Mixed Models for Primary Education in Rural Areas of Madagascar

Zdeněk Hlávka

Final Project for Postgraduate Programme in Biostatistics
at L.U.C. Diepenbeek
Supervised by Prof. Dr. J. K. Lindsey
July — September 1996

Abstract

In this report, four responses connected with system of primary education on Madagascar are being analyzed. Enrolment to school and admission to school are binary and are analyzed by logistic regression. Delay in going to school and delay in going to school in 1993 are counts and are modelled by Poisson type log-linear model. The responses are correlated within families and villages and therefore I will use models with random effects. These models will be compared with the fixed effects models given in report [1].

Contents

List of Tables

Chapter 1

Introduction

The aim of this report is to fit random effects models to data about primary education in Madagascar. It is likely that random effects models would provide better fit than fixed effects models, because of the structure of the dataset.

1.1 Dataset

The dataset was provided by Prof. Dr. J. K. Lindsey, his study [1] was supported and coordinated by the Ministry of National Education of Madagascar and UNESCO.

The dataset contains information about 4012 children. Investigators collected more than one thousand variables describing children, family, village, school and province characteristics. For my analyses, I will use only about forty explanatory variables (this number is not fixed because some variables in some analyses become nonapplicable).

I will provide only very short outline of the investigation. Detailed and more colorful description can be found in reports [2] or [1].

1.1.1 Investigation

It was decided to investigate only rural area of the island, where about 80% of population Madagascar live. The island is divided to 111 CISCOs. Of these 104 lie in rural area. In the 46 more populated CISCOs in rural area, four villages (two with and two without school) were randomly chosen. In 58 less populated CISCOs, only two villages (one with and one without school) were chosen. This gives a sample of 300 villages, stratified for CISCOs and presence of a school.

In each village investigators randomly chose five families. Thus a sample of 1500 families with 4012 children was obtained.

1.2 Some theory

In this section, I will try to describe models used in this report and how SABRE is dealing with them. I used logistic regression with random effects for modelling binary responses and random effects log-linear model with Poisson distribution for modelling count data.

1.2.1 Binary response

Binary response data are usually (most often) modelled by fixed effects logistic models. This "standard" model is given by this formula:

$$\log \frac{\pi_i}{1-\pi_i} = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_n x_{ni}, \tag{1.1}$$

where π_i is probability of success in i^{th} observation, x_{ij} are covariates and β's are parameters. Sometimes it can be more appropriate to model these data with random effects logistic models, this is true especially when more measurements are being taken on one individual.

In this report, I will use random effects model only with one level of random effects. This model can be written as

$$\log \frac{\pi_{it}}{1-\pi_{it}} = \beta_0+\beta_1 x_{1it}+\ldots+\beta_n x_{nit}+\gamma_i \quad \text{for } t = 1,\ldots,n_i,\ i = 1,\ldots,n, \tag{1.2}$$

where n is number of cases (e.g. villages) and n_i is number of individuals (children) belonging to i^{th} case. γ_i is random effect of i^{th} case. This model can be rewritten in terms of π_{it} as

$$\pi_{it} = \frac{\exp(\beta_0 + \beta_1 x_{1it} + \ldots + \beta_n x_{nit} + \gamma_i)}{1 + \exp(\beta_0 + \beta_1 x_{1it} + \ldots + \beta_n x_{nit} + \gamma_i)}. \tag{1.3}$$

From this, it immediately follows that the contribution of i^{th} case to the likelihood function is

$$L_i(\beta) = \prod_{t=1}^{n_i} \frac{\exp(\beta_0 + \beta_1 x_{1it} + \ldots + \beta_n x_{nit} + \gamma_i)^{y_{it}}}{1 + \exp(\beta_0 + \beta_1 x_{1it} + \ldots + \beta_n x_{nit} + \gamma_i)}. \tag{1.4}$$

In order to get rid of nuisance parameters γ_i, we can use the assumption that these are random effects and replace this likelihood with following integrated likelihood

$$L_i(\beta) = \int \prod_{t=1}^{n_i} \frac{\exp(\beta_0 + \beta_1 x_{1it} + \ldots + \beta_n x_{nit} + \gamma_i)^{y_{it}}}{1 + \exp(\beta_0 + \beta_1 x_{1it} + \ldots + \beta_n x_{nit} + \gamma_i)} f(\gamma_i) d\gamma_i \tag{1.5}$$

in which the nuisance parameters are integrated out with the help of their density $f(x)$.

This expression is in SABRE further replaced by

$$L_i(\beta,\omega)=\sum_{j=1}^{Q}p_j\prod_{t=1}^{n_i}\frac{\exp(\beta_0+\beta_1x_{1it}+\ldots+\beta_nx_{nit}+\omega z_j)^{y_{it}}}{1+\exp(\beta_0+\beta_1x_{1it}+\ldots+\beta_nx_{nit}+\omega z_j)}, \tag{1.6}$$

where p_j and z_j are standard normal probability and location respectively (SABRE can fit only one level of normal random effects) and ω is standard deviation of the random effect. The sum in 1.6 is rough approximation of integral in formula 1.5. Numbers p_j and ωz_j in 1.6 are actually "discrete version" of the density $f(\gamma_i)$ in 1.5. Number p_j is the probability concentrated in point ωz_j.

SABRE furthermore allows for mover–stayer model which gives following likelihood

$$L_i^{m-s}(p_0,p_1,\beta,\omega)=p_0\prod_{t=1}^{n_i}(1-y_{it})+p_1\prod_{t=1}^{n_i}y_{it}+(1-p_0-p_1)L_i(\beta), \tag{1.7}$$

where p_0 is probability of being stayer in state zero, p_1 is probability of being stayer in state one and $L_i(\beta)$ is the contribution of i^{th} case to likelihood as given above. SABRE estimates instead of probabilities p_0 and p_1 end-point parameters ψ_0 and ψ_1. The probabilities can be calculated from these end-point parameters in following way:

$$p_i=\frac{\psi_i}{1+\psi_1+\psi_0}\quad\text{for } i=0,1. \tag{1.8}$$

1.2.2 Counts

The random effects log-linear model with Poisson distribution used for count data is very similar to the model for binary responses. It is given by this formula:

$$\log\mu_{it}=\beta_0+\beta_1x_{1it}+\ldots+\beta_nx_{nit}+\gamma_i. \tag{1.9}$$

In the same way as with binary response we can arrive to contribution of i^{th} case to the likelihood:

$$L_i(\beta)=\int\prod_{t=1}^{n_i}\exp(\beta_0+\ldots+\gamma_i)^{y_{it}}\exp(-\exp(\beta_0+\ldots+\gamma_i))f(\gamma_i)d\gamma_i. \tag{1.10}$$

This integral is numerically evaluated in the same way as the integral for binary response. SABRE also here allows to include stayers in state zero. The likelihood for mover–stayer model can be obtained in the same way as likelihood for mover–stayer model for binary response.

1.3 SABRE

For fitting random effects models for binary and count data, I used program SABRE (Software for Analysis of Binary Recurrent Events). SABRE is downloadable on this WWW site: `http://www.cas.lancs.ac.uk/software/`. Here, I want to say few words about its advantages and limitations.

1.3.1 Advantages

The biggest advantage of SABRE is that it is able to fit models with random effects. Second advantage is that it is shareware and the only price you have to pay for using it is certain discomofort, because SABRE is not very user friendly.

1.3.2 Limitations

The only thing SABRE can do is to fit logistic model or log-linear model with Poisson distribution with ot without random effects. It can fit only one level of random effects and it can handle only normally distributed random effect.

Other limitation is that maximum number of variables is one hundred. This number can be increased if you change the source of SABRE (in cooperation with its authors and most likely in very complicated way).

Probably the worst problem (at least for my analyses) is that SABRE does not allow to weight out observations with missing values. Therefore I had to exclude observations with missing values in SAS, then I had to change the output in BRIEF, then I had to use WS-FTP to move the resulting file to my home directory on alpha (this is our server, where I ran SABRE) and then finally I could use TELNET to run SABRE and to analyse the data. This procedure is so long that it was impossible to follow exactly the same model building strategy as ChangLin Mei and A. F. Poblete in [2] and Lindsey in [1] and to use different data for checking level of significance of every variable.

1.3.3 Short overview

Basic commands in SABRE are:

- DATA: this command specifies the variables (if you use this command you start completely new analysis, all old variables are lost)
- READ: reads the data
- YVARIATE: specifies the response variable
- FACTOR: declares variable as categorical (creates a collection of dummy variables which behaves as one new variable)
- FIT: fits model with random effects

- DISPLAY E: displays estimates
- STOP: leaves SABRE

It is possible to survive only with these commands, however other useful commands are:

- CASE: specifies the case variable (as default is taken the first variable given in DATA command)
- INPUT: takes input from specified file
- FIT + and FIT -: addition or removal of explanatory variables (you cannot do both at once)
- LFIT: fits fixed effects logistic model (this model can be obtained also with FIT command if you use mass-point of one and end-points fixed at zero — commands MASS and ENDPOINTS)
- POISSON: option YES allows to analyse count data
- OUTPUT: specifes the log file
- MONITOR: reduces the information about model fitting
- ARITHMETIC: you can choose FAST or ACCURATE, it is very useful if there are some problems with overflow or underflow

This is only short overview of most useful commands. There are still few commands I did not mention. Very nice on-line manual (with complete list of commands) you can find on above mentioned WWW site.

1.4 Methodology

ChangLin Mei and A. F. Poblete in [2] and J. K. Lindsey in [1] used following model building strategy for fixed effects models:

1. Include important children characteristics
2. Include important family characteristics
3. Include important village characteristics
4. Include important school characteristics
5. Remove useless variables
6. Check province differences

They used as many observations as possible — they weighted out observations with missing values in variables currently in the model. It means that for checking significance level of variable BIOLOGIC they were using slightly different data than when they were checking significance of variable FIELD. Their model building was based on AIC — variable entered model if the change in deviance was greater than 2.

I tried to follow this strategy as closely as possible, but I had to make some changes in order to make my analyses feasible with SABRE.

The main difference is that I used only six different datasets for each response variable, because SABRE does not have any facility which allows to weight out observations with missing values. It means that in the first step (children characteristics), I used observations without missing values in children characteristics, in the second step (family characteristics), I used observations without missing values in all family characteristics and children characteristics currently in the model and so on. My model building was also based on AIC.

It is not very clear how to compare different models which fit different number of observations. One way how to do it is to use AIC. This can be calculated as the deviance plus two times number of parameters in the model plus two times number of deleted (or weighted out) observations. We can say that to weight out (or delete) an observation is the same as to use an extra parameter only for this observation. I will try to use this criterion mainly when the models will fit approximately the same number of observations.

I performed my model building separately for family and village random effects. It was not possible to use family and village random effects at once, because SABRE does not allow for more than one level of random effects.

1.4.1 Briefly about responses

There are five response variables in the dataset: enrolment to school, admission to school, delay in going to school, delay in going to school in 1993 and abandoning school.

Enrolment, admission and abandoning school are binary responses and I used logistic regression with random effects to model them. In the case of abandoning school, the random effects model did not lead to any improvement and I do not present any new models for this response.

Delays are nonnegative integers and I used Poisson type log-linear models with normal random effects to model them.

1.4.2 Explanatory variables

Most of the explanatory variables in the dataset are binary and there is not any problem with them. However, there are some variables which can take more than two values. Some of them, for example level of father's education, were included as continuous, others, like type of school in the village, were replaced

by series of binary variables. Variable which can take n different values is for the analyses replaced by n binary variables and I investigate significance of these binary variables separately. This method is very flexible because it allows to keep in the model only variables which are really important. For example variable TYPESCH (type of school in the village), which can take four different values, was replaced by four binary variables TYPESCH 1,. . . ,TYPESCH 4, indicators for each level of variable TYPESCH.

