Master thesis Econometrics and Management Science

RISK BEHAVIOUR OF CHILDREN MEASURED WITH THE COLUMBIA CARD TASK

May 9, 2017

Nienke F.S. Dijkstra
Student number: 356174
nienkefs.dijkstra@gmail.com

Supervisor: Prof. dr. Patrick J.F. Groenen
Co-reader: Prof. dr. Richard Paap
Erasmus Universiteit Rotterdam
Erasmus School of Economics

Abstract

Risk behaviour in an experimental risk task is a key predictor of real life behaviour and the development of a child. Existing research shows that smokers and gamblers are risk seeking on experimental risk tasks. The present study aimed to predict risk behaviour of children by means of the Columbia Card Task (CCT). The analysed dataset consists of 3326 children of around nine years old. This study is embedded in the Generation R program that investigates children's growth, development, and health from foetal life onwards. A finite mixture model with a censored Poisson regression is implemented to predict the number of cards these children turn over. Additionally, with this model the influence of the loss probability and gain and loss amount on the number of cards turned over is analysed. The results show that children with a low socioeconomic status perform worse on the CCT and are more risk seeking than children with a high socioeconomic status. Additionally, a difference between boys and girls is revealed. Boys tend to be more risk averse and score higher on the CCT than girls. Regarding the game settings, the loss probability and loss amount have a significant negative effect on the number of cards turned over. Besides, these factors seem to be more important to boys than to girls.

Keywords: Columbia Card Task, GenerationR, Risk Behaviour, Censored Poisson Regression, Finite Mixture Model

Supercomputing computations for the finite mixture model are supported by the NWO Physical Sciences Division (Exacte Wetenschappen) and SURFsara (Lisa compute cluster, www.surfsara.nl).

Contents

List of Figures

List of Tables

1. Introduction

Risk behaviour measured in an experimental environment could be an indicator for real-world risk behaviour such as smoking, gambling, and delinquency. Lejuez et al. (2002) argue that behavioural tasks are a useful addition to self-reported real-world risk behaviour. For example, Lejuez et al. (2003) prove that smokers and nonsmokers score significantly different in the risk task BART. Likewise, Collins et al. (1987) show the relationship between risk taking/rebelliousness and smoking at an older age. It is advantageous for prevention to correctly predict at a young age the likelihood of smoking. Similarly, risk tasks could be useful for the prevention of gambling and delinquency.

From another angle, risk behaviour is an important indicator for entrepreneurship. Entrepreneurs are known to be people who are willing to take risks. Chye Koh (1996) concludes that entrepreneurially minded people score higher on risk tasks than nonentrepreneurially minded people. However, other authors argue that this does not mean that entrepreneurs are high risk takers. According to Caird (1991), entrepreneurs are not gamblers, but take calculated risks. Miner (1990) goes even further and argues that a key entrepreneurial task is to avoid risk. Based on these statements one could hypothesize that entrepreneurs are successful in experimental risk tasks.

Measuring the risk behaviour at a young age is highly favourable. From the medical point of view an adequate prevention scheme can be adopted for children with a high risk of smoking, drinking, et cetera. Also within entrepreneurship it is advantageous to detect entrepreneurial minded children at a young age. In this way children can get the accompaniment they need to become a successful entrepreneur.

This paper focusses on risk behaviour of children and seeks for a relationship between risk behaviour and the IQ and socioeconomic status of children. Previous research has shown that the socioeconomic status is associated with risky behaviour (e.g. smoking, a sedentary lifestyle, and unhealthy eating) that may lead to chronic diseases (Lowry, Kann, Collins, & Kolbe, 1996). It is also shown that pathological gambling appears more often in the low socioeconomic classes (Welte, Barnes, Wieczorek, Tidwell, & Parker, 2002). However both studies are based on adults and not on children.

For this study risk behaviour is measured with the Columbia Card Task (CCT). In this card game money can be won or lost by turning over cards. Due to the design of the task it is possible to measure, next to risk aversion, the sensitivity to reward and punishment. What this research sets apart from previous literature on the CCT is that it focusses on risk behaviour of children instead of adults. Secondly, most studies analyse the outcome of the CCT on aggregated level (average per individual or per game setting). With this procedure you lose a lot of detailed information, for example how individuals react on previous gains or losses.

The design of the CCT in this study leads to censoring, therefore a censored Poisson regression is implemented to deal with the censoring and the counted dependent variable. To account for unobserved heterogeneity across individuals, individuals are divided into several homogeneous segments. This procedure is called Finite Mixture Modelling. The segment-specific constant, β-coefficients, and segment probabilities are estimated with the EM-algorithm.

Lastly, this study is embedded in the Generation R Study, a large population based multi-ethnic birth cohort, investigating children's growth, development, and health from foetal life onwards (Tiemeier et al., 2012). For this study children and their mothers are followed from the prenatal phase onwards. The data collection is intense by multiple surveys and biological and observational assessments. The CCT is one of the observational assessments that was conducted on nine-year-old children. The total cohort includes almost ten thousand children. The current study sample consists of 4551 children who participated in the CCT.

The remainder of this paper is structured as follows. The next chapter discusses related risk tasks and summarizes the available literature on the CCT. Chapter 3 provides an extensive description of the design of the CCT, explains the data cleaning process, and gives descriptive statistics of the data. Next to tables with summary statistics, graphs are included to show the distribution of several variables. Chapter 4 starts with a description of a censored Poisson model, and builds up to a finite mixture model. Additionally, the EM-algorithm and a decomposition of the model coefficients are elaborated. Chapter 5 presents the results of a linear regression on the average score on individual level and describes the results of the censored Poisson regression with finite mixtures. Lastly, Chapter 6 gives a conclusion and discusses the limitations.

2. Literature review

Risk behaviour is a widely studied phenomenon. Multiple tests are introduced to measure riskiness under experimental conditions. The Balloon Analogue Risk Task (BART), introduced by Lejuez et al. (2002) is similar to the Columbia Card Task (CCT). At every iteration the participant has the choose to continue or to stop and collect the points. An important difference is that in the CCT the probabilities are explicitly stated whereas in the BART they are not.

Other risk taking tasks are based on gambling. The Cambridge Gambling Task (Rogers et al., 1999) and the Game of Dice Task (Brand et al., 2005) are two examples where the participant has to bet on the outcome. In both tasks the riskier option has next to a higher maximal payoff and a lower minimal payoff, also a lower expected value compared to the less risky option. Consequently, the reason for choosing the less risky option is hard to distinguish. It could be based on either risk avoidance or the higher expected value. Similarly, choosing the riskier option can be driven by greater risk seeking or by increased reward sensitivity.

Figner and Weber (2011) claim that risk behaviour is often domain specific and depends on psychological processes. The authors argue that in laboratory experiments adolescents show no significant difference from other age groups in risk behaviour, whereas in real life adolescents take great risks in many domains (e.g. substance use, dangerous driving, and unsafe sex). According to the authors this difference could be explained by the fact that laboratory experiments make no distinction between affective and deliberative decision making. The CCT distinguishes between these two psychological processes. Moreover, Figner, Mackinlay, Wilkening, and Weber (2009), the inventors of the CCT, show that adolescents are riskier in affective situations, while there is no difference among age groups in deliberative situations.

The Columbia Card Task (CCT) is a relative new card game whereby participants can win or lose money by turning over cards. Originally there are two versions of the CCT: a hot CCT to measure affective decision making and a cold CCT for deliberative decision making. The most important difference is the time of the feedback. During the hot CCT the participant gets immediate feedback. After clicking on a card he or she immediately sees whether it is a gain or loss card. With the cold CCT the participant has to select a number of cards he or she wants to turn over before the first card is shown. At the end of the game it is revealed whether a loss card was faced or not. Note that this research only focusses on the outcome of a hot CCT.

A major advantage of the CCT, over other dynamic risk tasks, is that the game settings vary across trial. The gain amount, loss amount, and number of loss cards all can take two values, which results in eight different combinations of game settings. Since all different game settings are played at least once per respondent, a researcher could assess the influence of the loss probability on risk behaviour and the sensitivity to reward and punishment per respondent separately. Previous research has shown that individuals may react differently on these three parameters.

Penolazzi, Gremigni, and Russo (2012), for instance, investigate whether personality traits influence risk behaviour and whether this influence depends on the context of the decision making process (affective or deliberative decision making). A remarkable finding is that in the hot CCT only the loss amount and number of loss cards are significantly associated with the number of cards turned over, whereas in the cold CCT all three game settings seem to play a role. This finding is supported by the studies of Kluwe-Schiavon et al. (2015) and Holper and Murphy (2014). With an adaptation of the CCT to Brazilian Portuguese, Kluwe-Schiavon et al. (2015) find that in all CCT conditions (hot, warm, and cold) the number of loss cards is most frequently used, followed by the loss amount and gain amount (the latter had no significant effect in the hot CCT). In relation to the personality traits, Penolazzi et al. (2012) conclude that high sensation seekers on average turn over more cards than low sensation seekers in the hot CCT, however the main effect of sensation seeking in a linear regression is not significant. Besides, they find an interaction between BAS-reward responsiveness and both gain amount and loss amount in the hot CCT.

Contrary, Buelow (2015) finds no significant relationship between state mood, impulsive sensation seeking, and BIS/BAS (includes the BAS-reward responsiveness subscale) and the performance on the hot CCT. She finds, however, a negative correlation between working memory (measured with the Digit Span backward) and risk taking on the hot CCT. Additionally, Buelow shows that participants with a high score on the Digit Span backward as compared to participants with a low score, turn over less cards when the loss amount is 250 and the number of loss cards is three by any gain amount. Besides, she concludes that participants pay more attention to the game settings in the cold CCT than in the hot CCT. In particular, only the number of loss cards seems to be relevant in the hot CCT, which partly contradicts the findings of Penolazzi et al. (2012) and Kluwe-Schiavon et al. (2015).

