuals is present

Parameter	True	MNL[a]	ROL[b]	LCROL[c]	Mixed ROL[d]
β_{01}	1.00	-0.56	0.05	1.29	0.13
		(1.57)	(0.96)	(0.36)	(0.87)
β_{11}	0.75	0.53	0.38	0.78	0.50
		(0.24)	(0.38)	(0.17)	(0.26)
β_{21}	-0.30	-0.36	-0.16	-0.44	-0.21
		(0.18)	(0.19)	(0.33)	(0.17)
β_{02}	0.25	-0.93	-0.39	0.44	-0.50
		(1.19)	(0.64)	(0.24)	(0.76)
β_{12}	-0.50	-0.47	-0.25	-0.62	-0.41
		(0.10)	(0.26)	(0.21)	(0.13)
β_{22}	0.45	0.25	0.23	0.37	0.35
		(0.27)	(0.26)	(0.33)	(0.20)
β_{03}	-0.25	-0.90	-0.67	-0.15	-1.05
		(0.68)	(0.43)	(0.27)	(0.81)
β_{13}	1.00	0.53	0.50	1.08	0.74
		(0.48)	(0.51)	(0.21)	(0.28)
β_{23}	0.80	0.41	0.41	0.78	0.64
		(0.43)	(0.42)	(0.34)	(0.26)
$\beta_3\ (\mu)$	-0.75	-0.13	-0.28	-1.22	-0.64
		(0.59)	(0.45)	(0.46)	(0.12)
$\beta_3\ (\sigma)$	1.00	-	-	-	0.84
					(0.17)
p_0	0.00	-	-	0.40	-
				(0.40)	
p_1	0.00	1.00	-	0.00	-
				(0.00)	
p_2	0.00	-	-	0.00	-
				(0.00)	
p_3	1.00	-	1.00	0.60	1.00
				(0.40)	

[a]Multinomial logit model
[b]Rank-ordered logit model
[c]Latent-class rank-ordered logit model
[d]Mixed rank-ordered logit model with estimated standard deviation of β_3 equal to 0.84 with a mean standard error of 0.07

model shows a smaller bias for some of the parameters than the multinomial logit model. We can explain this by the fact that the rank-ordered logit model uses the entire ranking of an individual, which is in this data correctly specified. Looking at the latent-class rank-ordered logit results, we can see that this model also gives biased parameter estimates and performs about as bad as the standard rank-ordered logit model in terms of parameter estimates. What further draws our attention is the fact that the latent-class rank-ordered logit model estimates that p_0 equals 0.40 and that p_3 equals 0.60. So, according to the latent-class rank-ordered logit model there are individuals who cannot rank any of the

alternatives correctly. However, in this data all individuals can rank all alternatives correctly and only preference heterogeneity is present. This lets us draw the same conclusion as in the previous simulation, that it is difficult to differentiate the modelling of heterogeneity in ranking capabilities from the preference heterogeneity. We do not construct another latent-class rank-ordered logit model with p_1 and p_2 equal to zero, because we then end up with the standard rank-ordered logit model.
To have a deeper look into the estimated value of p_0 by the latent-class rank-ordered logit model, we make a plot of estimated p_0 values when the degree of preference heterogeneity (the standard deviation of the β_3 parameter) varies. This plot is displayed in figure 3.