1.4.3 Interpretation of end-points

The interpretation of end-point parameters can cause some difficulties. From end-point parameters it is possible to calculate probabilities p_0 and p_1. From formula 1.8 we can see that probabilities p_i are a bit smaller than end-points.

The probability p_0 is actually estimated proportion of population (in our case population of villages or families) which are stayers in state zero, i.e. some villages in which children do not go to school. However, it is necessary to say that in the mover–stayer model there are two kinds of villages in which children do not go to school: stayers in zero where children are not able to go to school and movers where children have chance to go to school, but it just happened that no children go there.

I suggest to be very cautious with interpretation of these parameters.

1.4.4 Presentation of results

The results of my model building are presented after each step, I present results in six tables for each response variable. First two columns of every table contain model with random effect of village, model with random effect of family can be found in second two columns of the table and in last two columns is the corresponding fixed effects model from Lindsey [1].

Deviance and number of observations for which the model was fitted (observations with missing values were deleted) are given for each model.

The pluses in the model with fixed effects mean that significant variable was not considered for the fixed effects analysis because it has non-random missing values.

The symbol AL[E] which appears in few tables means that corresponding explanatory variable was extrinsicaly aliased and the standard error therefore was not estimated.

Symbol AL[I] appears in the province characteristics and it means that standard error for province Antananarivo was not estimated because this province was chosen as baseline (aliased by design).

Chapter 2

School enrolment

In this chapter, I will try to find random effects models describing reasons why children were registered to school in the year 1993. From this analysis were excluded children who were six years old in the year of investigation (1994), because most of them did not have any chance to be enrolled to school in 1993.

I will compare these random effects models with fixed effects models given in Lindsey [1].

2.1 Children characteristics

Here, I want to find children characteristics important for being registered in school in 1993. Exclusion of observation with missing values in children characteristics leaves 2914 observations from 3693 children aged seven to sixteen in 1994.

The response variable is binary and therefore I am going to use logistic regression. The responses are correlated and to account for dependencies within family or village, I use models with random effects. SABRE cannot fit more than one level of random effects and I have to use only village or only family random effects. I performed separate model building with village and family random effect. Results of these analyses are given together with results of fixed effects model building from report [1] in table 2.1.

It seems that the best model is the one with village random effects. It gives deviance 2154.1 and it is certainly better than model with random effect of family, which gives deviance 2388.5 with one extra explanatory variable (CLOTHES) and one extra end-point parameter. The comparison of random effects models and the fixed effects model is not so straightforward. The fixed effects model is based on a bit different model building strategy and it does not include variable WATER because it has non-random missing values. Anyway, comparison based on AIC (deviance plus two times number of parameters plus

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.017	0.385	3.202	0.732	1.831	0.242
Children						
Age	-0.226	0.024	-0.285	0.032	-0.139	0.016
Water	1.233	0.163	1.530	0.203	+	
Shopping	0.809	0.153	1.262	0.180	0.729	0.089
Sex	-0.515	0.139	-0.799	0.181	-0.316	0.094
Meals	0.486	0.190	0.428	0.242		
Biologic	0.571	0.235	1.126	0.307	0.361	0.156
Rice	0.292	0.163	0.475	0.198	0.699	0.099
Clothes			0.376	0.242	0.495	0.122
Field					-0.178	0.106
Scale	1.792	0.149	1.645	0.381		
End-points						
End-point 0	0.061	0.019	0.097	0.025		
End-point 1	0.000	fixed	0.020	0.190		
Deviance	2154.1		2388.5		3250.1	
Observations	2914		2914		3078	

Table 2.1: Enrolment to school and children characteristics

two times number of deleted observations) suggests that random effects model with random effect of village gives the best fit. Deviance of the fixed effects model is 3250.1 and it fits 3078 observations.

The variable AGE is significant in all models and in all models it has negative sign. It means that children are less likely to go to school with increasing age. This can reflect the fact that older children have to stay at home to help their parents. However, this idea is not supported by other variables which have entered the model.

Variable WATER is not included in the fixed effect model because of its non-random missing values. The positive parameter estimate shows that children going to get water are more likely to be registered in school (or that children who attend school are more likely to fetch water). Other variables with positive influence on school enrolment are: SHOPPING, MEALS (only in random effects models), BIOLOGIC, RICE and CLOTHES (only with family random effects or fixed effects). The fixed effects model suggests negative influence of variable FIELD. This variable was not significant in any of random effects models.

The negative coefficient for variable SEX means that girls are less likely to be enrolled to school. It seems that on Madagascar it is the father who is the head of family.

The estimate of end-point parameters in village random effect model says that $0.061/(1 + 0.061) = 5.7\%$ of villages are stayers in zero where children are not able to be enrolled to school. There are no stayers in one.

End-points in the family random effect model say that 8.67% of families are stayers in zero and that only 1.83% of families are stayers in one (this is

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.194	0.563	3.856	0.699	2.708	0.450
Children						
Age	-0.263	0.033	-0.294	0.038	-0.153	0.021
Water	1.282	0.211	1.288	0.227	+	
Shopping	0.847	0.200	0.811	0.212	0.676	0.112
Sex	-0.512	0.190	-0.551	0.216	-0.334	0.120
Meals	0.697	0.267	0.454	0.300		
Biologic	0.712	0.324	1.152	0.371	0.164	0.245
Rice	0.139	0.208	0.256	0.225	0.642	0.124
Clothes			-0.092	0.276	0.308	0.148
Field					-0.330	0.128
Family						
Religion 3	-0.895	0.241	-1.281	0.284	-1.046	0.145
Religion 2					-0.296	0.140
Mread	0.505	0.204	0.699	0.250	0.473	0.121
Reason 2	-0.515	0.330			0.352	0.147
Reason 3	-0.676	0.416				
Reason 5	-1.009	0.538				
Reason 10			0.761	0.351		
Reason 11			0.630	0.313		
Mntdis	0.007	0.004				
Holidays 3	0.293	0.207				
Holidays 2					-0.327	0.136
Harea					0.009	0.003
Actsoc					0.361	0.110
Famsize					-0.111	0.044
Numbch					0.075	0.048
Fathed					0.168	0.050
Language 2					-0.330	0.141
Scale	1.438	0.178	1.367	0.251		
End-points						
End-point 0	0.037	0.015	0.050	0.013		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	1276.7		1364.3		2158.9	
Observations	1792		1779		2397	

Table 2.2: Enrolment to school and family characteristics

probably the upper class).

2.2 Family characteristics

In this section, I deleted observations which have missing values in family characteristics or in some children characteristic which is in model. The difference of eight observations between data used for model building for village random effect model and for family random effect model is caused only by missing values in variable CLOTHES.

Variables RICE and CLOTHES do not seem to be significant after adding family characteristics. Anyway, according to the model building strategy they will be kept in the model and they can be removed in the variable removal step.

Variable RELIGION 3 is significant in all models. This variable is indicator for other than Protestant or Catholic religion of the head of the family. This is mainly Anglican, Islam or traditional religion. Some of these religions can have different point of view on the necessity of sending children to school than Protestants and Catholics. Variable RELIGION 2, significant only in the fixed effects model, suggests that Catholics are bit less likely to send their children to school than Protestants.

Variable MREAD is also significant in all three models. Positive coefficient shows that children whose mother can read are more likely to attend school.

Reasons 2, 3 and 5 are significant in the model with village random effects. These reasons are "safe way to school", "free books in school" and "school's restaurant" respectively. All these reasons are in the model with negative sign and it seems that parents who send their children to school because of safe way to school, free books in school and school's restaurant are less likely to send their children to school. This can happen probably because the way to school is not safe, books in school are not for free and there is not any restaurant in the school. REASON 2 (safe way to school) is significant also in the fixed effects model.

Reasons 10 and 11, "director–teachers–parents relationship" and "competence of teachers" respectively, are significant only in the model with random effect of family. They have positive sign and parents who care about these reasons are in this model more likely to send their children to school.

Variable MNTDIS, distance from the house to the school in minutes, is important in the model with random effect of village. The positive sign looks a bit strange here. It looks as if the probability of being registered to school increases with distance to school.

Variable HOLIDAYS 3 in the model with village random effect stands for parents who are satisfied with the timing of main school holidays. Their children are more likely to attend school. Variable HOLIDAYS 2 from the fixed effects model suggests that parents who would prefer holidays during harvest period are less likely to enrol their children to school.

Variables HAREA (house area), ACTSOC (whether father of the child takes part in social activities), FAMSIZE (size of the family), NUMBCH (number of children), FATHED (level of education of the father) and LANGUAGE 2 (parents who want their children to learn only French) are significant only in the model with fixed effects. It seems that after accounting for the correlation structure, these variables are not important any more.

End-points seem to become a bit smaller after adding important family characteristics.

2.3 Village characteristics

Resulting models after adding village characteristics are given in table 2.3.

The deviance of the model with random effect of village is 961.8 and it fits 1426 observations. The model with family random effect fits 1691 observations with deviance 1273.2. The deviance of the fixed effects model is 2009.4 with 2216 observations. It would be possible to compare these models with AIC, but the difference in number of observations is in this stage of model building too big to get nice results.

Any village characteristic was not significant in the model with the village random effects. It seems that the random effect of village describes all differences between villages quite well.

In the model with random effect of family, significant variables are AGRV, PROTV and DISTCF. This is something completely different from the fixed effects model, where the significant variables are HERDV and VSIZE.

Variable AGRV says that children are less likely to attend school if they live in mainly agricultural village. This is quite understandable because these children probably have to help their parents with the work on the fields.

Variable PROTV is indicator for villages where the majority of people are Protestants. In the model with family random effects, children from these villages are less likely to be registered in school.

Variable DISTCF is distance to faritany, the district's capital. Child is less likely to be enrolled to school with increasing distance to faritany. Villages closer to faritany are probably more developed and people there more advanced.

In the model with fixed effects, children are less likely to attend school if their village's main activity is breeding animals. The appearance of this variable in the model is a bit surprising, because children characteristics connected with taking care of animals were not significant. Variable VSIZE with positive sign says that in the fixed effects model, children from bigger villages are more likely to be enrolled to school.

The proportion of stayers in zero looks reasonably small in both random effects models. In the family random effect model is a large proportion of stayers in state one. In this model 13.5% of families always register their children to school.

2.4 School characteristics

The models after adding school characteristics are given in table 2.4.