Another interesting association was discovered by Konnikova (2013). She compares the performance of high and low self-controllers in both a stress and non-stress condition in the hot CCT. The stress condition was created by limiting the amount of time for completing a trial. In both conditions on the CCT high self controllers turn over more cards than low self-controllers. In addition, Konnikova finds that high self-controllers do not use the information of the game settings. Low self-controllers, on the other hand, turn over significantly less cards when facing a higher loss amount.

According to Huang, Wood, Berger, and Hanoch (2015) the information use on the CCT is associated with age. They suggest that the lower deliberative capacities of older adults account for the lower amount of information use. This study, however, is conducted on young adults (mean age = 24.5) and older adults (mean age = 75.3) and provides no information on children. To my knowledge, van Duijvenvoorde et al. (2015) is the only study on the CCT that includes children. In this study the risk and return sensitivity are compared among three age groups; children (8-10 years), adolescents (16-19 years), and adults (24-34 years). Both risk aversion and return sensitivity increase with age. Interestingly, children appear to be risk insensitive on average. However, there was a large difference in performance on the CCT among children. In addition, the authors argue that all age categories understood the task information, because the effects of the game settings were significant and in the expected direction for all age categories.

Unlike the other literature on the CCT, Pripfl, Neumann, Köhler, and Lamm (2013) do not look at personal characteristics, but compare the risk behaviour of smokers and non-smokers. The results show that smokers significantly take more risk on the CCT than non-smokers. In particular, smokers as compared to non-smokers take more risk in trials with many loss cards and in trials where a high gain amount is combined with a low loss amount. It is likely that children who have not started smoking yet but will do so at a later age, show the same pattern and take more risk on the CCT.

3. Data description

The current study is embedded in the Generation R Study, a large population based multi-ethnic cohort study (Tiemeier et al., 2012). The Generation R Study was designed to analyse early environmental and genetic determinants of growth, development, and health from foetal life until young adulthood. All pregnant women living in Rotterdam, The Netherlands, with a delivery date between April 2002 and January 2006 were invited to participate. In total, 9778 pregnant women enrolled in the study. During pregnancy several maternal, paternal, and familial characteristics were collected. At several time points, parents filled in questionnaires about the development of their child. Additionally, at the age of six and nine children were invited to participate in various tests and assessments. One of these tests is the Columbia Card Task, which is conducted on nine-year-old children.

The next section elaborates the design of this task. Section 3.2 summarizes the data cleaning process. In Section 3.3 the distributions of important variables are visualized in graphs and a table with summary statistics of most variables is presented. Lastly, Section 3.4 explains how is dealt with missing data.

3.1 Columbia Card Task

The Columbia Card Task[1] (CCT) is a card game that measures riskiness. Figure 3.1 shows the layout of the game. Participants could win or lose money by turning over cards. The hot CCT played by the children in this research is closely related to the original one (Figner et al., 2009). There are 32 cards, divided in win and loss cards. By turning over a win card the participant earns points and by turning over a loss card he or she loses points and the game ends. At every step the participant has the choice between turning over a(nother) card or hitting the stop button to (voluntarily) stop this game. It is also possible to stop immediately without turning over any card. After a game has ended the earned points are summed and a potential loss amount is subtracted from it.

The values of the win and loss cards can vary per trial, even as the number of loss cards. The gain amount varies between ten and thirty, the loss amount between 250 and 750, and there are either one or three loss cards in a trial. Note that, in contrast to Figner et al. (2009), the loss cards are randomly distributed over the 32 cards and hence the game is not manipulated. These three parameters lead to eight different game settings and within a block of eight trials the sequence of the game settings is random. Every participant plays at least two blocks of eight trials[2]. In other words, every game setting is played at least two times. Because of the different game settings, the CCT measures next to riskiness also the complexity of information use and the sensitivity to reward and punishment. With the three parameters (gain amount, loss amount, and number of loss cards) it is possible to assess which of these three parameters affects participants' choices.

After a participant has played all trials, three trials are randomly selected and are paid out in real money. The participant has a start value of 200 cents (i.e. 2 euro) and

[1] See columbiacardtask.org for an example.

[2] During the data collection they decided to shorten the test. Instead of three blocks and 24 trials, two blocks and 16 trials were played.

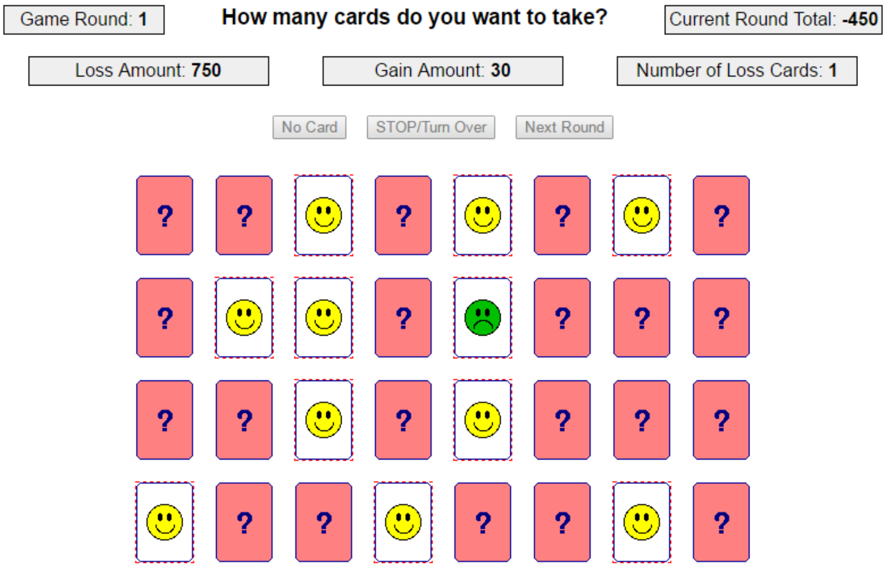

Figure 3.1: Screenshot of the Columbia Card Task

the total points of the selected trials will be added or deducted from this start value. A participant only receives money and does not have to pay any losses.

3.2 Data cleaning process

From the 9778 children who started in the Generation R Study 4551 children participated in the CCT, see Figure 3.2 for a flowchart of the data cleaning process. From these children only 4538 completed the task, that is, completed all 16 or 24 trials. To get a balanced data set only the first 16 trials of every participant are analysed, hence the data set contains 72608 trials. Some children clearly did not understand the game and turned over the 32^{nd} card (or 30^{th} card in case of three loss cards), which must be a loss card. The 104 children who did this are excluded from the data set. Furthermore, only 75% of the observations in the sample are analysed in this study (3326 children), such that in later studies the other 25% could still be used to test the model performance.

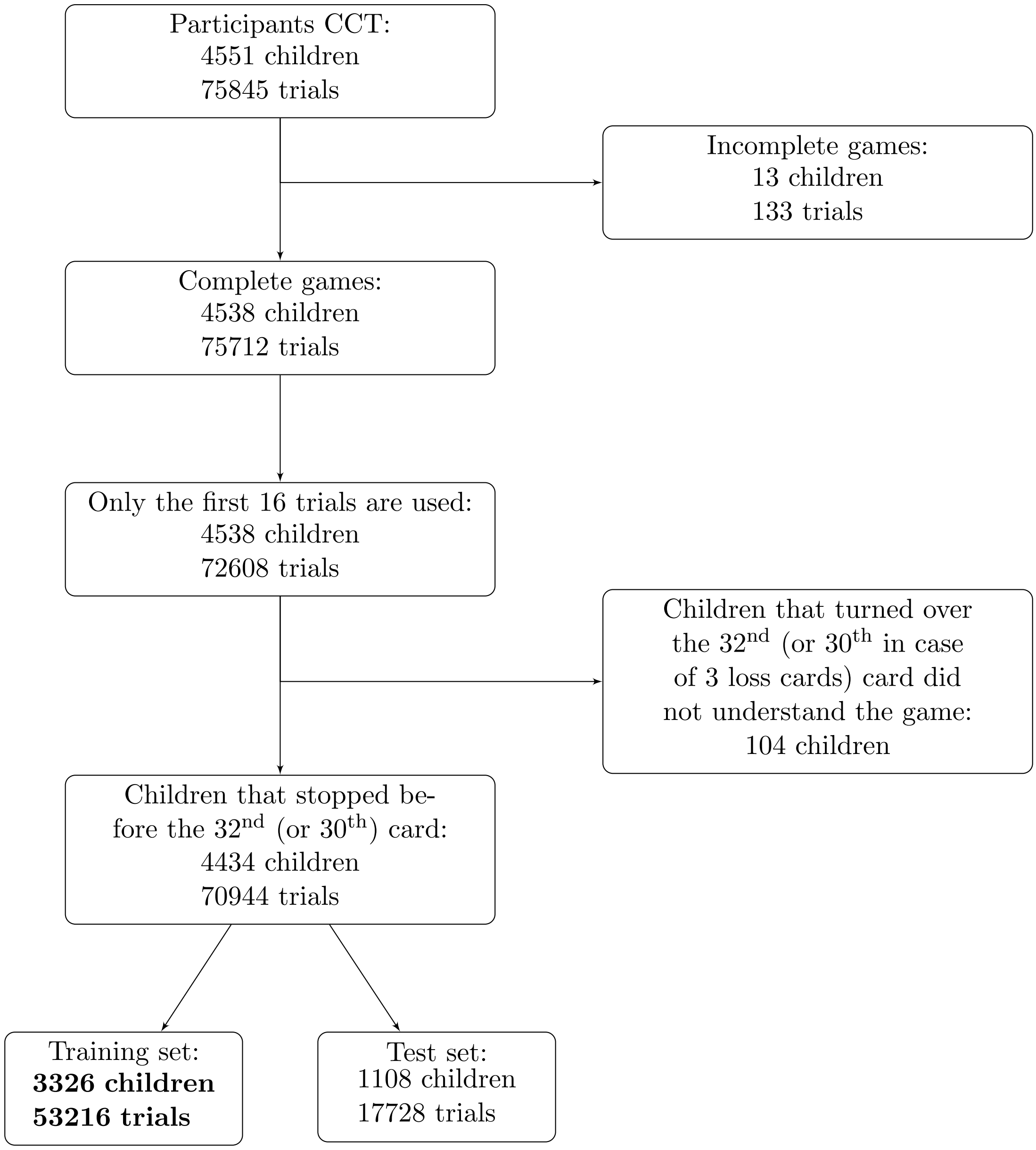

Figure 3.2: Flowchart of the data cleaning process

3.3 Data characteristics

When a loss card is turned over the trial is censored, because this terminates the game and it is unknown how many more cards the participant would have turned over when he or she did not face a loss card. More than two-third of the trials is censored, 48982 trials, and for some children every observation is censored (i.e. they faced a loss card in every trial). Figure 3.3 gives the distribution of the number of censored trials per child. As you can see, most children face around twelve loss cards.