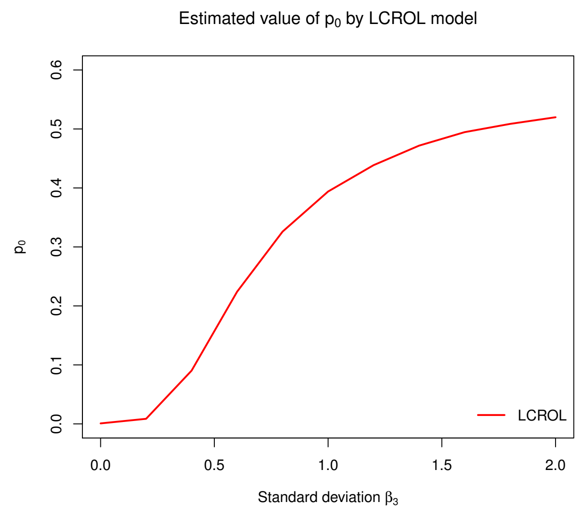

Figure 3: Estimated value of p_0 when standard deviation of β_3 varies

From figure 3 we conclude that when the degree of preference heterogeneity increases, then the probability that no one can rank any of the alternatives correctly estimated by the latent-class rank-ordered logit model also increases, while in this setting this probability equals zero. When the standard deviation of the β_3 variable equals 2.0 the estimated probability that no one can rank any of the alternatives correctly is even about 0.50. So, one needs to be careful to conclude that there is heterogeneity in ranking capabilities when the latent-class rank-ordered logit model indicates this. It could very well be that there only is preference heterogeneity across individuals.

When looking at the results of the mixed rank-ordered logit model in table 2, we see that this model also gives us biased estimates. This bias could be caused by we generate the data and in particular at the values of the alternative-specific variable. Also, the mixed rank-ordered logit model estimates a significant, at the 5% significance level, standard deviation of the β_3 parameter equal to 0.84. In this simulation, where only preference heterogeneity is present, there is no model which performs clearly better than the other models when looking at the parameter estimates and corresponding RMSE's.

Next, we further investigate the sensitivity in parameter estimates of the different models for preference heterogeneity. We do this by generating data in the exact same way as the simulation, but we now vary the standard deviation of the β_3 parameter. The standard deviation ranges from 0 to 2.0 with increments of 0.2. With every standard deviation we generate data, and estimate each model. As performance measure we, again, take the mean RMSE over the β parameters. We also generate ten simulations with 1000 individuals each for every value of the standard deviation of β_3. The results are displayed in figure 4.

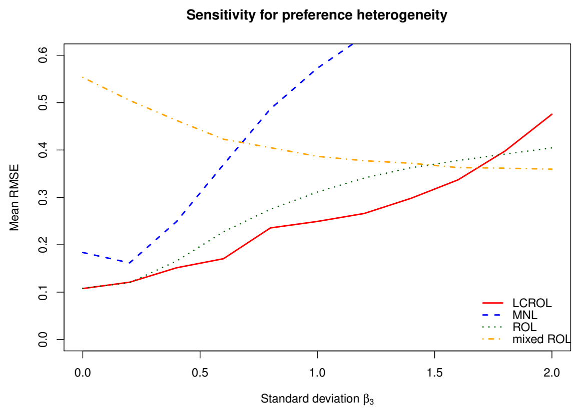

Figure 4: Sensitivity of different models for preference heterogeneity in the data

When the standard deviation of β_3 is equal to zero, there is no preference heterogeneity and no heterogeneity in ranking capabilities present in the data. All individuals can rank all alternatives correctly. Therefore the latent-class rank-ordered logit model and the standard rank-ordered-logit model are the same and perform best. We can also see this in the figure. The graphs of the latent-class rank-ordered logit and standard rank-ordered logit model are equal when there is a small amount of standard deviation of β_3 and thus a low degree of preference heterogeneity. When the standard deviation increases the latent-class rank-ordered logit model gives the best performance of the used models. Only when the degree of preference heterogeneity is pretty strong (when standard deviation is more than about 1.6) the mixed rank-ordered logit model performs best in terms of mean RMSE over the β parameters. The multinomial logit model performs the worst when the degree of preference heterogeneity increases. We first see a decrease in mean RMSE of the multinomial logit model when the degree of preference heterogenity increases, but this is likely caused by the fact that we only performed ten simulations for each model. To conclude, the latent-class rank-ordered logit model is also the best model in terms of mean RMSE over the β parameters when there is a moderate degree of preference heterogeneity, and when the degree of preference heterogeneity increases the mixed rank-ordered logit model seems the best option.

5.3 Application on gaming platform preferences

We now apply the different models on a real-life dataset. Therefore, we use the preferences of Dutch students on gaming platforms. Using this real-life setting, we can investigate whether it is now also difficult to differentiate heterogeneity in ranking capabilities from preference heterogeneity. In table 3 the parameter estimates resulting from the multinomial logit, rank-ordered logit, latent-class rank-ordered logit and mixed logit models using the game dataset are displayed.