The model with random effect of village still seems to be the best one. Its deviance is 1530.7, deviance of the model with random effect of family is 1646.1 and deviance of the fixed effects model is 1848.5. The number of observations currently in the model is almost the same in all three models.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	4.051	0.720	5.732	1.189	2.575	0.467
Children						
Age	-0.256	0.038	-0.334	0.044	-0.157	0.022
Water	1.274	0.235	1.363	0.259	+	
Shopping	0.881	0.224	0.805	0.241	0.684	0.116
Sex	-0.620	0.211	-0.760	0.244	-0.288	0.124
Meals	0.666	0.296	0.483	0.333		
Biologic	0.634	0.362	0.913	0.417	0.136	0.252
Rice	0.167	0.242	0.386	0.267	0.633	0.129
Clothes			0.390	0.314	0.182	0.156
Field					-0.254	0.132
Family						
Religion 3	-0.848	0.277	-1.476	0.567	-1.155	0.150
Religion 2					-0.292	0.144
Mread	0.410	0.230	0.581	0.283	0.472	0.126
Reason 2	-0.447	0.362			0.393	0.151
Reason 3	-0.923	0.439				
Reason 5	-0.719	0.577				
Reason 10			1.307	0.445		
Reason 11			0.729	0.395		
Mntdis	0.002	0.005				
Holidays 3	0.331	0.237				
Holidays 2					-0.269	0.140
Harea					0.010	0.003
Actsoc					0.339	0.114
Famsize					-0.096	0.045
Numbch					0.062	0.050
Fathed					0.144	0.052
Lang 2					-0.357	0.147
Village						
Agrv			-1.117	0.657		
Protv			-0.457	0.314		
Distcf			-0.001	0.000		
Herdv					-0.182	0.120
Vsize					0.001	0.000
Scale	1.358	0.186	1.196	0.457		
End-points						
End-point 0	0.048	0.019	0.080	0.027		
End-point 1	0.000	fixed	0.169	0.159		
Deviance	961.8		1273.2		2009.4	
Observations	1426		1691		2216	

Table 2.3: Enrolment to school and village characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.779	0.557	6.049	0.889	1.498	0.522
Children						
Age	-0.235	0.029	-0.302	0.036	-0.185	0.023
Water	1.264	0.187	1.458	0.220	+	
Shopping	0.810	0.176	0.839	0.203	0.590	0.121
Sex	-0.495	0.169	-0.648	0.202	-0.301	0.131
Meals	0.535	0.225	0.310	0.274		
Biologic	0.549	0.274	0.628	0.350	0.110	0.264
Rice	0.158	0.191	0.335	0.225	0.635	0.137
Clothes			0.168	0.269	0.340	0.164
Field					-0.202	0.138
Family						
Religion 3	-0.903	0.220	-1.566	0.287	-0.883	0.159
Religion 2					-0.203	0.150
Mread	0.452	0.185	0.782	0.256	0.505	0.134
Reason 2	-0.464	0.304			0.369	0.160
Reason 3	-0.782	0.375				
Reason 5	-0.814	0.467				
Reason 10			0.960	0.361		
Reason 11			0.773	0.321		
Mntdis	0.006	0.004				
Holidays 3	0.080	0.182				
Holidays 2					-0.452	0.148
Harea					0.011	0.003
Actsoc					0.267	0.121
Famsize					-0.072	0.048
Numbch					0.027	0.053
Fathed					0.154	0.056
Lang 2					-0.187	0.156
Village						
Agrv			-1.090	0.446		
Protv			-0.183	0.300		
Distcf			-0.000	0.000		
Herdv					-0.047	0.129
Vsize					0.001	0.000
School						
Typesch 1	-1.989	0.375	-1.949	0.399		
Typesch 2					0.716	0.309
Typesch 3					1.149	0.350
Typesch 4					0.095	0.554
Typecsch 2			-0.823	0.367	-0.200	0.175
Typecsch 3					-1.189	0.304
Numbsch					0.703	0.238
Scale	1.989	0.375	1.995	0.246		
End-points						
End-point 0	0.026	0.013	0.020	0.010		
End-point 1	0.022	0.063	0.000	fixed		
Deviance	1530.7		1646.1		1848.5	
Observations	2166		2196		2191	

Table 2.4: Enrolment to school and school characteristics

Variable TYPESCH 1 is important in models with random effects. It says that children from villages without a school are less likely to be registered in school. This corresponds also to the variables in the fixed effects model, where variables TYPESCH 2, TYPESCH 3 and TYPESCH 4 with positive signs are included instead of variable TYPESCH 1 with negative sign. The interpretation is practically the same.

The model with family random effect and the fixed effects model include variable TYPECSCH 2. This stands for villages with closed public school and it has negative influence on children's attendance to school. In the fixed effects model, variable TYPECSCH 3, presence of closed private school in the village, has also negative influence on children's enrolment to school. Variable NUMBSCH is significant only in the fixed effects models. It says that the child is more likely to be enrolled to school in village with bigger number of school.

After adding school characteristics, there disappeared the estimated large proportion of stayers in one in the model with random effect of family.

2.5 Variable removal

Now it is time to remove nonsignificant variables. All models after removal of unnecessary variables are given in table 2.5.

Removal of variables went as follows. As first, variables which entered the model as the last ones were removed. It means that first I tried to remove school characteristics, then village, family and children characteristics. The criterion for removal of variable was again AIC. Variable was removed if the following change in deviance was less than two.

In the model with random effect of village, variables HOLIDAYS 3 and RICE were removed as unimportant. In the model with random effect of family, variables PROTV and DISTCF were removed. Variables BIOLOGIC, LANGUAGE 2, NUMBCH and HERDV became nonsignificant in the fixed effects model.

The model with random effect of village seems to fit the data best even though it has to pay a big penalty due to deleted observations (if we use AIC to compare these models). It has deviance 1531.6 with 2166 observations, the model with random effect of family has deviance 1649.8 with 2196 observations. The deviance of the fixed effects model is 2041.6 with 2390 observations.

End-point parameters look reasonably small. About 2% of villages are stayers in zero or one and a bit less than 2% of families are stayers in one.

2.6 Provinces

The final models are given in table 2.6. Random effects models are almost the same as in the last section, the only difference is that now they use more

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.785	0.546	5.641	0.849	1.317	0.376
Children						
Age	-0.233	0.029	-0.292	0.035	-0.186	0.022
Water	1.306	0.181	1.504	0.219	+	
Shopping	0.808	0.176	0.815	0.203	0.566	0.114
Sex	-0.468	0.166	-0.549	0.187	-0.342	0.124
Meals	0.556	0.224				
Biologic	0.562	0.224	0.646	0.348		
Rice			0.408	0.220	0.564	0.128
Clothes					0.423	0.154
Field					-0.262	0.132
Family						
Religion 3	-0.902	0.219	-1.601	0.279	-0.875	0.150
Religion 2					-0.229	0.142
Mread	0.445	0.184	0.772	0.255	0.517	0.125
Reason 2	-0.456	0.304			0.346	0.150
Reason 3	-0.800	0.374				
Reason 5	-0.837	0.467				
Reason 10			0.927	0.357		
Reason 11			0.775	0.319		
Mntdis	0.006	0.004				
Holidays 2					-0.424	0.142
Harea					0.011	0.003
Actsoc					0.250	0.114
Famsize					-0.039	0.021
Fathed					0.183	0.052
Village						
Agrv			-1.123	0.448		
Vsize					0.001	0.000
School						
Typesch 1	-1.992	0.374	-1.990	0.391		
Typesch 2					0.708	0.289
Typesch 3					1.044	0.328
Typesch 4					0.076	0.521
Typecsch 2			-0.754	0.357	-0.243	0.162
Typecsch 3					-1.108	0.299
Numbsch					0.686	0.221
Scale	1.363	0.190	2.001	0.237		
End-points						
End-point 0	0.027	0.014	0.019	0.010		
End-point 1	0.027	0.063	0.000	fixed		
Deviance	1531.6		1649.8		2041.6	
Observations	2166		2196		2390	

Table 2.5: Enrolment to school — variable removal

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	4.066	0.603	5.995	0.830	1.695	0.411
Children						
Age	-0.238	0.029	-0.293	0.032	-0.190	0.022
Water	1.317	0.181	1.610	0.213	+	
Shopping	0.809	0.178	0.731	0.192	0.518	0.118
Sex	-0.482	0.164	-0.611	0.175	-0.389	0.126
Meals	0.524	0.225				
Biologic	0.499	0.270	0.580	0.335		
Rice			0.187	0.211	0.535	0.133
Clothes					0.446	0.158
Field					-0.350	0.135
Family						
Religion 3	-0.773	0.227	-1.156	0.284	-0.738	0.161
Religion 2					-0.263	0.147
Mread	0.450	0.182	0.777	0.245	0.419	0.130
Reason 2	-0.355	0.311			0.376	0.160
Reason 3	-0.777	0.373				
Reason 5	-0.678	0.432				
Reason 10			0.823	0.362		
Reason 11			0.244	0.305		
Mntdis	0.005	0.004				
Holidays 2					-0.479	0.146
Harea					0.014	0.003
Famsize					-0.053	0.022
Fathed					0.206	0.053
Village						
Agrv			-1.453	0.460		
School						
Typesch 1	-1.804	0.348	-2.116	0.352		
Typesch 2					0.943	0.295
Typesch 3					1.191	0.338
Typesch 4					0.458	0.535
Typecsch 2			-0.403	0.302	0.005	0.168
Typecsch 3					-1.219	0.319
Numbsch					0.664	0.229
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	-0.668	0.355	-1.402	0.369	-0.580	0.182
Mahajanga	0.549	0.406	0.291	0.404	0.459	0.221
Antisranana	-0.922	0.419	-1.524	0.437	-0.815	0.221
Toamasina	-0.441	0.373	-0.564	0.393	0.010	0.216
Toliara	0.560	0.526	0.868	0.562	0.902	0.301
Scale	1.258	0.177	2.378	0.210		
End-points						
End-point 0	0.024	0.012	0.000	fixed		
End-point 1	0.028	0.047	0.000	fixed		
Deviance	1545.7		1957.4		1982.6	
Observations	2230		2602		2390	

Table 2.6: Enrolment to school and provinces

observations (only observations with missing value in variables in model were deleted) and include effect of province.

All provinces are being compared with province Antananarivo. There are six provinces in the dataset and therefore we need five parameters to include their effects.

Mahajanga and Toliara are provinces which look better than province Antananarivo. The difference between provinces Mahajanga and Antananarivo does not look significant in the model with random effect of family. Provinces Fianarantsoa, Toamasina and Toliara are significantly worse (the difference between provinces Antananarivo and Toliara does not look significant in the fixed effects model).

The effect of provinces is in all models approximately the same.

Now we can compare the model with family random effect and fixed effects model very easily. The model with random effect of family has less parameters (20 instead of 26), fits more observations (2602 instead of 2390) and even though it has better deviance (1957.4 instead of 1982.6). The comparison of the model with random effect of village and the model with random effect of family is not so straightforward. Comparison based on AIC says that the model with random effect of family is the better one. This result is probably due to the big difference in number of observations in models, in previous steps of model building the model with random effect of village seemed to be the better one.

There is small proportion of stayers in zero in the model with random effect of village, end-point 0 parameter extimate is 0.024, this says that 2.3% of villages are stayers in zero. 2.7% of villages are estimated to be stayers in one — villages in which all children are always registered to school and where children do not have any chance to avoid school.

Both end-points are fixed at zero in the model with random effect of family. This model does not include any stayers.

Chapter 3

School admission

In this chapter, I will try to find reasons why children started school in year 1993. Children who were six years old in year 1994 are again excluded from analyses, because most of them were not six years old when school year 1993 started and did not have the chance to start school at that time.

The response variable is binary and therefore I will use logistic regression.

3.1 Children characteristics

In this section, I excluded observations with missing values in children characteristics. The resulting sample size is 989 observations.

The best fitting model is now without any doubt the model with random effect of village. The deviance of this model is 826.3, deviance of the model with random effect of family is 925.3. Deviance of the fixed effects model is even worse, 1040.2, the difference in number of observations is only 20.

The model with village random effect contains only variables AGE, WATER and SHOPPING. The model with family random effect contains these three variables together with BROTHER and ANIMALS. The fixed effects model contains also these three variables together with SEX, MEALS and FIELD.

The probability of starting school decreases with age, as we would expect. The negative influence of age is very similar in all three models. Older children are less likely to be admitted to school.

In this model, we get also very positive influence of fetching water. This is the same result as in the model for enrolment to school in the previous chapter. It has exactly the same interpretation as before: fetching water increases the probability of going to school.

Variable SHOPPING is the last variable common in all three types of models. Children who can do small shopping are again more likely to be admitted to school. This variable can be connected with the attitude of parents towards

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	4.782	0.528	5.874	0.815	3.111	0.335
Children						
Age	-0.543	0.052	-0.733	0.093	-0.413	0.034
Water	1.718	0.280	2.257	0.407	1.237	0.168
Shopping	0.683	0.258	1.394	0.331	0.667	0.153
Brother			0.899	0.460		
Animals			0.756	0.423		
Sex					-0.048	0.156
Meals					-0.586	0.247
Field					-0.393	0.211
Scale	1.285	0.224	1.819	0.484		
End-points						
End-point 0	0.150	0.041	0.179	0.047		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	826.3		925.3		1040.2	
Observations	989		989		1009	

Table 3.1: Admission to school and children characteristics

their children.