Table 3.1 gives the average number of cards the participants turned over per game setting. For these statistics only the noncensored trials are included. As expected, the average number of cards turned over decreases by an increase in risk. It also decreases by a higher loss amount, although with a smaller magnitude. Furthermore, an increase in

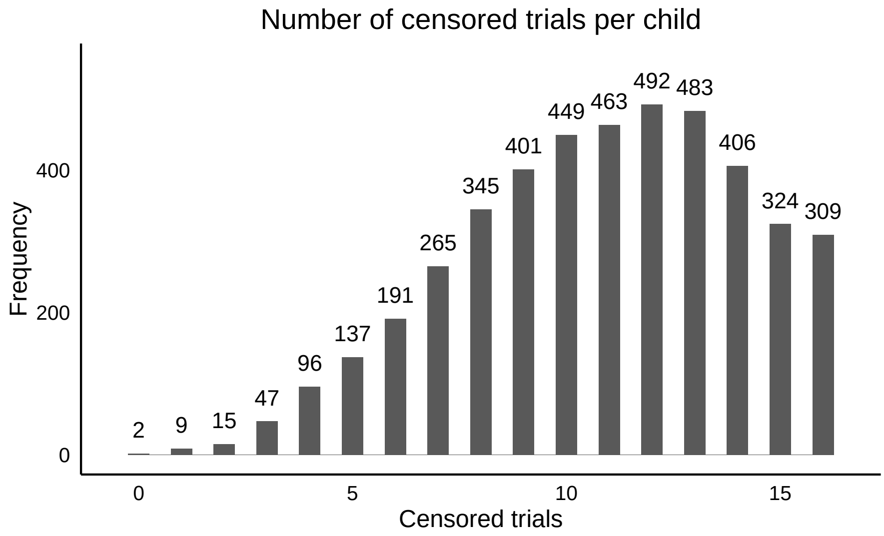

Figure 3.3: Histogram of censored trials per child

gain amount leads to a small decrease in number of cards turned over. Table 3.2, on the other hand, presents the number of cards that a risk neutral person would turn over (i.e. maximizes the expected value). Except for the less profitable setting (gain amount 10, loss amount 750, and 3 loss cards), the average of the participants is lower for all game settings, meaning that these participants are on average risk averse.

Table 3.1: Average number of cards turned over for the noncensored trials per game setting

1 loss card

		loss amount 250	loss amount 750
gain	10	11.6	11.2
amount	30	10.7	10.6

3 loss cards

		loss amount 250	loss amount 750
gain	10	7.1	6.7
amount	30	6.9	6.7

Table 3.2: Optimal number of cards to turn over when maximizing the expected value

1 loss card

		loss amount 250	loss amount 750
gain	10	16	16
amount	30	16	16

3 loss cards

		loss amount 250	loss amount 750
gain	10	10	0
amount	30	8	10

Adding the earned points and subtracting potential losses results in a score per trial. The distribution of the average score over all trials per child is presented in Figure 3.4. The mean of this distribution is -165, which means that most children ended with a negative score over all trials and this game was not profitable for them (in terms of earning money). A risk neutral strategy would on average result in a score of 92.

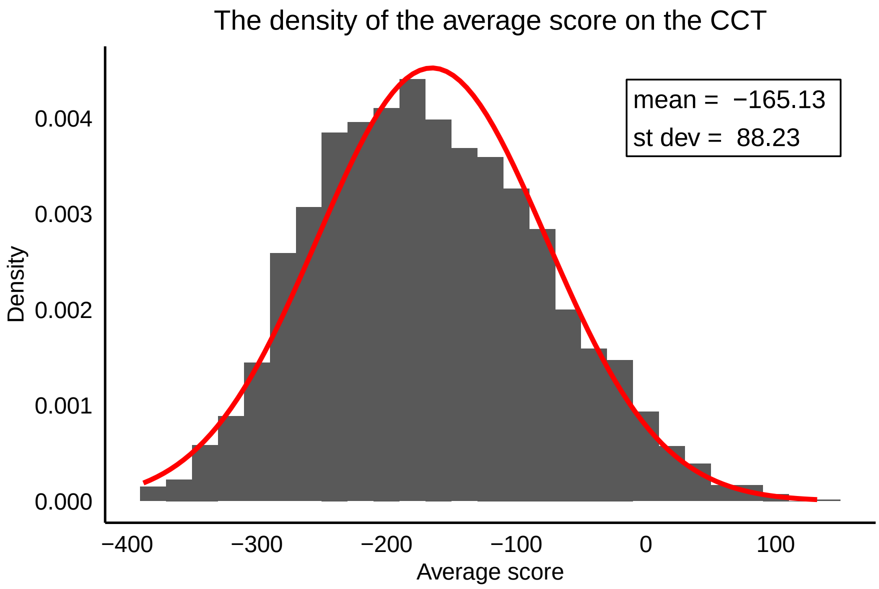

Figure 3.4: A density histogram of the average score of children

Table 3.3 gives a summary of the data analysed in this paper. As mentioned earlier 4434 children completed and understood the CCT. A little bit more than half of these children are girls (50.7%). The average age is 9.8 years (±0.26) and the average IQ is 102 (±14.7). The latter is measured with the SON-R 2.5-7 at the age of six. Information about the mother is available in ethnicity and education. Most mothers have a Dutch or Other Western nationality. Other common ethnicities are Surinamese, Turkish, and Moroccan. Regarding the education level, most mothers have a degree from secondary school or higher. Furthermore, the household income for most families is above €2000 per month.

Table 3.3: Summary of the data

Participants	4434
Censored trials (%)	67.5
Sex (% girl)	50.7
Age (years)	9.8 (±0.26)
IQ (N = 3830)	102 (±14.7)
Ethnicity mother (%, N = 4331)	
Dutch	59.8
African	4.8
Asian, non Western	5.7
Moroccan	4.3
Dutch Antilles	2.1
Surinamese	7.1
Turkish	5.9
Other Western	10.1
Education mother (%, N = 4085)	
No education or primary education	6.7
Secondary education	42.2
Higher education	51.2
Household income per month in euro's (%, N = 3652)	
<2000	20.5
2000 - 4000	43.8
>4000	35.7

3.4 Data imputation

Some observations have missing values in the individual characteristics, such as the ethnicity (2% missing) and education (13% missing) of the mother, household income (18% missing), and IQ (14% missing) of the child. These values are imputed with predictive mean matching (PMM), which is applicable when dealing with categorical variables. The general idea of single imputation with PMM is that missing values are filled with data from similar participants. Let $\boldsymbol{x}$ be a variable with missing values and $\boldsymbol{Z}$ be a set of variables without missing data for the cases that are missing in $\boldsymbol{x}$. Estimate regression coefficients $\boldsymbol{b}$ by regressing the complete cases of $\boldsymbol{x}$ on $\boldsymbol{Z}$. Then, randomly draw a set of new coefficients $\boldsymbol{b}^*$ from a multivariate normal distribution with mean $\boldsymbol{b}$ and the estimated covariance matrix of $\boldsymbol{b}$. Generate predicted values for all cases in $\boldsymbol{x}$ with $\boldsymbol{b}^*$. For each case with missing data in $\boldsymbol{x}$, select a set from the cases with observed data in $\boldsymbol{x}$ whose predicted values are close to the predicted value of the case with missing data. From the cases with observed values randomly choose one case and assign its observed value to the case with missing value. In this analysis the set $\boldsymbol{Z}$ contains the age, gender, weight at birth, and IQ of the child, and the age at delivery, ethnicity, and education of the mother, and the household income.

4. Methodology

This chapter is concerned with the methods and techniques applied in this research. Since the variable of interest, number of cards turned over, is a nonnegative integer and is censored the data is modelled with a censored Poisson regression. Besides, the censored Poisson regression is extended with a finite mixture model to account for the unobserved heterogeneity across individuals. Individuals are assigned with a certain probability to a segment such that the difference within segments is as small as possible and the difference between segments is as large as possible. For each segment the model contains a different intercept. Commonly, the parameters of a finite mixture model are estimated with the EM-algorithm, a method that is often used in case of missing data. Lastly, this chapter covers a technique to decompose the coefficients vector and obtain coefficients for all levels of the categorical variables.

4.1 Censored Poisson Regression

The variable of interest, number of cards turned over, is a count variable from 0 to 31 (or 29 in a trial with three loss cards). Because the Poisson regression has a discrete nonnegative distribution, this fits the data better than a classical regression. The Poisson regression model belongs to the family of generalised linear models, which is discussed in detail by Nelder and Wedderburn (1972).

The Poisson distribution has the following probability mass function,

$$\Pr[Y = y] = \frac{\exp(-\mu)\mu^y}{y!}, \qquad y = 0, 1, 2, ...,$$

where μ is the rate parameter. An important property of the Poisson distribution is equidispersion, that is equality of the mean and variance, $\mathrm{E}[Y] = \mathrm{V}[Y] = \mu$.

To use this distribution in a regression framework the parameter μ is expressed in terms of the covariates and individual subscript i and trial subscript t are added.

$$\mu_{it} = f(\boldsymbol{x}_{it}), \qquad i = 1, ..., N, \quad t = 1, ..., T.$$

It is important to choose $f(\boldsymbol{x}_{it})$ in such way that μ_{it} is nonnegative for all values of $\boldsymbol{x}_{it}$. Note that the Poisson regression is heteroskedastic, because of the equidispersion property, $\mathrm{V}[y_{it} \mid \boldsymbol{x}_{it}] = f(\boldsymbol{x}_{it})$.