As explanatory variables we use an intercept, the weekly numbers of hours spent on gaming and an alternative-specific dummy variable which equals 1 if an individual owns a gaming platform and 0 otherwise. For the use of the mixed rank-ordered logit model we assume that the ownership coefficient differs across individuals and that this parameter follows the normal distribution. As base category we take the PC.

Table 3: Parameters estimates and standard errors (in brackets) of the MNL, ROL, LCROL and mixed ROL models using the gaming platform preferences data

Parameter	MNL	ROL	LCROL	Mixed ROL
Xbox (*intercept*)	0.87	1.37	1.48	1.43
	(0.49)	(0.29)	(0.48)	(0.32)
Playstation (*intercept*)	0.49	0.88	1.05	0.94
	(0.46)	(0.27)	(0.45)	(0.30)
PSP (*intercept*)	-0.07	0.74	0.39	0.79
	(0.59)	(0.28)	(0.48)	(0.32)
GameCube (*intercept*)	0.46	-0.03	-3.48	0.02
	(0.59)	(0.30)	(1.36)	(0.32)
GameBoy (*intercept*)	-1.47	0.03	-2.71	0.08
	(1.00)	(0.29)	(1.20)	(0.31)
Xbox (*hours*)	-0.09	-0.17	-0.13	-0.18
	(0.07)	(0.05)	(0.06)	(0.05)
Playstation (*hours*)	-0.10	-0.12	-0.10	-0.14
	(0.07)	(0.04)	(0.06)	(0.05)
PSP (*hours*)	-0.10	-0.23	-0.35	-0.24
	(0.11)	(0.05)	(0.12)	(0.06)
GameCube (*hours*)	-0.39	-0.18	-0.01	-0.19
	(0.24)	(0.05)	(0.14)	(0.06)
GameBoy (*hours*)	-0.06	-0.23	-0.23	-0.24
	(0.18)	(0.05)	(0.14)	(0.06)
Ownership (μ)	1.74	0.95	1.69	0.97
	(0.38)	(0.19)	(0.36)	(0.21)
Ownership (σ)	-	-	-	0.75
				(0.74)
p_0	-	-	0.23	-
			(0.07)	
p_1	1.00	-	0.19	-
			(0.08)	
p_2	-	-	0.07	-
			(0.08)	
p_3	-	-	0.08	-
			(0.09)	
p_4	-	-	0.00	-
			(0.16)	
p_5	-	1.00	0.44	1.00

When we compare the parameter estimates resulting from the multinomial logit model and the rank-ordered logit model, we see a large difference in parameter estimates. This gives the suggestion that not all individuals could rank all alternatives correctly. When we look at the results of the latent-class rank-ordered logit model, we see that this model suggests that there is heterogeneity in ranking capabilities across individuals. We estimate the probability that an individual cannot rank a single item correctly at 23%. However, based on the results of our simulation study we cannot draw the conclusion that there is unobserved heterogeneity across individuals that easily. The results of the mixed rank-ordered logit model give even more reason to doubt about the suggestion that there only is heterogeneity in ranking capabilities. Using this model, the estimated standard deviation of the ownership variable is 0.75 when we assume that all individuals can rank all alternatives correctly. However, this estimated standard deviation is not even significant at the 10% significance level. So, in this real-life application we obviously do not know whether there is heterogeneity in ranking capabilities across individuals, preference heterogeneity or a combination of both present. The different models we use lead to conflicting conclusions. To be sure which kind of heterogeneity is present one have to know more background information which could be obtained by asking more questions.