Variables BROTHER and ANIMALS appear only in the model with random effect of family. Both have positive influence on children's admittance to school. Variable BROTHER stands for taking care of brother or sister, variable ANIMALS means that child takes care of animals of the family. It seems that taking care of something improves chances of going to school.

Variables SEX, MEALS and FIELD appear only in the fixed effects model. All of them have negative coefficients. Variable SEX is not significant and was left in the model on purpose. Children who prepare meals and children who work on the field are in this model less likely to be admitted to school.

The end-points in both random effect models look quite similar. In the model with random effect of village approximately 13% of villages are stayers in zero, in the model with random effect of family about 15% of families are stayers in zero. This big proportion of stayers in zero is quite surprising.

The models with children characteristics are given in table 3.1.

3.2 Family characteristics

Resulting models after adding family characteristics are given in table 3.2.

Variables RELIGION 3 and MREAD seem to be really important because they are significant in all three types of models. Negative coefficient for RELIGION 3 means that children are less likely to be admitted to school if their father's religion is other than Protestant or Catholic. This negative influence of "other religion" is very similar as in models for enrolment to school.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	6.154	0.886	5.955	1.002	3.328	0.410
Children						
Age	-0.598	0.072	-0.694	0.097	-0.452	0.039
Water	1.418	0.373	1.319	0.439	1.059	0.190
Shopping	1.068	0.349	1.010	0.380	0.471	0.174
Brother			-0.650	0.437		
Animals			-0.159	0.414		
Sex					-0.049	0.178
Meals					-0.494	0.267
Field					-0.439	0.237
Family						
Religion 3	-0.879	0.397	-1.500	0.437	-0.826	0.182
Mread	0.938	0.380	0.957	0.418	0.858	0.180
Fread	-0.904	0.445				
Language 2	-0.721	0.498				
Holidays 2	-0.793	0.431				
Holidays 3			0.706	0.443		
Actsoc					0.383	0.173
Distwat			0.000	0.000	0.000	0.000
Scale	1.214	0.311	1.962	0.540		
End-points						
End-point 0	0.095	0.036	0.009	0.049		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	461.8		476.1		834.3	
Observations	563		536		872	

Table 3.2: Admission to school and family characteristics

Positive coefficient for variable MREAD is also very similar as in models for enrolment to school. The best for being admitted to school is to have a father who is either Catholic or Protestant and a mother who can read. It is interesting that variable FREAD (whether the father of the family can read or not — significant only in the model with village random effects) has negative sign. The negative coefficient could be caused by some father's who prefer to educate their children themselves or some father's who really did not like the school when they were forced to attend it.

Variable LANGUAGE 2 is significant only in the model with random effect of village. Negative sign means that parents who want their children to learn only French do not always admit their children to school.

Variable HOLIDAYS 2 appears also only in model with random effect of village (exactly the same as in the model for school enrolment). Parents who would prefer to have the main school holidays during harvest season are less likely to admit their children to school. Very similar interpretation has variable HOLIDAYS 3 in the model with family random effect. It says that people who are satisfied with the timing of main holidays are more likely to admit their children to school than people who would prefer to have the holidays during

harvest period (binary variable HOLIDAYS 2) or during rainy season (variable HOLIDAYS 1).

Variable ACTSOC appears only in the fixed effects model. It shows that fathers who take part in social and religious activities are more likely to care about the level of education of their children.

Variable DISTWAT is significant in the model with random effect of family and in the fixed effect model. It shows interesting connection between the distance to the nearest source of drinkable water and probability of being admitted to school. The positive sign which suggests that probability of being admitted to school increases with increasing distance to the source of water is very surprising.

The chance of having a village where no children are admitted to school almost 10% in the model with random effect of village. In the model with random effect of family less than 1% of families are stayers in zero which never admit their children to school.

3.3 Village characteristics

The model with characteristics of village is given in table 3.3.

Deviance of the model with random effect of village is 396.8 and it fits 484 observations. It seems to be better than the model with random effect of family, which has deviance 425.9 with 492 observations. The fixed effects model has deviance 788.6 with 831 observations.

Only variable HERDV has entered the model with village random effect. It shows that children are less likely to be admitted to school in villages where the main activity is breeding animals. This variable appears also in the model with random effect of family with very similar coefficient.

Variable VSIZE is significant in the model with random effect of family and in the fixed effects model. It shows quite acceptable fact that probability of being admitted to school increases with the size of the village.

Variables PROTV (whether majority of people in village are Protestants) and AGRV (whether the village is mainly agricultural) appear only in the model with random effect of family. Corresponding parameter estimates suggest that children living in village where majority of people are Protestants are less likely to be admitted to school and that children who live in mainly agricultural villages are more likely to be admitted to school. The positive coefficient for variable AGRV is unexpected. Anyway, this variable will later become non-significant and will be removed in the variable removal step of model building.

The end-point parameters for village random effect slightly increased. The end-point 0 parameter for the model with random effect of family increased from less then 1% to more than 6%. There are 6% of families staying in zero in the current model.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.103	0.976	5.136	1.496	3.222	0.423
Children						
Age	-0.609	0.085	-0.761	0.123	-0.463	0.041
Water	1.809	0.425	1.489	0.502	0.934	0.195
Shopping	0.098	0.400	0.355	0.453	0.439	0.181
Brother			0.334	0.553		
Animals			-0.360	0.557		
Sex					-0.080	0.183
Meals					-0.526	0.278
Field					-0.319	0.242
Family						
Religion 3	-0.530	0.450	-2.094	0.575	-0.903	0.187
Mread	0.522	0.413	0.983	0.532	0.781	0.185
Fread	0.817	0.493				
Language 2	0.196	0.517				
Holidays 2	-0.413	0.475				
Holidays 3			0.568	0.491		
Actsoc					0.440	0.180
Distwat			0.000	0.000	0.000	0.000
Village						
Herdv	-1.165	0.477	-1.156	0.535		
Vsize			0.007	0.003	0.002	0.001
Protv			-1.044	0.597		
Agrv			1.417	0.955		
Scale	1.047	0.468	2.040	0.578		
End-points						
End-point 0	0.112	0.052	0.065	0.046		
End-point 1	0.090	0.077	0.000	fixed		
Deviance	396.8		425.9		788.6	
Observations	484		492		831	

Table 3.3: Admission to school and village characteristics

3.4 School characteristics

In table 3.4, models after adding school characteristics are given.

The model with random effect of village has deviance 624.8 and it fits 782 observations, the model with random effect of family has deviance 590.4 and it fits 697 observations. The fixed effects model with deviance of 740.4 fits 831 observations. Comparison of these models based on AIC would choose as the best one the model with random effect of village.

Both random effects model contain variable TYPESCH 1. This variable was very important also in the model for school enrolment. The variable TYPESCH 1 means that there is not any open school in the village and it has definitely negative influence on admission to school. The same says variable OPSCH in the fixed effects model. OPSCH can be calculated as $1 - \text{TYPESCH 1}$, its existence is due to different coding of variables in report [1]. The positive coefficient for

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.688	0.715	3.772	1.557	2.045	0.468
Children						
Age	-0.583	0.059	-0.700	0.093	-0.486	0.043
Water	1.511	0.317	1.717	0.402	1.051	0.203
Shopping	0.203	0.296	0.557	0.402	0.242	0.188
Brother			0.333	0.442		
Animals			0.109	0.426		
Sex					-0.145	0.190
Meals					-0.323	0.291
Field					-0.304	0.251
Family						
Religion 3	-0.616	0.341	-1.001	0.395	-0.642	0.198
Mread	1.002	0.325	1.186	0.392	0.824	0.193
Fread	0.023	0.346				
Language 2	-0.442	0.417				
Holidays 2	-0.672	0.362				
Holidays 3			0.229	0.369		
Actsoc					0.371	0.186
Distwat			0.000	0.000	0.000	0.000
Village						
Herdv	-0.717	0.365	-0.313	0.429		
Vsize			0.002	0.002	0.002	0.001
Protv			-0.512	0.477		
Agrv			0.987	0.735		
School						
Typesch 1	-2.591	0.534	-1.408	1.059		
Opsch					1.232	0.391
Numbsch			1.085	0.814	0.439	0.269
Scale	1.562	0.310	1.433	0.503		
End-points						
End-point 0	0.035	0.027	0.090	0.045		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	624.8		590.4		740.4	
Observations	782		697		831	

Table 3.4: Admission to school and school characteristics

variable OPSCH has exactly the same interpretation as the negative coefficient for variable TYPESCH 1.

The model with random effect of family and the fixed effects model contain variable NUMBSCH, number of school in the village. The more schools, the better.

Proportions of stayers have decreased again. There are about 3% of villages and less than 1% of families staying in zero, i.e. not admitting their children to school.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.761	0.641	2.874	0.788	2.209	0.442
Children						
Age	-0.586	0.058	-0.686	0.087	-0.507	0.043
Water	1.542	0.313	1.843	0.390	1.043	0.197
Shopping			0.634	0.338		
Field					-0.346	0.238
Family						
Religion 3	-0.646	0.345	-0.917	0.369	-0.671	0.195
Mread	1.007	0.307	1.076	0.373	0.835	0.189
Holidays 2	-0.652	0.362				
Actsoc					0.360	0.183
Distwt			0.000	0.000	0.000	0.000
Village						
Herdv	-0.696	0.367				
Vsize			0.005	0.002	0.001	0.001
School						
Typesch 1	-2.612	0.521				
Opsch					1.817	0.257
Numbsch			2.373	0.468		
Scale	1.563	0.313	1.063	0.418		
End-points						
End-point 0	0.035	0.028	0.136	0.033		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	626.4		593.7		757.4	
Observations	782		697		848	

Table 3.5: Admission to school — variable removal

3.5 Variable removal

In this section, variables which became nonsignificant during adding new variables to the model were removed.

In the model with village random effect, variables SHOPP, FREAD and LANGUAGE 2 were removed as unnecessary. In the model with family random effect, variables BROTHER, ANIMALS, HOLIDAYS 3, HERDV, PROTV, AGRV and TYPESCH 1 became nonsignificant. In the fixed effects model, I removed variables SHOPP, SEX and MEALS.

The model with village random effect has deviance of 626.4 with 782 observations. The deviance of the model with random effect of family is 593.7, but it fits only 697 observations. Deviance of the fixed effect model is 757.4 with 848 observations.

The models after removal of unnecessary variables are given in table 3.5.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.978	0.744	2.969	0.854	2.502	0.502
Children						
Age	-0.591	0.056	-0.687	0.077	-0.496	0.041
Water	1.874	0.315	1.626	0.362	1.227	0.206
Shopping			0.633	0.320		
Field					-0.408	0.242
Family						
Religion 3	-0.559	0.359	-0.752	0.372	-0.403	0.210
Mread	0.897	0.290	1.107	0.358	0.829	0.193
Holidays 2	-0.703	0.362				
Actsoc					0.307	0.184
Distwt			0.000	0.000	0.000	0.000
Village						
Herdv	-0.534	0.364				
Vsize			0.003	0.002		
School						
Typesch 1	-2.584	0.504				
Opsch					1.824	0.258
Numbsch			2.334	0.427		
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	-1.067	0.623	-0.344	0.558	-0.588	0.294
Mahajanga	0.200	0.654	1.400	0.681	0.437	0.363
Antisranana	-1.532	0.661	-0.646	0.611	-1.044	0.350
Toamasina	0.080	0.565	0.102	0.580	0.071	0.329
Toliara	0.274	0.828	1.310	0.934	0.465	0.495
Scale	1.653	0.292	1.333	0.360		
End-points						
End-point 0	0.019	0.026	0.114	0.031		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	680.2		721.9		771.5	
Observations	875		862		889	

Table 3.6: Admission to school and provinces

3.6 Provinces

In table 3.6, final models after including effect of province are given.