The cumulative mass function of the Poisson distibution is added to the model, to accommodate the situation of censored observations. The rationale is that the probability of all outcomes above the censoring value is equal to one minus the probability of all outcomes below and including the censoring value. The likelihood function of a censored Poisson regression is then given by

$$L(\boldsymbol{\mu}) = \prod_{i=1}^{N}\prod_{t=1}^{T}\left[\left(\frac{\mu_{it}^{y_{it}}\exp(-\mu_{it})}{y_{it}!}\right)^{d_{it}}\left(1 - \sum_{k=0}^{C_{it}}\frac{\mu_{it}^{k}\exp(-\mu_{it})}{k!}\right)^{1-d_{it}}\right],$$

where d_{it} is zero if trial t of individual i is censored and one otherwise, and C_{it} is the number of cards turned over including the loss card in that trial.

4.2 Finite Mixture Model

To take account of the heterogeneity across individuals the censored Poisson regression is extended to a finite mixture model (Aitkin & Rubin, 1985). The idea behind this model is that the population consists of a finite number of homogeneous subpopulations. A finite mixture model is more parsimonious than a model with fixed effects, because it has only K, the number of subpopulations, additional free parameters instead of N, the number of individuals. Furthermore, a finite mixture model relies on fewer distributional assumptions than a model with random effects. It is assumed that the subpopulations follow the same distribution, but have different parameters. All individuals are assigned to the subpopulations with a certain probability $\boldsymbol{\pi}$. Inserting this segment probability into the likelihood function results in

$$L(\boldsymbol{\mu}) = \prod_{i=1}^{N} \sum_{s=1}^{K} \pi_s f_s(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{is}, \boldsymbol{x}_i), \tag{4.1}$$

where $f_s(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{is}, \boldsymbol{x}_i)$ is the likelihood contribution of individual i belonging to segment s, given by

$$f_s(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{is}, \boldsymbol{x}_{it}) = \prod_{t=1}^{T} \left[\left(\frac{\mu_{its}^{y_{it}} \exp(-\mu_{its})}{y_{it}!} \right)^{d_{it}} \left(1 - \sum_{k=0}^{C_{it}} \frac{\mu_{its}^{k} \exp(-\mu_{its})}{k!} \right)^{1-d_{it}} \right].$$

Commonly, μ_{its} is defined with multiplications of predictor variables. However, in this research we choose to model the mean μ_{its} by a linear combination of predictor variables as

$$\mu_{its} = \max(\alpha_s + \boldsymbol{x}_{it}'\boldsymbol{\beta}, \epsilon),$$

with ϵ sufficiently close to zero, for $i = 1, ..., N$, $t = 1, ..., T$, and $s = 1, ..., K$. The maximum of $\alpha_s + \boldsymbol{x}_{it}'\boldsymbol{\beta}$ and ϵ is taken to avoid the possibility of a negative μ_{its}. Moreover, a major advantage of this definition compared to the more common one is that this model is linear instead of multiplicative. The interpretation of coefficients in a linear model is more straight forward. Note that the intercept α_s is segment specific and takes K different values. Furthermore, the segment probabilities sum up to one, $\sum_{s=1}^{K} \pi_s = 1$.

4.3 EM-algorithm

Maximizing the likelihood function in (4.1) to estimate the parameters is complicated when maximizing across the entire parameter space. Augmenting the observed data with additional information about the segment memberships would greatly simplify this maximization, because it enables us to compute estimates of the parameters on segment-level. However, the segment membership of an observation is unobserved. The EM-algorithm (Dempster, Laird, & Rubin, 1977) is applicable for computing maximum likelihood estimates from incomplete data.

Assume that the segment memberships are known, then the log complete data likelihood function is given by

$$\ell_c(\boldsymbol{\mu}) = \sum_{i=1}^{N} \sum_{s=1}^{K} I[S_i = s](\log(\pi_s) + \log(f_s(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{is}, \boldsymbol{x}_i))), \tag{4.2}$$

where $I[S_i = s]$ is an indicator function equal to 1 if observation i belongs to segment s. The EM-algorithm iteratively maximizes this log complete data likelihood function.

In the first step of the EM-algorithm, the E-step, the expectation of the log complete data likelihood function in (4.2) is calculated with respect to the unobserved $\boldsymbol{S}$, given the observed $\boldsymbol{y}$ and the preliminary estimate $\boldsymbol{\mu}$. The only stochastic components are the segment memberships denoted by S_i, hence an expression for $\mathrm{E}[I(S_i = s) \mid \boldsymbol{y}_i]$ is needed. The posterior segment probabilities, are defined as

$$p_{is} := \mathrm{E}[I(S_i = s) \mid \boldsymbol{y}_i] = \frac{f_s(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{is}, \boldsymbol{x}_i)\pi_s}{\sum_{j=1}^{K} f_j(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{ij}, \boldsymbol{x}_i)\pi_j},$$

for $i = 1, ..., N$ and $s = 1, ..., K$. Subsequently, the expected log complete data likelihood function can be written as

$$\mathrm{E}[\ell_c(\boldsymbol{\mu}) \mid \boldsymbol{y}] = \sum_{i=1}^{N}\sum_{s=1}^{K} p_{is}(\log(\pi_s) + \log(f_s(\boldsymbol{y}_i \mid \boldsymbol{\mu}_{is}, \boldsymbol{x}_i))). \tag{4.3}$$

In the second step, the M-step, (4.3) is maximized with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\pi}$. Setting the first order derivative with respect to π_s equal to zero results in an update for π_s given by

$$\pi_s = \frac{1}{N}\sum_{i=1}^{N} p_{is} \qquad \forall\ s = 1, ..., K.$$

In words, the estimates of the prior probabilities π_s are equal to the average posterior probabilities in each segment.

The maximization with respect to α_s and $\boldsymbol{\beta}$ is a bit more complicated and hard to solve analytically. Numerical optimization with for example the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm can be used to find estimates for α_s and $\boldsymbol{\beta}$. The BFGS algorithm is an iterative gradient method. In general the procedure of iterative gradient methods is to update the current estimate by a factor based on the gradient,

$$\hat{\boldsymbol{\beta}}_{m+1} = \hat{\boldsymbol{\beta}}_m + \boldsymbol{A}_m \boldsymbol{g}_m,$$

where $\boldsymbol{g}_m$ is the gradient vector evaluated at $\hat{\boldsymbol{\beta}}_m$ and $\boldsymbol{A}_m$ is a $q \times q$ matrix depending on $\hat{\boldsymbol{\beta}}_{\boldsymbol{m}}$. In the Newton-Raphson algorithm $\boldsymbol{A}_m$ is minus the inverse Hessian, whereas the BFGS algorithm uses an estimate for the Hessian. In each step of the BFGS algorithm the change in the gradient is used to obtain a better approximate of the Hessian. Extensive literature on the BFGS algorithm is written by Møller (1993) and Watrous (1988) among others. Since the only criteria in the M-step is an increase in the log likelihood function, only a few maximization steps of the BFGS algorithm are performed. This will presumably fasten the optimization process.

The gradients with respect to α_s and $\boldsymbol{\beta}$ for the BFGS algorithm can be derived analytically. Let $\phi(\mu_{its})$ be the probability mass function of the Poisson distribution, then the gradient vector of α_s is given by

$$\frac{\partial \mathrm{E}[\ell_c(\boldsymbol{\mu})]}{\partial \alpha_s} = \left[\sum_{i=1}^{N} p_{is} \sum_{t=1}^{T} d_{it}\left(\frac{y_{it}}{\mu_{its}} - 1\right) - (1 - d_{it})\frac{\sum_{k=0}^{C_{it}} \phi(\mu_{its})(\frac{k}{\mu_{its}} - 1)}{1 - \sum_{k=0}^{C_{it}} \phi(\mu_{its})}\right]_+$$

for all segments $s = 1, ..., K$ and the gradient vector of $\boldsymbol{\beta}$ is given by

$$\frac{\partial \mathrm{E}[\ell_c(\boldsymbol{\mu})]}{\partial \boldsymbol{\beta}} = \left[\sum_{s=1}^{K}\sum_{i=1}^{N} p_{is} \sum_{t=1}^{T} \left[d_{it}\left(\frac{y_{it}}{\mu_{its}} - 1\right) - (1 - d_{it})\frac{\sum_{k=0}^{C_{it}} \phi(\mu_{its})\left(\frac{k}{\mu_{its}} - 1\right)}{1 - \sum_{k=0}^{C_{it}} \phi(\mu_{its})}\right] \boldsymbol{x}_{it}'\right]_+ .$$

The brackets $[...]_+$ denote that this is the gradient if $\alpha_s + \boldsymbol{x}'_{it}\boldsymbol{\beta} > \epsilon$, otherwise the gradient is equal to zero.

The parameter estimates found in the M-step are used to compute the posterior probabilities p_{is} in the E-step. The E- and M-step are alternated until the likelihood function does not further improve.

The standard errors of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\pi}$ can be easily derived from the diagonal elements of the inverse Hessian matrix of the likelihood function in (4.1). However, take into account the restrictions $\sum_{s=1}^{K} \pi_s = 1$ and $0 \leq \pi_s \leq 1$. Due to the first restriction one can only compute the Hessian for $K-1$ elements of $\boldsymbol{\pi}$. The second restriction could cause problems, because the Hessian is approximated with the R package `NumDeriv` (Gilbert & Varadhan, 2012). To find the numerical Hessian, the function is evaluated at many different values for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\pi}$. All these values should be valid, thus all elements of $\boldsymbol{\pi}$ should lie between zero and one.