6 Conclusion

To get reliable parameter estimates of a ranking task in a survey when there is unobserved heterogeneity in ranking capabilities across individuals present, the best model one can use, based on the results of our simulation study, is the latent-class rank-ordered logit model. This model also gives the best parameter estimates when the degree of heterogeneity in ranking capabilities increases. When there is only preference heterogeneity present, one could also use the latent-class rank-ordered logit model when this preference heterogeneity across individuals is not large, otherwise the mixed rank-ordered logit model is the best model one can use. When comparing the modelling of heterogeneity in ranking capabilities and preference heterogeneity across individuals, we find that the mixed rank-ordered logit model indicates the presence of preference heterogeneity when there only is heterogeneity in ranking capabilities present in the data. Also, the latent-class rank-ordered logit model indicates that there is heterogeneity in ranking capabilities present while there only is preference heterogeneity present in the data. When the degree of preference heterogeneity increases, the latent-class-rank-ordered logit model also estimates an ever increasing probability that nobody can rank any of the alternatives correctly. In the application of the models on real-life data, the latent-class rank-ordered logit model and mixed rank-ordered logit model lead to a contradictory conclusion on which kind of heterogeneity is present. Therefore, when you investigate real-life data you have to look further into the data when you search for which kind of heterogeneity is present in the data. You could also ask someone to indicate if he or she had difficulties filling in the ranking task.

In our research we only incorporated the modelling of heterogeneity in ranking capabilities and preference heterogeneity only once at a time in a model. To make a model which can deal with both types of heterogeneity, one could construct a mixed latent-class rank-ordered logit model. We also evaluate the models by computing the root mean squared error (RMSE) over the estimated parameters. To get further insights into the model performance one could use other performance measures and also compare the prediction accuracy of the models. Another limitation of this research is that we only use a relatively small amount of simulations to evaluate the models. One could increase this number and check if the same conclusions hold. Finally, one could simulate data in a different way. In our simulation the alternative-specific variable have a large contribution to the utility obtained by an individual compared to the individual-specific variables. In further research, one could change this and see whether this changes the parameter estimates, especially in the simulation with only preference heterogeneity where all models are not that accurate.

References

Beggs, S., Cardell, S., & Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Econometrics*, *17*(1), 1–19.

Busse, L. M., Orbanz, P., & Buhmann, J. M. (2007). Cluster analysis of heterogeneous rank data. In *Proceedings of the 24th international conference on machine learning* (pp. 113–120).

Calfee, J., Winston, C., & Stempski, R. (2001). Econometric issues in estimating consumer preferences from stated preference data: a case study of the value of automobile travel time. *Review of Economics and Statistics*, *83*(4), 699–707.

Fok, D., Paap, R., & Van Dijk, B. (2012). A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. *Journal of Applied Econometrics*, *27*(5), 831–846.

Hajivassiliou, V. A., & Ruud, P. A. (1994). Classical estimation methods for ldv models using simulation. *Handbook of Econometrics*, *4*, 2383–2441.

Hausman, J. A., & Ruud, P. A. (1987). Specifying and testing econometric models for rank-ordered data. *Journal of Econometrics*, *34*(1-2), 83–104.

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior.* Institute of Urban and Regional Development, University of California.

Murphy, T. B., & Martin, D. (2003). Mixtures of distance-based models for ranking data. *Computational Statistics & Data Analysis*, *41*(3), 645–655.

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, *80*(4), 647–657.

Train, K. (2000). Halton sequences for mixed logit. *Department of Economics, UCB*.

A Appendix

Table 4: Distribution of the rank of the four alternatives averaged over the 1000 simulations for the simulation when only unobserved heterogeneity in ranking capabilities across individuals is present

	1^{st}	2^{nd}	3^{rd}	4^{th}
Alternative 1	196	344	276	184
Alternative 2	113	231	332	324
Alternative 3	92	171	303	434
Alternative 4	599	254	89	58

Table 5: Distribution of the rank of the four alternatives averaged over the 1000 simulations for the simulation when only preference heterogeneity across individuals is present

	1^{st}	2^{nd}	3^{rd}	4^{th}
Alternative 1	303	273	264	160
Alternative 2	268	251	243	238
Alternative 3	162	200	244	394
Alternative 4	267	276	249	208

Table 6: Distribution of the rank of the six gaming platforms

	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	6^{th}
Xbox	18	26	25	10	9	2
Playstation	17	27	19	11	9	7
PSP	7	9	18	28	16	12
GameCube	7	9	9	16	19	30
GameBoy	2	6	8	15	31	28
PC	39	13	11	10	6	11