All models now fit approximately the same number of observations and we can compare them more easily. Clearly, the best model is the one with random effect of village. The model with random effect of family comes as second. The fixed effects model seems to be the worst one (at least from the point of view of AIC).

Really important children characteristics (significant in all three models) are AGE and WATER. Importance of age in this context is clear. The significance of fetching water if somewhat surprising. The only explanation, I can think about, is that fetching water is a typical home task for children attending school.

RELIGION 3 and MREAD are family characteristics important in all three types of models. Other than Protestant or Catholic religion with negative and mother's ability to read with positive sign.

Important village characteristics look quite different in different types of models.

School characteristics are closer to each other. Variable TYPESCH 1 and OPSCH have actually the same meaning (only the sign is different) and the variable NUMBSCH (number of schools) is also their very close relative.

Comparison of provinces shows that provinces Fianarantsoa and Antisranana have worse admission to school than province Antananarivo. Province Mahajanga (and maybe province Toliara) is better than province Antananarivo. Province Toamasina does not show any significant difference from province Antananarivo.

Chapter 4

Delay in starting school

In this chapter, delay in starting school in years 1989—1993 will be analyzed.

For this analysis, I can use data about 2447 children who started school during this period. From variables considered in previous analyses, I am not going to use variable AGE, because of the obvious deterministic connection between age and delay in going to school. On the other hand, I will include variable YEAR, which can describe linear trend in delays in these years. This variable is defined as zero for children who started school in year 1989, as one for children who started school in year 1990 and so on.

Children in Madagascar should be admitted to school at the age of six. However, in the dataset were some children who seemed to start school earlier. Children who started school at five years were included in analysis as children with no delay. Children who started school in the age of four, three or even two years were considered as coding errors and were excluded.

4.1 Children characteristics

Exclusion of observations with missing values in any of children characteristics leaves 2072 observations from the total 2447 observations.

The best fit is provided by the model with random effect of village. It has deviance 2758.4 and it fits 2072 observations. The model with random effect of family is slightly worse. Deviance of this model is worse ,2788.2, and it estimates one extra (end-point) parameter. The deviance of the fixed effects model is 3492.7 and this model gives the worst fit.

Many variables in the fixed effects model were not included in the model, because of their non-random missing values. However, if these variables were kept in the model, all three models would contain exactly the same variables.

Variable YEAR has negative sign and in all models it shows decrease in delay in starting school during years 1989—1993.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.827	0.057	0.846	0.058	0.789	0.032
Children						
Year	-0.074	0.012	-0.074	0.013	-0.057	0.010
Field	0.242	0.036	0.235	0.039	+	
Meals	0.157	0.042	0.114	0.046	+	
Animals	0.179	0.039	0.167	0.042	+	
Water	0.110	0.036	0.112	0.039	+	
Clothes	0.110	0.042	0.118	0.045	0.231	0.032
Shopping	-0.099	0.034	-0.102	0.035	+	
Rice	0.086	0.036	0.059	0.038	+	
Sex	-0.059	0.035	-0.063	0.038	-0.133	0.028
Scale	0.367	0.022	0.407	0.024		
End-points						
End-point 0	0.000	fixed	0.012	0.007		
Deviance	2758.4		2788.2		3492.7	
Observations	2072		2072		2232	

Table 4.1: Delay and children characteristics

The "most significant" variable seems to be the variable FIELD. Children who have to work on a field, have delay on average $e^{0.242} = 1.75$ with the village random effect (or $e^{0.235} = 1.72$ with the family random effect) times longer than children who do not have to work on the field.

Other variables with positive sign and therefore negative[1] influence on delay are: MEALS, ANIMALS, WATER, CLOTHES and RICE. Children who have to prepare meals, take care of animals, go to get water, must wash clothes or prepare rice, delay more than children who do not have to do these home tasks.

Children characteristic which makes delay shorter is SHOPPING. Children allowed to do some small shopping delay approximately 1.15 times less than children who do not do any shopping.

Variable SEX with negative sign means that girls delay less than boys. The reason for this can be that boys have to work at home harder than girls.

Estimate of end-point parameter in the model with village random effect is zero. The proportion of stayers in state 0 in the model with family random effect is a bit greater than 1% and it means that in this model are 1.2% of families who always admit children to school without any delay.

4.2 Family characteristics

The models for delay after adding family characteristics are given in table 4.2. Number of obsevations is the same for both random effects models, because

[1] With negative influence I mean that it makes delay worse — longer.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.919	0.096	0.956	0.105	0.980	0.084
Children						
Year	-0.089	0.015	-0.077	0.015	-0.064	0.011
Field	0.194	0.046	0.191	0.047	+	
Meals	0.083	0.055	0.049	0.058	+	
Animals	0.115	0.049	0.110	0.052	+	
Water	0.074	0.047	0.096	0.048	+	
Clothes	0.109	0.053	0.098	0.054	0.214	0.038
Shopping	-0.139	0.044	-0.144	0.044	+	
Rice	0.114	0.046	0.110	0.047	+	
Sex	-0.091	0.046	-0.071	0.048	-0.142	0.033
Family						
Numbch	0.041	0.009	0.025	0.011	0.026	0.007
Fathed	-0.057	0.020	-0.093	0.022	-0.095	0.014
Harea	-0.002	0.001	-0.001	0.000	-0.002	0.001
Reason 7	-0.000	AL[E]	-11.200	133.0		
Actsoc	-0.081	0.046	-0.083	0.049		
Language 1	0.182	0.120	0.203	0.129		
Language 3					-0.222	0.073
Holidays 1			0.098	0.050	+	
Mntdis					0.002	0.001
Scale	0.378	0.027	0.374	0.029		
End-points						
End-point 0	0.000	fixed	0.010	0.009		
Deviance	1697.8		1741.4		2600.8	
Observations	1323		1323		1730	

Table 4.2: Delay and family characteristics

these models contain exactly the same children characteristics. Important family characteristics look also very similar.

Deviance of the model with village random effect is 1697.4 and it fits 1323 observations, the model with family random effect has deviance 1741.4 with 1741.4 observations. The fixed effects model seems to be the worst one. It has deviance 2600.8 and it fits 1730 observations.

Variables NUMBCH, FATHED and HAREA are included in all three models. The coefficients fot these variables are also similar, especially for the model with random effect of family and fixed effects model. Variable NUMBCH, number of children in the family, says that delay increases with increasing number of children in the family. Variable FATHED is the level of education of the father of the family. It can take values between zero and six, so children of man with second cycle of superior education delay 2 times less in the model with random effect of village and 3.6 times less in the other models than children of a man who never went to school. Variable HAREA is the area of the house in square metres and it seems to reflect the wealth of the family. Delay in starting school becomes shorter with increasing area of the house.

All three models also include some variable describing language preferences of parents. Variable LANGUAGE 1 in models with random effects shows that parents who prefer to learn their children only Malagasy send their children to school later and variable LANGUAGE 3 in fixed effects model says that parents who prefer to learn their children both Malagasy and French send their children to school earlier. This can reflect the ambitions of parents of the child. More ambitious parents who want their children to speak more languages are likely to admit them to school earlier.

Only models with random effects contain variable ACTSOC. It has negative sign and we can say that children go to school earlier if their father takes part in various social and religious activities.

Variable HOLIDAYS 1 is significant in the model with family random effects. Parents who would prefer main holidays during rainy season enrol their children to school later. Variable HOLIDAYS was significant also in the fixed effect analysis, but was not included in the model, because it has non-random missing values.

Variable MNTDIS is significant only in the fixed effects model, saying that delay is increasing with the distance to school (in minutes).

Variable REASON 7 (school's opening time) reduced deviance significantly, but SABRE produced confusing estimates of parameter and its standard error. This is due probably to insufficient number of people who send children to school because of the school's opening time. I excluded this variable from later analyses, even though it is significant here. The parameter estimates do not make any sense and only cause some numerical problems in SABRE.

Estimates of end-points parameters remain almost the same as in the previous section.

4.3 Village characteristics

In table 4.3, all models after adding village characteristics are given.

The model with random effect of village has deviance 1630.1. The model with random effect of family has almost the same deviance, 1631.5, but it fits less observations with more explanatory variables. If we use AIC to compare all three models (numbers of observations do not look very different), we conclude that the model with random effect of village now gives the best fit.

Variables DISTCF, CATHV and AGRV seem to be the most important, because they are significant in all three types of models. However, the variable AGRV was not included in the fixed effects model, because it has non-random missing values.

The delay in enrolment is increasing with distance to faritany, district's capital (variable DISTCF). Villages closer to faritany are probably more advanced and people there are more educated.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.571	0.180	0.640	0.157	0.983	0.106
Children						
Year	-0.060	0.016	-0.069	0.017	-0.059	0.013
Field	0.235	0.048	0.201	0.052	+	
Meals	0.150	0.056	0.106	0.061	+	
Animals	0.166	0.052	0.147	0.055	+	
Water	0.137	0.049	0.093	0.050	+	
Clothes	0.083	0.057	0.083	0.060	0.167	0.045
Shopping	-0.084	0.046	-0.090	0.046	+	
Rice	0.076	0.049	0.049	0.050	+	
Sex	-0.082	0.046	-0.102	0.049	-0.150	0.038
Family						
Numbch	0.029	0.010	0.024	0.011	0.024	0.008
Fathed	-0.052	0.019	-0.058	0.023	-0.086	0.016
Harea	-0.002	0.000	-0.001	0.001	-0.002	0.001
Actsoc	-0.069	0.047	-0.098	0.053		
Language 1	0.164	0.113	0.222	0.131		
Language 3					-0.248	0.092
Holidays 1			0.087	0.054	+	
Mntdis					0.002	0.001
Village						
Distcf	0.000	0.000	0.000	0.000	0.000	0.000
Cathv	0.180	0.074	0.104	0.058	0.091	0.037
Protv			-0.116	0.065		
Agrv	0.208	0.118	0.194	0.106	+	
Merch	-0.135	0.086			-0.086	0.038
Herdv					0.092	0.038
Scale	0.347	0.033	0.400	0.029		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	1630.1		1631.5		1953.8	
Observations	1236		1223		1304	

Table 4.3: Delay and village characteristics

Variable CATHV with positive sign says that in villages where majority of people are Catholics, children have on average longer delay. I do not have any idea what the possible reason for this could be. The model with random effect of family contains also variable PROTV, which shows that in villages with majority of Protestants children delay less.

Variable AGRV is indicator for mainly agricultural villages. In these villages, children have longer delay in going to school. The reason could be that these children have to work on field. This interpretation is supported also by significant children characteristics FIELD.

Variable MERCH is significant in the model with random effect of village and in the fixed effects model. It suggests that delay in starting school is shorter in villages with a shop. The significance of children characteristic SHOPPING

looks quite interesting in this context.

Variable HERDV (whether the village's main activity is breeding animals) is significant only in the fixed effects model. It shows that children in these villages have longer delay, because they have to take care of animals (significant variable ANIMALS).

End-point parameter again did not change much, in both random effects models it is now fixed at zero. It means that in these models, there are not any families or villages sending a priori all children to school without any delay.

4.4 School characteristics

In this step, important school characteristics were included. The models are given in table 4.4.

Variable TYPECSCH 3 is the only one significant school characteristic and it is significant only in the model with random effect of family. TYPECSCH 3 means that there is closed private school in the village. Children in villages with closed private school have in this model longer delay than children in other villages.

In the model with random effect of village, all previously included variables still seem to be important. Variables HOLIDAYS 1 and RICE in the model with random effect of family look a bit suspicious. Their importance will be checked in the next step, removal of unnecessary variables.

The end-point parameters are in both models with random effects still fixed at zero.