These restrictions can be met with a simple trick. Define

$$\pi_s = \frac{\exp(\tau_s)}{1 + \sum_{m=1}^{K-1} \exp(\tau_m)} \qquad \forall\ s = 1, ..., K-1$$

and $\pi_K = 1 - \sum_{s=1}^{K-1} \pi_s$. Without loss of generality, we can choose π_K to be the smallest value of all elements in $\boldsymbol{\pi}$. This definition leads to $K-1$ parameters $\boldsymbol{\tau}$ that can take all real values. Rewrite the likelihood function such that it depends on $\boldsymbol{\tau}$ instead of $\boldsymbol{\pi}$. Then the Hessian can be approximated for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\tau}$. Subsequently, the standard errors for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the square roots of the diagonal elements of the inverse negative Hessian matrix evaluated at the maximum likelihood estimates. The standard errors for $\boldsymbol{\pi}$ can be obtained by simulating from the multivariate normal distribution with mean $\boldsymbol{\tau}$ and covariance matrix $\boldsymbol{\Sigma}_\tau$, the inverse negative Hessian of $\boldsymbol{\tau}$, and computing the covariance matrix of all corresponding values of $\boldsymbol{\pi}$. The standard errors are the square roots of the diagonal elements of the covariance matrix of $\boldsymbol{\pi}$.

Selecting the number of segments is one of the challenges in a finite mixture model. A widely used selection technique is comparing information criteria such as the Bayesian Information Criterium (BIC). A finite mixture model is estimated for different numbers of segments and the values of the BIC are compared. The model with the lowest value for the BIC determines the number of segments. Nevertheless, interpretability and segment sizes are important criteria as well and are included in the segment selection.

4.4 Decomposition of coefficients vector

Many predictors in this study are categorical variables. A disadvantage of categorical variables is that there is always a reference group in a model with an intercept. There are no separate coefficients for these reference groups and in a model with multiple categorical variables the intercept describes all the reference groups of the categorical variables. For instance, if female, Dutch ethnicity, low education, and low income are the reference groups, the intercept is interpreted as the effect of Dutch females who have a low education and low income.

To accommodate every level of the categorical variables with a coefficient, the original coefficients vector $\boldsymbol{b} \in \mathbb{R}^k$ can be transformed into $\boldsymbol{b}^* \in \mathbb{R}^m$, where $\mathbb{R}^k$ is the original space and $\mathbb{R}^m$ additionally includes all reference groups as single variables. In the extended coefficient vector, the coefficients of the levels belonging to the same categorical variable have to sum to zero for all categorical variables. For the interaction terms the coefficients must sum to zero per level. For example, for the interaction term between sex (boy/girl) and ethnicity (Dutch/other) the coefficients for boy-Dutch and boy-other must sum to zero even as boy-Dutch and girl-Dutch.

The extended coefficients vector can be obtained with the following procedure. Let $\boldsymbol{X}_0$ be an $n \times m$ matrix with all levels of the categorical predictors (including intercept and interaction terms). For the example with two categorical variables sex (B/G) and ethnicity (D/O) and their interactions, $\boldsymbol{X}_0$ looks like

$$\boldsymbol{X}_0 = \begin{array}{c} \text{Intercept} \quad \text{B} \quad \text{G} \quad \text{D} \quad \text{O} \quad \text{B.D} \quad \text{G.D} \quad \text{B.O} \quad \text{G.O} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

A finite mixture model with segment specific intercepts contains multiple intercepts. In this case the posterior probabilities are included in $\boldsymbol{X}_0$ instead of an intercept equal to one.

Matrix $\boldsymbol{X}_0$ contains linear dependent columns, because the levels of the categorical variables are linearly dependent. In addition, the interaction terms are linear dependent, because of the linear dependence in the levels of the categorical variables. Let $\boldsymbol{X}_1$ be $\boldsymbol{X}_0$ but with dummy coding with a single level as reference group per categorical variable,

$$\boldsymbol{X}_1 = \begin{array}{c} \text{Intercept} \quad \text{B} \quad \text{G} \quad \text{D} \quad \text{O} \quad \text{B.D} \quad \text{G.D} \quad \text{B.O} \quad \text{G.O} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}.$$

Note that $\boldsymbol{X}_1$ without the zero columns is equal to the predictor matrix used in the regression.

Lastly, define $\boldsymbol{J}$ as a block diagonal matrix with centering matrices with the size equal to the number of levels per variable in $\boldsymbol{X}_0$. For the interaction terms, $\boldsymbol{J}$ contains the kronecker product, denoted by $\otimes$, of the two corresponding centering matrices

$$\boldsymbol{J} = \begin{bmatrix} \boldsymbol{I}_K & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_1 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}_2 \otimes \boldsymbol{C}_1 \end{bmatrix}.$$

Here, $\boldsymbol{I}_K$ is the identity matrix with size equal to the number of intercept coefficients, and $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are both a two dimensional centering matrix corresponding to the categorical variables sex and ethnicity, respectively.

The fitted values are equal to $\boldsymbol{X}_1\boldsymbol{b}$, where $\boldsymbol{b}$ is the original coefficients vector with added zero's for the reference groups. Logically, the fitted values may not change after decomposition. That is, the following equation must hold

$$\boldsymbol{X}_1\boldsymbol{b} = \boldsymbol{A}\boldsymbol{b}^*,$$

where $\boldsymbol{b}^*$ is the transformed vector including all levels of categorical variables. It appears that $\boldsymbol{A} = \boldsymbol{X}_0\boldsymbol{J}$. To solve this equation for $\boldsymbol{b}^*$, premultiply $\boldsymbol{X}_1\boldsymbol{b}$ with the Moore-Penrose inverse of $\boldsymbol{X}_0\boldsymbol{J}$. To compute the standard errors of $\boldsymbol{b}^*$, recall that $\boldsymbol{b}^* = (\boldsymbol{X}_0\boldsymbol{J})^+\boldsymbol{X}_1\boldsymbol{b} = \boldsymbol{G}\boldsymbol{b}$, where $(\boldsymbol{X}_0\boldsymbol{J})^+$ is the Moore-Penrose inverse of $\boldsymbol{X}_0\boldsymbol{J}$. Hence $\boldsymbol{b}^*$ is distributed as $\boldsymbol{b}^* = \boldsymbol{G}\boldsymbol{b} \sim \mathcal{N}(\boldsymbol{G}\boldsymbol{\mu}, \boldsymbol{G}\boldsymbol{\Sigma}\boldsymbol{G}')$, where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are respectively the expectation and covariance matrix of $\boldsymbol{b}$, both added with zeros for the reference groups. The standard errors are the square roots of the diagonal elements of $\boldsymbol{G}\boldsymbol{\Sigma}\boldsymbol{G}'$.

5. Results

This chapter presents the findings of this research. The first analysis investigates which variables influence the performance on the CCT. A linear regression is implemented and the average score obtained by the children is regressed on predictors such as IQ and variables measuring the socio-economic status. Secondly, the number of cards turned over can be predicted with a finite mixture model with a censored Poisson regression. Besides, this model examines the influence of the game settings on the number of cards turned over. First, the process of selecting the number of segments is discussed. Subsequently, the segment populations are presented to try to identify the differences between segments. Finally, the parameter estimates are interpreted and interesting conclusions are drawn.

5.1 Linear regression on average score

The overall success in the CCT is measured with the average score per participant. To determine what characteristics influence the overall success in the CCT a linear regression on the average score per individual is performed. Table 5.1 gives the coefficients and their standard errors between parentheses. The continuous variables, age and IQ are standardised and the decomposition described in Section 4.4 is applied to provide all levels of the categorical variables with a coefficient. The relatively large negative intercept is in accordance with the mean of the average scores (see Figure 3.4). Without regarding any characteristics the average score is around -174. A higher age and IQ will increase this base score with respectively 3.6 and 9.1 points per standard deviation increase. Similarly, boys score on average 16 points higher than girls. Besides, children with a high-educated mother score significantly higher on the CCT. Likewise, children from a Dutch mother perform significantly better than the base average. The coefficients for the other ethnicities do not significantly differ from the base average, however Dutch children seem to score substantially better than Surinamese and Turkish children. The same holds for the monthly household income. Although no category scores significantly different than the base average, children from a family with a low income seem to score worse than children from a family with a high income. Note that instead of ‘higher’ one could also say ‘less worse’, because most children still have a negative score on the CCT.

Table 5.1: Regression on individual level with average score as dependent variable

Variables	Coefficients (st error)
Intercept	-174.04 (2.89)
Age	3.57 (1.50)
Girl	-8.17 (1.50)
Boy	8.17 (1.50)
IQ	9.10 (1.60)
Ethnicity mother	
Dutch	8.47 (3.15)
African	0.39 (6.41)
Asian, non western	5.55 (6.01)
Morrocan	0.97 (6.57)
Dutch Antilles	-2.26 (9.38)
Surinamese	-9.64 (5.36)
Turkish	-8.44 (5.87)
Other western	4.95 (4.78)
Education mother	
No education or primary education	-5.66 (4.72)
Secondary education	-3.39 (2.84)
Higher education	9.06 (3.10)
Household income per month in euro's	
<2000	-5.45 (2.74)
2000-4000	2.23 (2.08)
>4000	3.22 (2.48)

The R-squared of this regression is 0.05

5.2 Finite mixture model

Next to the performance of the CCT, we can also predict the number of cards turned over by a child. Additionally, the influence of the game settings on the number of cards turned over can be analysed. For this analysis a finite mixture model with a censored Poisson regression is implemented. As discussed in Chapter 4 the mean of the Poisson regression is chosen as $\mu_{its} = \max(\alpha_s + \boldsymbol{x}_{it}'\boldsymbol{\beta}, \epsilon)$, where $\epsilon = 10^{-16}$. Furthermore, in this analysis the EM-algorithm has converged if the relative difference between the likelihood value and updated likelihood value is smaller than 10^{-14}. This strong criteria is necessary to ensure the gradients of α_s and $\boldsymbol{\beta}$ are equal to zero up to three decimals.

The number of segments is unknown on forehand, therefore the model is computed for several numbers of segments and the final model is selected based on the BIC and interpretability. Table 5.2 presents statistics of five models with different numbers of segments. The segment specific intercepts $\boldsymbol{\alpha}$ are the ones obtained after the decomposition as described in Section 4.4. If the models are compared based on only the BIC the model with nine segments would have been selected. However, segments one to four have similar segment specific intercepts in this model. In other words, the expected average number of cards turned over is almost the same in these segments. Therefore, it is questionable whether these four segments are indeed four distinct segments or that they can be summarized in fewer segments. The model with eight segments seems to comprise the first four segments from the nine-segments model into three segments, namely the first three. The differences of the segment specific intercepts in the eight-segments model is larger, although, one could still argue that segments one and two and segments five and six are close. Furthermore, the last segment is relatively small and describes only one percent of the population. This could also indicate that a model with fewer segments is better. However, the last segment is discriminative from the other segments based on the segment characteristics presented in Table 5.3. Thus, further analysis is based on the model with eight segments, because the interpretablility of the eight-segments model is better than the nine-segments model, and the difference in BIC is rather small between these two models. The seven-segments model is not considered, because the BIC of this model is much higher than that of the eight- and nine-segments model.

Table 5.2: Statistics of the models with respectively 6, 7, 8, 9, 10, and 11 segments

	Segment											BIC
	1	2	3	4	5	6	7	8	9	10	11	
6												167747
Segmentspecific intercept α	22.89	19.01	18.96	8.82	9.75	5.90						
Segment probability π	0.266	0.221	0.220	0.124	0.113	0.057						
7												167478
Segmentspecific intercept α	22.61	21.03	18.43	14.40	9.13	7.98	5.42					
Segment probability π	0.228	0.212	0.186	0.145	0.092	0.091	0.046					
8												167309
Segmentspecific intercept α	22.57	21.89	19.59	15.16	8.97	9.78	6.64	4.55				
Segment probability π	0.211	0.204	0.183	0.141	0.099	0.091	0.057	0.013				
9												167300
Segment specific intercept α	21.45	20.75	20.20	19.90	15.34	11.34	8.70	6.60	4.51			
Segment probability π	0.167	0.162	0.158	0.155	0.120	0.088	0.083	0.054	0.013			
10												167321
Segment specific intercept α	22.63	21.89	21.30	20.99	11.97	8.70	10.50	6.60	5.68	4.51		
Segment probability π	0.167	0.162	0.158	0.155	0.088	0.083	0.078	0.054	0.042	0.013		
11												167343
Segment specific intercept α	22.90	22.17	21.56	21.24	12.11	8.70	8.59	6.60	6.71	4.51	1.13	
Segment probability π	0.167	0.162	0.157	0.155	0.088	0.083	0.063	0.054	0.049	0.013	0.008	

Table 5.3 presents characteristics of the segment populations of the eight segments in the model and the average population over all segments. For instance, the average score obtained in the sixth segment is lowest, whereas it is highest in the last segment. As said previously segment eight is distinctive from the other segments. Besides the high score achieved in the last segment, the number of cards turned over and the number of censored trials are lowest in this segment. Moreover, segment eight mainly represents boys with a mother with an Asian (non Western), Surinamese, or Other Western ethnicity. Girls and children with a mother with a Dutch, African, or Dutch Antilles ethnicity are underrepresented in segment eight. In addition, the IQ of the child is relatively low in segment eight. Regarding the maternal education, the lowest category is well represented, whereas the middle category is not.

The children assigned to segment seven also perform relatively good on the CCT. These children are mainly boys with a mother from African or Dutch Antilles descent. In particular, the Moroccan ethnicity is uncommon in segment seven. Moreover, children assigned to this segment are primarily part of a household with a monthly income above 4000 euro's. Especially, households with a monthly income between 2000 and 4000 are rare in segment seven.

Segment six, the segment with the lowest average score, describes mainly girls with a Surinamese or Turkish background and relatively less children with a Dutch background. Additionally, most mothers in this segment have a secondary education and mothers with a higher education are mostly assigned to other segments. Likewise, most households in this segment belong to the lowest category of monthly household income and in particular do not belong to the highest category of household income.

Another segment where children with a low average score are assigned to is segment four. Similar as in segment six, children with a Turkish mother are described by this segment. The difference with segment six is that in segment four the IQ is closer to the average IQ over all segments. Likewise, segment four does not discriminate on sex.

Noteworthy, the segments describing children who preform worse on the CCT, segments four and six, also represent the children who turn over many cards and face many loss cards. On the other hand, children represented by segment seven and eight perform best on the CCT and turn over less cards and consequently, face less loss cards. Being risk averse seems to be the best strategy in this game. Another interesting fact is that the two segments with the lowest IQ, segment six and eight, are the segments that perform respectively worst and best on the CCT. Furthermore, the segments where children perform well on the CCT, segment seven and eight, have a high population of boys, whereas segment six, the least performing segment, has the highest population of girls. This is in accordance with the results from the first analysis, see Table 5.1, that showed that boys on average score higher on the CCT than girls do.

Segment one has a large Dutch and Moroccan population and contains a bit more girls than the average population. On the other characteristics the segment population in segment one is similar to the average population over all segments. The same holds for segment two, which only discriminates on the characteristic IQ. This segments represents children with a high IQ, although the difference with the average IQ is small.

The population of segment five contains a large sample of children with an African mother compared to the average population. Contrary, children with a Turkish mother are uncommon in this segment. In addition, the average score obtained by the children represented by segment five is above average. Segment three can be seen as the segment representing an ordinary child. This segment does not discriminate on any of the characteristics.

Table 5.3: Characteristics of segment populations of the model with eight segments

	Segment								
	1	2	3	4	5	6	7	8	Total
Age	9.76	9.78	9.78	9.77	9.76	9.76	9.74	9.77	9.77
Sex (% girl)	52.1	51.6	49.1	50.0	49.3	53.7	46.5	47.3	50.6
IQ	102.7	103.1	102.2	101.6	102.6	99.2	101.2	98.2	102.1
Ethnicity Mother (%)									
Dutch	63.4	60.8	58.0	59.6	60.0	54.9	57.3	43.0	59.5
African	4.3	4.8	5.0	3.6	6.5	5.7	6.4	2.8	4.9
Asian, non Western	4.3	6.2	5.5	6.3	5.4	4.7	6.2	12.0	5.6
Moroccan	5.5	4.6	5.3	3.9	4.3	3.7	2.2	4.5	4.6
Dutch Antilles	1.6	1.7	2.4	2.3	1.9	3.5	4.0	0.4	2.2
Surinamese	7.1	6.9	6.6	7.3	5.1	9.8	6.8	14.2	7.1
Turkish	4.5	6.3	6.5	7.4	4.0	8.3	4.5	5.4	6.0
Other Western	9.2	8.7	10.7	9.6	12.9	9.5	12.4	17.6	10.1
Education mother (%)									
No education or primary education	5.1	5.1	4.3	4.1	4.9	5.9	4.8	11.3	4.9
Secondary education	36.1	37.6	38.5	39.1	37.5	46.9	37.7	30.2	38.4
Higher education	58.8	57.3	57.2	56.8	57.6	47.2	57.5	58.4	56.6
Household income per month in euro's (%)									
<2000	21.2	21.6	22.8	21.8	22.7	29.9	27.7	29.7	23.0
2000 - 4000	43.8	44.8	43.7	44.8	41.9	44.3	33.6	40.6	43.4
>4000	35.0	33.6	33.5	33.4	35.5	25.8	38.7	29.7	33.6
Average score	-161.3	-184.9	-131.0	-214.9	-102.0	-261.2	-75.9	-55.4	-165.1
# cards turned over	9.2	10.6	7.8	11.8	6.4	12.3	5.0	3.4	9.3
# censored trials	10.6	12.2	8.9	13.9	7.1	15.6	5.5	4.0	10.8

The parameter estimates of the finite mixture model are reported in Table 5.4 and Table 5.5. The variables previous score and second previous score denote the scores obtained in respectively the trial before the current trail and two trials before the current trial. The continuous variables previous score, second previous score, age, and IQ are standardised for this analysis.

Because of the coefficients decomposition as described in Section 4.4, the intercepts in Table 5.4 represent the expected average number of cards to turn over per segment without correcting for any of the observed variables. Segment one is the most risk seeking segment and segment eight is the most risk averse segment. However, due to the differences in segment population the actual average number of cards turned over is different, see Table 5.3.

Furthermore, the segment probabilities $\boldsymbol{\pi}$ show that the prior probability to be assigned to segment one is highest and the prior probability to be assigned to segment eight is lowest. Hence segment one has the largest population and segment eight has the smallest population.

Table 5.4: Intercepts $\boldsymbol{\alpha}$ after decomposition and segment probabilities $\boldsymbol{\pi}$ with standard errors

	Segments							
	1	2	3	4	5	6	7	8
$\boldsymbol{\alpha}$	22.57	21.89	19.59	15.16	8.97	9.78	6.64	4.55
st. error	0.240	0.233	0.208	0.161	0.147	0.104	0.107	0.124
$\boldsymbol{\pi}$	0.211	0.204	0.183	0.141	0.099	0.091	0.057	0.013
st. error	0.011	0.011	0.009	0.010	0.009	0.006	0.005	0.002

Table 5.5 reports the β-coefficients after decomposition. Note that the model is a linear model, so the estimates should be interpreted linearly. For example, according to the model a girl will turn over 0.534 ($= 0.267 + |-0.267|$) cards more than a boy. Moreover, older children seem to be more risk averse than younger children. If the age increases with one standard deviation (0.26 years) the expected number of cards turned over decreases with almost 0.3 cards. So, someone who is one year older will turn over 1.1 cards less. Likewise, a higher IQ leads to a lower expected number of cards turned over. However, the latter effect is rather small. The IQ should increase with 14.7 points (one standard deviation) before the expected number of cards turned over decreases with half a card. Recall that Table 5.3 shows that a risk averse strategy is more profitable in this game.