From the point of view of AIC, the model with family random effect now seems to give the best fit. Both random effects models look better than the fixed effects model.

4.5 Variable removal

In this section, variables which became nonsignificant after adding other groups of variables were removed. The results are given in table 4.5.

Only variable CLOTHES was removed from children characteristics and it was only in the model with random effect of family.

None of the family characteristics has been removed.

Removed village characteristics were DISTCF and AGRV for model with village random effect and PROTV and AGRV in the model with family random effect. It seems that variable AGRV was not included in the fixed effects model for good reasons.

Both models with random effects now contain sixteen explanatory variables, fixed effects model contains only twelve explanatory variables, but a lot of significant variables have not been considered there, because they have non-random

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.594	0.172	0.702	0.136	0.983	0.106
Children						
Year	-0.065	0.015	-0.077	0.015	-0.059	0.013
Field	0.218	0.045	0.217	0.046	+	
Meals	0.163	0.052	0.128	0.054	+	
Animals	0.197	0.048	0.165	0.048	+	
Water	0.166	0.046	0.129	0.045	+	
Clothes	0.079	0.053	0.067	0.052	0.167	0.045
Shopping	-0.119	0.043	-0.101	0.041	+	
Rice	0.069	0.045	0.047	0.044	+	
Sex	-0.070	0.043	-0.078	0.043	-0.150	0.038
Family						
Numbch	0.037	0.009	0.027	0.010	0.024	0.008
Fathed	-0.055	0.018	-0.070	0.020	-0.086	0.016
Harea	-0.002	0.001	-0.001	0.001	-0.002	0.001
Actsoc	-0.085	0.044	-0.093	0.046		
Language 1	0.189	0.109	0.197	0.119		
Language 3					-0.248	0.092
Holidays 1			0.059	0.047	+	
Mntdis					0.002	0.001
Village						
Distcf	0.000	0.000	0.000	0.000	0.000	0.000
Cathv	0.183	0.075	0.135	0.051	0.091	0.037
Protv			-0.060	0.058		
Agrv	0.160	0.114	0.119	0.096	+	
Merch	-0.096	0.082			-0.086	0.038
Herdv					0.092	0.038
School						
Typecsch 3			0.339	0.140		
Scale	0.346	0.029	0.397	0.025		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	1854.0		2027.4		1953.8	
Observations	1406		1540		1304	

Table 4.4: Delay and school characteristics

missing values.

The model with random effect of family now seems to be the best one. It has deviance 2069.6 with 1574 observations. The model with random effect of village (deviance 1890.5 and 1440 observations) still looks much better than the fixed effects model with deviance 1953.8 and only 1304 observations.

4.6 Provinces

The final model with provinces is given in table 4.6.

Comparison of models with AIC shows that the best fit is given by the

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.824	0.109	0.811	0.092	0.983	0.106
Children						
Year	-0.065	0.014	-0.076	0.015	-0.059	0.013
Field	0.219	0.044	0.233	0.044	+	
Meals	0.158	0.051	0.138	0.052	+	
Animals	0.194	0.048	0.158	0.047	+	
Water	0.150	0.045	0.124	0.044	+	
Clothes	0.076	0.052			0.167	0.045
Shopping	-0.122	0.042	-0.104	0.040	+	
Rice	0.086	0.045	0.071	0.043	+	
Sex	-0.071	0.043	-0.066	0.042	-0.150	0.038
Family						
Numbch	0.036	0.009	0.027	0.010	0.024	0.008
Fathed	-0.059	0.018	-0.076	0.020	-0.086	0.016
Harea	-0.002	0.001	-0.001	0.001	-0.002	0.001
Actsoc	-0.074	0.043	-0.086	0.045		
Language 1	0.189	0.106	0.212	0.115		
Language 3					-0.248	0.092
Holidays 1			0.068	0.046	+	
Mntdis					0.002	0.001
Village						
Distcf					0.000	0.000
Cathv	0.171	0.069	0.148	0.046	0.091	0.037
Merch	-0.118	0.081			-0.086	0.038
Herdv					0.092	0.038
School						
Typecsch 3			0.378	0.138		
Scale	0.346	0.028	0.393	0.025		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	1890.5		2069.6		1953.8	
Observations	1440		1574		1304	

Table 4.5: Delay — variable removal

model with random effect of family. This model has deviance 2337.6 with 1779 observations. The model with random effect of village has deviance 2087.1 and it fits 1605 observations. The fixed effects model has smaller deviance of 1939.2, but it uses only some 1304 observations. However, the comparison of models with random effects was affected by exclusion of some observations due to variable MERCH in the model with village random effect. This variable now seems to be nonsignificant and removal of this variable could completely change results of this comparison.

In the model with random effect of village, variables MERCH and ACTSOC do not seem to be significant anymore. In the model with random effect of family, variables HOLIDAYS 1 and ACTSOC also lost something from their previous significance. The reason can be that province effects which now entered

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.772	0.116	0.880	0.100	0.988	0.114
Children						
Year	-0.072	0.014	-0.081	0.014	-0.061	0.013
Field	0.219	0.042	0.230	0.042	+	
Meals	0.177	0.048	0.163	0.048	+	
Animals	0.172	0.046	0.119	0.045	+	
Water	0.137	0.043	0.120	0.042	+	
Clothes	0.082	0.049			0.184	0.045
Shopping	-0.129	0.038	-0.120	0.038	+	
Rice	0.099	0.044	0.091	0.041	+	
Sex	-0.071	0.040	-0.061	0.040	-0.158	0.038
Family						
Numbch	0.033	0.008	0.022	0.009	0.022	0.008
Fathed	-0.059	0.017	-0.080	0.019	-0.090	0.016
Harea	-0.002	0.001	-0.002	0.001	-0.002	0.001
Actsoc	-0.046	0.041	-0.029	0.043		
Language 1	0.136	0.102	0.088	0.108		
Language 3					-0.211	0.094
Holidays 1			0.051	0.043	+	
Mntdis					0.002	0.001
Village						
Distcf					0.000	0.000
Cathv	0.115	0.059	0.096	0.045	0.093	0.039
Merch	-0.026	0.065			-0.090	0.040
Herdv					0.071	0.041
School						
Typecsch 3			0.232	0.128		
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	0.204	0.094	0.128	0.064	-0.050	0.056
Mahajanga	-0.031	0.087	-0.070	0.066	-0.045	0.060
Antisranana	-0.082	0.102	-0.102	0.083	-0.144	0.071
Toamasina	0.233	0.105	0.164	0.073	-0.066	0.069
Toliara	-0.088	0.092	-0.104	0.081	-0.269	0.079
Scale	0.369	0.035	0.381	0.025		
End-points						
End-point 0	0.000	fixed	0.011	0.007		
Deviance	2087.1		2337.6		1939.2	
Observations	1605		1779		1304	

Table 4.6: Delay and provinces

the model explain part of the variability which was explained by these now nonsignificant variables before.

All provinces are being compared with province Antananarivo which was chosen as baseline.

Comparison of provinces shows that province Fianarantsoa has longer delay than province Antananarivo in random effects models, the difference between

these two provinces is not significant in fixed effects model.

Province Mahajanga does not show significant difference in comparison with province Antananarivo, province Antisranana looks almost the same, maybe a bit better.

Province Toamasina shows again different results in fixed and random effects models. For this province fixed effects model does not show any significant difference, random effects models show that children in province Toamasina have longer delay in enrolment to school.

On the contrary, province Toliara was significantly better in the fixed effects model, but it shows only some small nonsignificant improvement in random effects models.

End-point parameters are again almost the same, the end-point parameter in the model with random effect of family again shows small proportion of families staying in zero.

Chapter 5

Delay in 1993

The reasons for starting school late may change during few years and therefore in this chapter, only delay for children who were admitted to school in year 1993 will be analyzed. These analyses should provide most recent information about reasons for late start of school.

In this chapter, variable AGE is not considered from the same reasons as in the previous chapter (because DELAY + 6 = AGE). There were 606 children admitted to school in 1993.

5.1 Children characteristics

We start the model building again with important children characteristics. Resulting models are given in table 5.1.

All three model use almost the same number of observations and we can compare them very easily. The worst one is the fixed effects model. It has deviance 783.6 and it fits 528 observations. Deviance of the model with family random effect is 728.1 with 526 observations. The best fitting model is now the model with village random effects with deviance 715.9 and 526 observations.

There are six variables significant in all three models, two variables are significant in fixed effects model and in the model with random effect of family, and one variable is significant only in the fixed effects model. Fortunately, there is not any problem with non-random missing values.

All variables significant in all three models make the delay in enrolment to school longer. These variables are WATER, RICE, FIELD, MEALS, ANIMALS and CLOTHES. All these variables were also included in the model for delay in years 1989—1993. It seems that child who has to fetch water, prepare rice, work on the field, prepare meals, feed animals and wash clothes of the family, really does not have enough free time to start school already in six years.

Two variables common in the model with family random effect and the

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.273	0.082	0.427	0.083	0.409	0.129
Children						
Water	0.239	0.083	0.178	0.082	0.179	0.097
Rice	0.294	0.083	0.188	0.080	0.167	0.068
Field	0.298	0.095	0.319	0.095	0.289	0.080
Meals	0.239	0.118	0.179	0.120	0.271	0.098
Animals	0.171	0.086	0.185	0.089	0.160	0.078
Clothes	0.159	0.102	0.161	0.107	0.160	0.092
Shopping			-0.128	0.074	-0.101	0.061
Afood			-0.272	0.180	-0.355	0.156
Biologic					0.168	0.112
Sex					0.043	0.066
Scale	0.441	0.052	0.425	0.044		
End-points						
End-point 0	0.004	0.017	0.000	fixed		
Deviance	715.9		728.1		783.6	
Observations	526		526		528	

Table 5.1: Delay in 1993 and children characteristics

fixed effects model are SHOPPING and AFOOD. Both have negative sign and it means that they make delay in admission to school shorter. This is quite unexpected for the variable AFOOD, which means that child has to go to get food for the animals of the family.

Variable BIOLOGIC is significant only in the fixed effects model. Variable SEX was left in the fixed effects model on purpose even though it is not significant.

The estimate of end-point parameter in the model with random effect of village is practically zero, there is less the half percent of villages without any delayed children. Estimate of end-point parameter in the model with family random effect is fixed at zero.

5.2 Family characteristics

In table 5.2, models after including significant family characteristics are given.

Number of observations in models with random effects is almost the same and it is easy to see that the model with village random effects gives better fit than the model with family random effects.

There are three family characteristics common in all three models, two are common only in the two random effects models and one is unique in each model.

The most important family characteristics, common in all models, are number of children, house area and reason 2. Variables NUMBCH and HAREA very highly significant also in the model for delay in 1989—1993. In comparison with

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.272	0.199	0.502	0.198	0.140	0.172
Children						
Water	0.159	0.108	0.180	0.104	0.209	0.074
Rice	0.416	0.102	0.328	0.098	0.160	0.070
Field	0.343	0.116	0.321	0.119	0.238	0.084
Meals	0.119	0.181	0.070	0.191	0.250	0.105
Animals	0.186	0.106	0.204	0.108	0.166	0.081
Clothes	0.034	0.140	0.025	0.144	0.168	0.096
Shopping			-0.116	0.094	-0.127	0.064
Afood			-0.259	0.214	-0.368	0.160
Biologic					0.200	0.131
Sex					-0.081	0.069
Family						
Numbch	0.067	0.019	0.057	0.020	0.056	0.013
Harea	-0.005	0.002	-0.004	0.002	-0.004	0.002
Reason 2	0.252	0.166	0.315	0.173	0.211	0.068
Reason 9			-0.433	0.316		
Actsoc	-0.212	0.100	-0.164	0.094		
Activ	-0.191	0.123	-0.266	0.125		
Religion 3	0.178	0.106				
Mread					-0.118	0.066
Scale	0.410	0.058	0.364	0.058		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	414.0		422.3		707.0	
Observations	333		332		492	

Table 5.2: Delay in 1993 and family characteristics

the model for delay in 1989—1993, I miss only variable FATHED, measuring level of education of the father of the family.