Furthermore, children with a Dutch or Moroccan mother are likely to be risk averse, whereas children with an African, Dutch Antilles, or Other Western background are risk seeking. Although, the coefficient of the Turkish ethnicity is not significant, Turkish children are likely to be more risk averse than children with an African, Dutch Antilles, and Other Western descent. Moreover, a low maternal education positively effects the number of cards turned over, whereas a high maternal education negatively effects the number of cards turned over. Similarly, a monthly household income below 2000 euro's has a positive effect on the number of cards turned over, whereas a monthly household income between 2000 and 4000 has a negative effect.

Due to the different game settings we can investigate the influence of the loss probability and the sensitivity to reward and punishment. According to the model, in a trial with one loss card on average two and a half cards more are turned over than there are in a trial with three loss cards. The game setting loss amount also shows the expected sign. In a trial with a high loss amount the expected number of cards turned over is lower. Against all odds, the expected number of cards turned over is lower in a trial with a high gain amount than it is in a trail with a low gain amount. Recall that Table 3.1, which presents the average number of cards turned over per game setting shows the same pattern.

The scores obtained in previous rounds have a significant effect on the risk behaviour in the current round. If someone scores high in the previous round he or she is more likely to turn over more cards in the current round. As expected, the effect of the score in the previous round is stronger than the effect of the score in the second previous round. The effect of the previously obtained scores on the current risk behaviour could indicate the presence of a learning effect.

In addition, this model includes interaction terms between the game settings and sex. According to the model, the sex boy together with a loss amount of 250 accounts for an additional 0.161 ($= 0.302 - 0.267 + 0.126$) cards to be turned over. In a trial with loss amount 750, a boy is expected to turn over 0.695 ($= |-0.302 - 0.267 - 0.126|$) cards less than the base average (i.e. the segment specific intercepts). Hence, the effect the loss amount has on the number of cards turned over by a boy is 0.856 ($= 0.161+|-0.695|$). This effect is smaller for girls, namely 0.624 ($= (0.302+0.267-0.126)+(-0.302+0.267+0.216)$). To conclude, boys seem to be more sensitive to punishment in the CCT than girls are.

The same pattern can be seen when looking at the number of loss cards in a trial. The effect of the number of loss cards is larger for boys than for girls, namely 2.81 versus 2.298. This effect is for both boys and girls larger than the effect of the loss amount. Hence, both boys and girls seem to be more influenced by the probability of loss than by the height of a possible punishment.

For the game setting gain amount the pattern is different. Girls seem to be more sensitive to reward than boys, 0.892 versus 0.534. However, according to the model, girls turn over an additional 0.713 cards when the gain amount is ten and turn over 0.179 cards less than the base average when the gain amount is thirty. For boys both values are negative, in other words regardless the gain amount, boys turn over less cards than the base average.

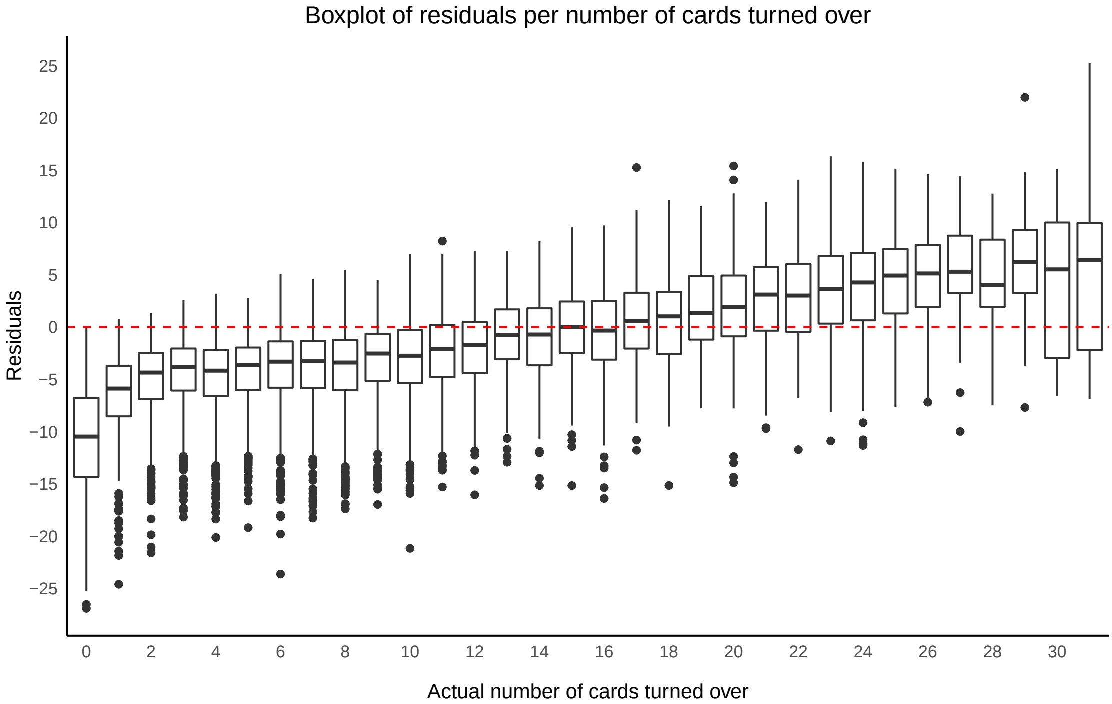

Figure 5.1: Boxplot with residuals per number of cards turned over of the model with eight segments

Figure 5.1 displays the distribution of the residuals per number of cards turned over. This plot is based on only the noncensored cases, because the model predicts how many cards someone wants to turn over, not how many cards someone actually turns over. The model would have fitted perfectly if the boxes were small and would have lied around the red horizontal line. As you can see, the model, on average, overestimates the number of cards turned over if the actual number of cards turned over is below twelve. If the actual number of cards turned over is above twenty, the model mainly underestimates the number of cards turned over. The size of the boxes describes the variation in the residuals. If a box is small, 75% of the residuals have similar values. For example, the box for 13 cards is smaller than the box for 14 cards. Because both boxes lie around the red horizontal line, we can conclude that the model, in 75% of the cases, better predicts if the actual number is 13 than it does if the actual number is 14. Overall, the boxes for a low number of cards are smaller than the boxes for a high number of cards. This means that the model more accurately predicts when the actual number of cards turned over is low. However, for the low actual number of cards turned over the residuals have more outliers, meaning that the predictions are not as accurately as can be expected from the box sizes. Besides, note that the increasing size of the boxes is largely due to the heteroskedasticity of the Poisson model. Contrary, the graph shows that the model also has difficulties predicting the correct value if the actual value is zero.

Table 5.5: β-coefficients with standard errors between brackets

	β - coefficients (st error)
Age	-0.287 (0.048)
Girl	0.267 (0.054)
Boy	-0.267 (0.054)
IQ	-0.547 (0.045)
Ethnicity Mother	
Dutch	-0.863 (0.120)
African	0.469 (0.235)
Asian, non Western	-0.230 (0.226)
Moroccan	-0.585 (0.241)
Dutch Antilles	0.944 (0.324)
Surinamese	0.175 (0.186)
Turkish	-0.606 (0.366)
Other Western	0.697 (0.180)
Education mother	
No education or primary education	0.554 (0.174)
Secondary education	0.008 (0.120)
Higher education	-0.561 (0.105)
Householdincome per month in euro's	
<2000	0.606 (0.116)
2000 - 4000	-0.587 (0.077)
>4000	-0.020 (0.082)
Gain amount (10)	0.284 (0.021)
Gain amount (30)	-0.284 (0.021)
Loss amount (250)	0.302 (0.021)
Loss amount (750)	-0.302 (0.021)
# loss cards (1)	1.277 (0.022)
# loss cards (3)	-1.277 (0.022)
Previous score	0.894 (0.022)
Second previous score	0.583 (0.022)
Interaction terms	
Gain amount (10) : girl	0.162 (0.021)
Gain amount (30) : girl	-0.162 (0.021)
Gain amount (10) : boy	-0.162 (0.021)
Gain amount (30) : boy	0.162 (0.021)
Loss amount (250) : girl	-0.126 (0.021)
Loss amount (750) : girl	0.126 (0.021)
Loss amount (250) : boy	0.126 (0.021)
Loss amount (750) : boy	-0.126 (0.021)
# loss cards (1) : girl	-0.128 (0.021)
# loss cards (3) : girl	0.128 (0.021)
# loss cards (1) : boy	0.128 (0.021)
# loss cards (3) : boy	-0.128 (0.021)

6. Conclusion and Limitations

This research focusses on risk behaviour of children measured with the Columbia Card Task (CCT). Besides, it investigates whether risk behaviour can be predicted from the IQ and socio-economic status of a child. The design of the CCT additionally allows us to analyse different factors that may influence the decision process.

An analysis on the performance of children on the CCT shows that age and sex are important predictors. The performance on the CCT increases with an increase in age, and boys score on average higher than girls. Furthermore, Dutch children score significantly higher on the CCT than the average and Surinamese and Turkish children score substantially lower than Dutch children.

Besides the performance on the CCT, one could also measure the risk behaviour with the CCT. The more cards a child turns over, the more risk seeking he or she is. A finite mixture model with eight segments shows that boys are more risk averse than girls. Besides, older children and children with a higher IQ turn over less cards. From the segment populations we can conclude that a risk averse strategy is more profitable than a risk seeking strategy. Furthermore, children with a Dutch or Moroccan mother are more risk averse than the average. Contrary, children with an African, Dutch Antilles, or Other Western background turn over more cards than the average. In addition, there seems to be a difference between children with a Turkish descent and children with an African, Dutch Antilles, or Other Western descent, where the first group is more risk averse than the latter. Regarding the maternal education and monthly household income, the lower these two variables the more risk seeking a child is on average.

An advantage of the CCT is that it allows researchers to measure several underlying factors of risk behaviour. Due to the different game settings a researcher could distinguish between the effect of the loss probability and the sensitivity to reward and punishment on risk behaviour. Both the loss amount and loss probability show the expected sign, however, according to the model, a higher gain amount results in a lower expected number of cards turned over. Furthermore, the effect of the number of loss cards is greater than the effect of the loss amount. This result is in accordance with previous findings (Kluwe-Schiavon et al., 2015). Besides, boys are, according to the model, more influenced by the loss probability and are more sensitive to punishment than girls are. In addition, previous results on the CCT have a significant effect on the current game. This could indicate the presence of a learning effect.

It is hard to label the eight segments based on their characteristics, however the segment populations can be used to compare segments that perform well on the CCT with segments that perform worse. Segment seven and eight perform best on the CCT and describe mainly boys. Note that the linear regression on the average score also indicates that boys score better than girls. Segment eight, in addition, has a low population of Dutch children, which does not corresponds to the results of the linear regression model. On the other hand, segment six, the segment that performs worst on the CCT, also contains few Dutch children. Besides, both segment six and segment eight have a low average IQ score. Moreover, segment six and seven show that children from a household with a low income perform worse on the CCT than children from a household with a high income. This result is in accordance with the findings from the linear regression on the average score.

A major limitation in this study is the censoring of the data. More than two-third of the trials are censored, making it more difficult to come to accurate estimates. This could have been solved by manipulating the game such that the loss card is at the last possible card, as Figner et al. (2009) did. However, if the game is manipulated in this way participants have to complete more trials to ensure that they do not discover the manipulation of the game, making the research more expensive. Besides, relatively many children continued turning over cards in all trials until they faced a loss card. Hence, it is questionable whether manipulating the game like this will lead to more accurate results.

The CCT may be considered as an advanced risk task, since participants have to consider three different game settings during the decision making process. According to van Duijvenvoorde et al. (2015) children in the age category 8-10 years understand the CCT. However, their argumentation is based on the significant result that the effects of all game settings were in the expected direction. However, we found a significant effect of the gain amount in the opposite direction. In this research we excluded 104 children that turned over the last card, which is definitely a loss card, because they clearly did not understand the task. In addition the data set contains 309 children who faced a loss card in all 16 trials. These children are not excluded from the analysis, because these children could just be extremely risk seeking. On the other hand, one could argue that these children did not understand the function of the stop button, with which they could stop the trial and collect the earned points. However, after the first three trials the children were reminded of the stop button, therefore we found the risk seeking argument more plausible in this case.

The finite mixture model used to analyse the risk behaviour has some limitations as well. The parameters are estimated with the EM-algorithm, which is known to be a slowly converging algorithm. However, even after considering the slow convergence of the EM-algorithm this model is extremely time consuming. We first considered a model with multiplicative effects (i.e. defining the mean as $\boldsymbol{\mu} = \exp(\boldsymbol{x}'\boldsymbol{\beta})$), which is more common in a Poisson regression model, but a linear model appeared to converge faster, presumably because it suffers from less numerical instability.

It is important that the mean is specified such that it cannot have negative values, because a Poisson regression model is not defined if the mean is negative. Besides, in this case a negative mean does not make any sense, because it is impossible to turn over less than zero cards. In this research the nonnegativity is guaranteed by taking the maximum of $\alpha_s + \boldsymbol{x}'_{it}\boldsymbol{\beta}$ and a value sufficiently close to zero, but one could also choose to set a penalty on negative values of $\alpha_s + \boldsymbol{x}'_{it}\boldsymbol{\beta}$ to ensure nonnegativity. For further research different specifications for the mean could be considered.

Another improvement that could be considered is a different link function and/or distribution for a smoother likelihood function and faster convergence. For instance, a negative binomial distribution could be applied. The advantage of this distribution is that it has a finite upper bound, whereas the Poisson distribution has no upper bound (i.e. the upper bound is infinite). In the CCT participants can turn over a maximum of 31 cards, so a distribution with a upper bound at 31 might better fit the data. Additionally, the negative binomial distribution does not assume equidispersion. However, this distribution is not examined for this research.

An assumption of the Poisson distribution is equidispersion, equal mean and variance. However, in practice this assumption is often violated. Tests to check the validity of this assumption are designed for noncensored data, but not for censored data. When the data is censored the true mean and true variance can not be computed easily. Therefore, the best way to check for equidispersion when the data is censored is to compare a censored Poisson regression model, which assumes equidispersion, with a model that does not assume equidispersion. Examples of models that do not assume equidispersion are a censored generalized Poisson regression model (Famoye & Wang, 2004) and a model with a negative binomial distribution. However this is out of the scope of this research.

Finally, the analysis could be improved by limiting the randomness in the analysis, leading to more robust results. In particular, the EM-algorithm, which is used to find the maximum likelihood estimates of the parameters in the finite mixture model, is sensitive to local maxima. It is advisable to use random starts to reduce the problem of local maxima. However, due to the time constraint using random start values was not possible in this analysis. Moreover, single imputation is performed to fill missing data values. Again, multiple imputations will reduce the randomness and give more robust estimates of the missing values. Selecting the number of segments is now largely based on the Bayesian Information Criterion (BIC), while other criteria could have led to a different number of segments. A better approach might be to use cross validation to determine the number of segments (Grimm, Mazza, & Davoudzadeh, 2017). However, again due to the time constraint this was not possible in this analysis.

References

Aitkin, M., & Rubin, D. B. (1985). Estimation and hypothesis testing in finite mixture models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 67–75.

Brand, M., Fujiwara, E., Borsutzky, S., Kalbe, E., Kessler, J., & Markowitsch, H. J. (2005). Decision-making deficits of korsakoff patients in a new gambling task with explicit rules: associations with executive functions. *Neuropsychology*, *19*(3), 267.

Buelow, M. T. (2015). Predicting performance on the columbia card task effects of personality characteristics, mood, and executive functions. *Assessment*, *22*(2), 178–187.

Caird, S. (1991). The enterprising tendency of occupational groups. *International Small Business Journal*, *9*(4), 75–81.

Chye Koh, H. (1996). Testing hypotheses of entrepreneurial characteristics: A study of hong kong mba students. *Journal of managerial Psychology*, *11*(3), 12–25.

Collins, L. M., Sussman, S., Rauch, J. M., Dent, C. W., Johnson, C. A., Hansen, W. B., & Flay, B. R. (1987). Psychosocial predictors of young adolescent cigarette smoking: A sixteen-month, three-wave longitudinal study1. *Journal of Applied Social Psychology*, *17*(6), 554–573.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, 1–38.

Famoye, F., & Wang, W. (2004). Censored generalized poisson regression model. *Computational statistics & data analysis*, *46*(3), 547–560.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the columbia card task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 709.

Figner, B., & Weber, E. U. (2011). Who takes risks when and why? determinants of risk taking. *Current Directions in Psychological Science*, *20*(4), 211–216.

Gilbert, P., & Varadhan, R. (2012). numderiv: accurate numerical derivatives (Vol. 1) [Computer software manual].

Grimm, K. J., Mazza, G. L., & Davoudzadeh, P. (2017). Model selection in finite mixture models: A k-fold cross-validation approach. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(2), 246–256.

Holper, L., & Murphy, R. O. (2014). Hemodynamic and affective correlates assessed during performance on the columbia card task (cct). *Brain imaging and behavior*, *8*(4), 517–530.

Huang, Y. H., Wood, S., Berger, D. E., & Hanoch, Y. (2015). Age differences in experiential and deliberative processes in unambiguous and ambiguous decision making. *Psychology and aging*, *30*(3), 675.

Kluwe-Schiavon, B., Sanvicente-Vieira, B., Viola, T. W., Veiga, E., Bortolotto, V., &

Grassi-Oliveira, R. (2015). Assessing affective and deliberative decision-making: adaptation of the columbia card task to brazilian portuguese. *The Spanish journal of psychology*, *18*, E89.

Konnikova, M. (2013). *The limits of self-control: Self-control, illusory control, and risky financial decision making* (Unpublished doctoral dissertation). Columbia University.

Lejuez, C. W., Aklin, W. M., Jones, H. A., Richards, J. B., Strong, D. R., Kahler, C. W., & Read, J. P. (2003). The balloon analogue risk task (bart) differentiates smokers and nonsmokers. *Experimental and clinical psychopharmacology*, *11*(1), 26.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., ... Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the balloon analogue risk task (bart). *Journal of Experimental Psychology: Applied*, *8*(2), 75.

Lowry, R., Kann, L., Collins, J. L., & Kolbe, L. J. (1996). The effect of socioeconomic status on chronic disease risk behaviors among us adolescents. *Jama*, *276*(10), 792–797.

Miner, J. B. (1990). Entrepreneurs, high growth entrepreneurs, and managers: Contrasting and overlapping motivational patterns. *Journal of business venturing*, *5*(4), 221–234.

Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, *6*(4), 525–533.

Nelder, J., & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370-384.

Penolazzi, B., Gremigni, P., & Russo, P. M. (2012). Impulsivity and reward sensitivity differentially influence affective and deliberative risky decision making. *Personality and Individual Differences*, *53*(5), 655–659.

Pripfl, J., Neumann, R., Köhler, U., & Lamm, C. (2013). Effects of transcranial direct current stimulation on risky decision making are mediated by hotand colddecisions, personality, and hemisphere. *European Journal of Neuroscience*, *38*(12), 3778–3785.

Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J., & Robbins, T. W. (1999). Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. *The Journal of Neuroscience*, *19*(20), 9029–9038.

Tiemeier, H., Velders, F. P., Szekely, E., Roza, S. J., Dieleman, G., Jaddoe, V. W., ... others (2012). The generation r study: a review of design, findings to date, and a study of the 5-httlpr by environmental interaction from fetal life onward. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(11), 1119–1135.

van Duijvenvoorde, A. C., Huizenga, H. M., Somerville, L. H., Delgado, M. R., Powers, A., Weeda, W. D., ... Figner, B. (2015). Neural correlates of expected risks and returns in risky choice across development. *Journal of neuroscience*, *35*(4), 1549–1560.

Watrous, R. L. (1988). Learning algorithms for connectionist networks: Applied gradient methods of nonlinear optimization. *Technical Reports (CIS)*, 597.

Welte, J. W., Barnes, G. M., Wieczorek, W. F., Tidwell, M.-C., & Parker, J. (2002). Gambling participation in the usresults from a national survey. *Journal of gambling studies*, *18*(4), 313–337.