Variable NUMBCH has positive sign and the estimate is very similar in all three models. The positive sign means that children with more siblings delay on average more than other children. Maybe, their parents do not want to put them to school before their older brothers of sisters finish the school.

Variable HAREA is probably related to the wealth of family and larger house area makes the delay shorter.

REASON 2 is "safe way to school". It seems that parents who care about safe way to school enrol their children to school approximately 1.9 times later than parents who do not care about safe way to school so much. These parents are probably afraid to send small children on unsafe way and wait for their children to grow up.

REASON 9 (director–teachers relationship) is significant only in the model with random effect of family. People who care about director–teachers relationship tend to put their children to school earlier than normal people.

Variables ACTSOC and ACTIV are significant only in models with random

effects. Both have negative coefficients and it means that children whose father takes part in different social and religious activities and children whose father is a farmer go to school earlier than other children.

Variable RELIGION 3 is significant only in the model with random effect of village. Positive coefficient suggests that people with other than Catholic or Protestant religion put their children to school later (if they put them to school at all — this variable was highly significant in models for school enrolment and admission).

Variable MREAD is significant only in the fixed effects model. Children whose mother can read show shorter delay than other children.

Both end-point parameters are now fixed at zero.

5.3 Village characteristics

Resulting models after adding village characteristics are given in table 5.3.

The model with village random effects has deviance 353.0 and it still seems to be a bit better than the model with family random effects with deviance 394.6, two extra explanatory variables and twenty extra observations.

Surprisingly, all models show almost the same important village characteristics. In all three models appear variables DISTCF, MARKET and PROTV. Only in the fixed effects model appears variable CATHV.

Comparison with the model for delay in 1989—1993 does not look so bad. The variable MARKET in 1993 looks similar like variable MERCH in 1989—1993, variable PROTV and CATHV have also similar interpretation and corresponding signs of coefficients. Only surprise is given by variable DISTCF, which in comparison with the other model changed sign and became significant in all three models (in previous chapter, variable DISTCF was removed from two models in the variable removal step).

The negative coefficient for variable DISTCF, distance to faritany, is a bit unexpected. It seems that children from villages farther from faritany delay less than children from other villages.

Interpretation of variable MARKET is the same as interpretation of variable MERCH in model for delay 1989—1993. Children from more wealthy villages show shorter delay than children from other villages.

Negative coefficients for variable PROTV shows that children from fillages where majority of people are Protestants have shorter delay than other children. This corresponds to the variable CATHV in the model for years 1989—1993, which showed longer delay for children in Catholic villages. The fixed effects model also here includes variable CATHV, which shows longer delay for children in mostly Catholic villages.

End-points in both models are fixed at zero, there is not any family or village staying at zero.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.410	0.245	0.779	0.232	0.308	0.208
Children						
Water	0.262	0.112	0.237	0.107	0.195	0.083
Rice	0.340	0.111	0.224	0.102	0.185	0.078
Field	0.376	0.127	0.355	0.126	0.246	0.093
Meals	0.388	0.168	0.350	0.169	0.317	0.119
Animals	0.150	0.125	0.176	0.121	0.133	0.090
Clothes	0.116	0.149	0.209	0.136	0.178	0.105
Shopping			-0.121	0.096	-0.102	0.071
Afood			-0.084	0.222	-0.209	0.187
Biologic					0.262	0.159
Sex					-0.087	0.074
Family						
Numbch	0.054	0.022	0.035	0.022	0.046	0.015
Harea	-0.004	0.003	-0.004	0.002	-0.004	0.002
Reason 2	0.246	0.193	0.220	0.193	0.234	0.077
Reason 9			-0.317	0.339		
Actsoc	-0.110	0.104	-0.070	0.100		
Activ	-0.076	0.138	-0.185	0.137		
Religion 3	0.203	0.118				
Mread					-0.161	0.073
Village						
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Market	-0.303	0.215	-0.275	0.181	-0.304	0.121
Protv	-0.214	0.152	-0.230	0.129	-0.215	0.093
Cathv					0.104	0.072
Scale	0.390	0.065	0.368	0.062		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	353.0		394.6		573.5	
Observations	279		299		414	

Table 5.3: Delay in 1993 and village characteristics

5.4 School characteristics

Models after adding significant school characteristics are given in table 5.4.

The best fitting model is the model with village random effect. It has deviance 469.8 and it fits 365 observations. If we use AIC, then the model with random effect of family and the fixed effects model look almost the same. The model with family random effect has deviance 537.1 and fits 391 observations. The deviance of the fixed effects model is 563.7 with 406 observations.

All models include different school characteristics, but with similar (a bit unexpected) positive sign. It means that the the delay in going to school is getting longer with increasing number of schools in the village.

Only variable NUMBSCH, number of schools in the village, is significant in the model with village random effect. For every extra school in the village, the

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.103	0.251	0.585	0.199	0.315	0.210
Children						
Water	0.298	0.102	0.234	0.098	0.216	0.085
Rice	0.351	0.099	0.231	0.093	0.181	0.080
Field	0.313	0.110	0.270	0.109	0.242	0.096
Meals	0.364	0.147	0.319	0.148	0.341	0.121
Animals	0.158	0.105	0.145	0.105	0.130	0.091
Clothes	0.057	0.133	0.176	0.122	0.195	0.106
Shopping			-0.118	0.085	-0.096	0.073
Afood			-0.117	0.217	-0.180	0.187
Biologic					0.242	0.159
Sex					-0.095	0.075
Family						
Numbch	0.069	0.019	0.046	0.019	0.043	0.015
Harea	-0.004	0.002	-0.004	0.002	-0.004	0.002
Reason 2	0.031	0.157	0.007	0.162	0.216	0.079
Reason 9			-0.148	0.251		
Actsoc	-0.129	0.094	-0.062	0.089		
Activ	-0.097	0.121	-0.086	0.122		
Religion 3	0.079	0.107				
Mread					-0.167	0.074
Village						
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Market	-0.349	0.174	-0.331	0.150	-0.347	0.124
Protv	-0.173	0.137	-0.229	0.115	-0.234	0.099
Cathv					0.115	0.074
School						
Numbsch	0.217	0.115				
Typesch 4			0.314	0.166		
Typesch 2					0.264	0.131
Typecsch 2					0.397	0.264
Scale	0.403	0.060	0.382	0.053		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	469.8		537.1		563.7	
Observations	365		391		406	

Table 5.4: Delay in 1993 and school characteristics

average delay increases $e^{0.217} = 1.65$ times. It probably takes a long time to decide which school is the best one.

Variable TYPESCH 4 is indicator for villages with both private and public school. It seems that children of parents who have to choose between private and public school have delay in going to school $e^{0.314} = 2.1$ times longer than other children.

The fixed effects model includes variables TYPESCH 2 (whether there is an open public school in the village) and TYPECSCH 2 (whether there is a closed public school in the village). Longer delay in villages with closed public school

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.089	0.214	0.522	0.168	0.557	0.144
Children						
Water	0.298	0.100	0.236	0.097	0.190	0.083
Rice	0.337	0.096	0.221	0.092	0.208	0.079
Field	0.329	0.106	0.279	0.107	0.259	0.095
Meals	0.345	0.137	0.314	0.147	0.386	0.117
Clothes			0.175	0.121	0.203	0.106
Shopping			-0.128	0.085		
Sex					-0.123	0.071
Family						
Numbch	0.062	0.018	0.044	0.018	0.044	0.015
Harea	-0.005	0.002	-0.004	0.002	-0.003	0.002
Reason 2					0.235	0.077
Mread					-0.143	0.073
Village						
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Market	-0.371	0.174	-0.359	0.146	-0.344	0.122
Protv			-0.239	0.114	-0.300	0.094
School						
Numbsch	0.163	0.109				
Typesch 4			0.345	0.165		
Typesch 2					0.279	0.130
Scale	0.405	0.057	0.385	0.053		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	484.1		540.0		592.7	
Observations	373		391		410	

Table 5.5: Delay in 1993 — variable removal

looks reasonable and is to be expected, the result for variable TYPESCH 2 (longer delay in villages with open public school) is surprising and I do not have any reasonable explanation.

End-point parameters in both random effects model remain fixed at zero.

5.5 Variable removal

Now it is the right time to remove nonsignificant variables which have entered model in previous stages of model building.

These are variables ANIMALS, CLOTHES, REASON 2, ACTSOC, ACTIV, RELIGION 3 and PROTV in the model with random effects of village, variables ANIMALS, AFOOD, REASON 2, REASON 9, ACTSOC and ACTIV in the model with random effect of village and variables ANIMALS, SHOPPING, AFOOD, BIOLOGIC and TYPECSCH 2 in the fixed effects model.

There still remain some variables with unexpected parameter estimates. Significance of variable TYPESCH 2 in the fixed effects model could be due to

ignoring the correlation structure of data. Negative coefficient for DISTCF (distance to faritany) would deserve a closer look. Now it seems that the delay for children living closer to faritany is longer than for children living farther from faritany; there could be also some connection with the positive coefficient for number of school in the village.

According to AIC, the model with village random effects is the best fitting one. It has deviance 484.1 and it fits 373 observations. The model with family random effects with deviance 540.0 and 391 observations is on the second position. The worst fit is given by the fixed effects model. Its deviance is 592.7 with 410 observations.

5.6 Provinces

The final models after adding province effect are given in table 5.6.

Variable NUMBSCH in the model with village random effects seems to be the only one which does not look significant after provinces have been included in the model.

Significantly worse than province Antananarivo seem to be provinces Fianarantsoa (only in the model with the random effect of village) and Toamasina (only in models with random effects).

Children from province Mahajanga have the same delay in enrolment to school as children in province Antananarivo.

Province Antisranana gives shorter delay than province Antananarivo in the model with random effect of family and fixed effects model. Province Toliara gives significantly shorter delay only in the fixed effects model.

Deviance of the model with village random effects is 595.7 with 446 observations, deviance of the model with the family random effects is 578.4 and 422 observations. Deviance of the fixed effects model is 575.4 with 410 observations. AIC leads to conclusion that the best fitting model is the model with random effect of village.

In both random effects models, end-point parameters are fixed at zero. This means that in these models, no stayers in zero, i.e. families or villages where all children go to school always without any delay, are included.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.271	0.212	0.513	0.186	0.510	0.175
Children						
Water	0.283	0.092	0.247	0.096	0.247	0.086
Rice	0.309	0.087	0.259	0.092	0.261	0.081
Field	0.291	0.100	0.286	0.106	0.247	0.096
Meals	0.285	0.123	0.246	0.141	0.381	0.118
Clothes			0.207	0.121	0.199	0.107
Shopping			-0.164	0.083		
Sex					-0.130	0.071
Family						
Numbch	0.045	0.016	0.037	0.018	0.042	0.015
Harea	-0.044	0.018	-0.003	0.002	-0.004	0.002
Reason 2					0.275	0.079
Mread					-0.147	0.076
Village						
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Market	-0.400	0.171	-0.350	0.136	-0.364	0.124
Protv			-0.213	0.120	-0.261	0.103
School						
Numbsch	0.072	0.098				
Typesch 4			0.254	0.164		
Typesch 2					0.212	0.136
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	0.200	0.098	0.132	0.136	0.092	0.116
Mahajanga	0.033	0.156	-0.048	0.142	-0.023	0.121
Antisranana	-0.113	0.187	-0.273	0.170	-0.334	0.145
Toamasina	0.284	0.187	0.240	0.169	0.125	0.145
Toliara	-0.164	0.225	-0.178	0.194	-0.287	0.173
Scale	0.386	0.225	0.376	0.051		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	595.7		578.4		575.4	
Observations	446		422		410	

Table 5.6: Delay in 1993 and provinces

Chapter 6

Discussion

In this chapter, I just want to discuss results given in previous chapters. I will try to point out some interesting things and to draw some general conclusions about my analyses.

6.1 Comparison of models

The comparison of models based on AIC can be obscured by missing values. Some very good fitting model in which is significant some variable with a lot of missing values may look worse than it actually is, sometimes I had the impression that AIC penalizes observations which were weighted out too much.

If we use AIC, then for enrolment to school and delay in years 1989—1993, the best fitting final model with provinces was the model with family random effects.

On the other hand, the model with random effect of village provided best fit for delay in 1993 and admission to school.

Now it is hard to say which random effects model describes the structure of the data better. However, we can notice that the difference in number of observations is bigger in the final models for enrolment and delay in 1989—1993 than in models for admission and delay in 1993. If we take this fact into account, it seems that the random effect of village in general describes the structure of data better than random effect of family.

These models were compared also in appendix B, where models from report [1] improved by adding random effects can be found. All three types of models gave similar deviance only for abandonment of school. In all other models, the best fit was provided by the model with random effect of village, the model with random effect of family was the second one and the worst fit gave the fixed effects model.

My conclusion is that the model with random effect of village describes the

structure of the data better than other models and that AIC just says that for enrolment to school and for delay in 1989—1993 is better the model with random effect of family, because it does not weight out so many observations and that for admission to school and delay in 1993 is better the model with random effect of village because it better describes the correlation structure of data.

6.2 Most important variables

In this section, I will try to describe results of my analyses from a bit different point of view than in previous chapters. In following text, I will consider variables significant in all three types of models to be most important and I will concentrate on them.

Children characteristics

Children chracteristic important for all responses is fetching water. Children who have to go to get water are more likely to go to school, but they go there with longer delay.

Younger children and children who can do small shopping are more likely to go to school.

During years 1989—1993, delay in starting school became smaller. Longer delay can be due to the sex of a child (boys delay longer) or work on field, preparing of meal or rice, fetching water or taking care of animals. Children who can do small shopping start school with shorter delay. Sex of a child, taking care of animals and doing small shopping do not seem to be important in the year 1993.

Family characteristics

In both model for enrolment and model for admittance, the most important family characteristics are RELIGION 3 and MREAD. Children whose father's religion is other than Protestant or Catholic and whose mother can not read are less likely to go to school.

Number of children and area of the house are important for delay in starting school. Level of father's education seemed to be important only before year 1993.

Village characteristics

There is not any village characteristic important for enrolment or admission to school in all three types of models.

Children in mostly Catholic villages have longer delay before year 1993. Children from villages with market and children from villages farther from faritany in the year 1993 did not delay so long as other children.

School characteristics

Children are less likely to be registered to school if they live in village without any open school.

Any school characteristic is not important for starting school late in all three models. Some models suggested that delay in starting school is increasing with increasing number of schools in the village.

Provinces

If we compare primary education in other provinces with situation in province Antananarivo, then it seems that smaller enrolment with longer delay show provinces Fianarantsoa, Antisranana and Toamasina. Better situation in primary education is in provinces Mahajanga and Toliara.

6.3 Problems

My intention was to follow model building strategy in report [1] as closely as possible and to obtain some comparable results. This was not always possible. I had to change the model building strategy, because I was using different software and it can happen that some differences in the final model are due to these adjustments and not due to including random effects in the model. Other differences can follow from a bit different way of coding of variables.

The dataset is very big and a lot of people spent a lot of work on collecting this data and on putting these data to computer. In such big dataset some errors always appear, but I hope that most of them were already discovered and that I was using almost correct data.

Missing values are big problem in this dataset. Deleting observations with missing values is not the best way how to deal with them. I used this method because I wanted to follow the model building strategy in [1] and [2] and because in other ways (like including indicator variables for missing values) it would take too long time to build some model, especially with random effects.

6.4 The end

I would like to finish this report in some optimistic way and so I will just say that from some models we can see that the system of primary education on Madagascar has significantly improved in last few years and I hope that this report will help to make it even better.

Bibliography

[1] J. K. Lindsey, R. Rakotondrazaka, C. Tibi: *Primary Education in Madagascar: Admission, Attendance, Delay, Drop-outs*, 1996

[2] ChangLin Mei, A. F. Poblete: *Analyses about Malagasy Children in Primary Education*, 1995

[3] On-line manual for SABRE: *SABRE — software for the analysis of binary recurrent events*, `http://www.cas.lancs.ac.uk/software/sabre.html`, 1996

Appendix A

List of variables

Response

- ENROLMENT Whether child was registered at school in 1993
- ADMISSION Whether child was admitted to school in 1993
- DELAY IN 1993 The delay of children admitted to school in 1993
- DELAY IN 1989—1993 The delay of children in 1989—1993
- DROPOUTS Whether child abandoned school in years 1990—1993

Children

- SEX This does not need any explanation
- AGE I believe that this is also self explanatory
- BIOLOGIC Whether the child was born in the family
- WATER Whether the child goes to get water
- RICE Whether the child prepares rice
- BROTHER Whether the child takes care of brothers or sisters
- CLOTHES Whether the child must wash clothes of a family
- ANIMALS Whether the child takes care of animals of a family
- FIELD Whether the child works at field
- MEALS Whether the child helps to prepare meals

- SHOPPING Whether the child does some shopping
- AFOOD Whether the child goes to get food for the animals

Time trend

- YEAR How many years after 1989 was child admitted to school

Family

- HAREA House area in squared meters
- HWALL Whether the material of construction of house is hard
- ACTIV Whether the father of child is farmer
- FAMSIZE Size of the family
- NUMBCH Number of children
- RELIGION Religion of the head of a family
  - 1 Protestant
  - 2 Catholic
  - 3 Other
- ACTSOC Whether the father of a family participates in various social activities
- FATHED Level of education of the father of a family
- MREAD Whether mother of a child reads
- FREAD Whether father of a child reads
- DISTWAT Distance from the house to the nearest source of drinkable water
- MNTDIS Distance to school (in minutes)
- REASON Why the parents decided to register their child to school
  - 1 School not far away
  - 2 Safe way to the school
  - 3 Free books in school
  - 4 Free material in school

  - 5 School's restaurant
  - 6 School's schedule
  - 7 School's opening time
  - 8 (unknown)
  - 9 Director–teachers relationship
  - 10 Director–teachers–parents relationship
  - 11 Competence of teachers
- LANGUAGE What language do parents want their children to learn at school
  - 1 Malagasy
  - 2 French
  - 3 Both
- HOLIDAYS When would parents prefer the main school holidays
  - 1 During rainy season
  - 2 During harvest period
  - 3 No changes

Village

- AGRV Whether the village is mainly agricultural
- HERDV Whether the village's main activity is breeding animals
- FISHDV Whether the village's main activity is fishing
- MARKET Whether there is a market in the village
- MERCH Whether there is a shop in the village
- VSIZE Number of houses in the village
- DISTCF Distance to the district's capital (faritany)
- CATHV Whether majority of people in the village are catholics
- PROTV Whether majority of people in the village are protestants

School

- NUMBSCH Number of schools in village
- TYPESCH Type of open school in village
  - 1 No school
  - 2 Public school
  - 3 Private school
  - 4 Both types
- TYPECSCH Type of closed school
  - 1 None
  - 2 Public school
  - 3 Private school

Provinces

- PROVINCE Code for province
  - 1 Antananarivo
  - 2 Fianarantsoa
  - 3 Mahajanga
  - 4 Antisranana
  - 5 Toamasina
  - 6 Toliara

Appendix B

Comparisons of models

Tables containing the same models as in report [1] are in this appendix. The models "improved" with random effect of village are in first two columns of each table, in second two columns models with random effect of family can be found and the original fixed effect models are given in the last two columns.

It is possible to compare deviances, standard errors of estimates, significance of estimates and to draw some conclusion about importance of random effects for modelling these data.

The number of observations should be the same for all models, however, sometimes there are some differences due to discovered errors mainly in coding of missing values. These differences usually are not very big and estimates of parameters do not change much.

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.260	0.359	3.256	0.256	1.831	0.242
Children						
Age	-0.210	0.023	-0.248	0.029	-0.319	0.016
Sex	-0.406	0.132	-0.248	0.029	-0.316	0.094
Biologic	0.412	0.226	0.533	0.275	0.361	0.156
Rice	0.829	0.153	1.076	0.180	0.699	0.099
Clothes	0.449	0.172	0.583	0.214	0.495	0.122
Shopping	1.044	0.143	1.376	0.173	0.729	0.089
Field	-0.026	0.158	0.031	0.185	-0.178	0.106
Scale	1.619	0.125	1.127	0.364		
End-points						
End-point 0	0.080	0.021	0.141	0.033		
End-point 1	0.000	fixed	0.207	0.201		
Deviance	2437.0		2716.9		3250.1	
Observations	3078		3078		3078	

Table B.1: Enrolment and children characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.279	0.587	3.500	0.701	2.708	0.450
Children						
Age	-0.210	0.027	-0.241	0.032	-0.153	0.021
Sex	-0.357	0.154	-0.519	0.174	-0.334	0.120
Biologic	0.374	0.300	0.364	0.358	0.164	0.245
Rice	0.702	0.171	0.966	0.192	0.643	0.124
Clothes	0.234	0.196	0.317	0.223	0.308	0.148
Shopping	0.952	0.166	1.142	0.179	0.676	0.112
Field	-0.142	0.176	-0.232	0.195	-0.330	0.128
Family						
Harea	0.006	0.004	0.008	0.004	0.009	0.003
Famsize	-0.075	0.063	-0.082	0.082	-0.111	0.044
Numbch	0.050	0.067	0.075	0.089	0.075	0.048
Fathed	0.104	0.073	0.258	0.091	0.168	0.050
Mread	0.498	0.176	0.712	0.226	0.473	0.121
Religion 2	-0.024	0.193	-0.336	0.245	-0.296	0.140
Religion 3	-0.740	0.226	-1.386	0.285	-1.046	0.145
Actsoc	0.208	0.165	0.368	0.198	0.361	0.110
Reason 2	-0.375	0.273	-0.220	0.348	0.352	0.147
Holidays 2	-0.476	0.202	-0.527	0.243	-0.327	0.136
Language 2	0.209	0.241	0.145	0.261	-0.330	0.141
Scale	1.155	0.200	1.418	0.226		
End-points						
End-point 0	0.069	0.022	0.051	0.014		
End-point 1	0.036	0.077	0.000	fixed		
Deviance	1841.4		1948.2		2158.9	
Observations	2397		2397		2397	

Table B.2: Enrolment and family characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	3.237	0.609	3.537	0.738	2.575	0.467
Children						
Age	-0.213	0.028	-0.243	0.033	-0.157	0.022
Sex	-0.296	0.159	-0.459	0.180	-0.288	0.124
Biologic	0.360	0.309	0.327	0.370	0.136	0.252
Rice	0.676	0.177	0.930	0.200	0.633	0.129
Clothes	0.172	0.204	0.222	0.236	0.182	0.156
Shopping	0.935	0.169	1.165	0.187	0.684	0.116
Field	-0.116	0.180	-0.160	0.202	-0.254	0.132
Family						
Harea	0.007	0.004	0.010	0.005	0.010	0.003
Famsize	-0.065	0.064	-0.074	0.086	-0.096	0.045
Numbch	0.021	0.069	0.055	0.095	0.062	0.050
Fathed	0.082	0.075	0.214	0.098	0.144	0.052
Mread	0.481	0.181	0.702	0.242	0.472	0.126
Religion 2	0.000	0.198	-0.362	0.259	-0.292	0.144
Religion 3	-0.873	0.236	-1.616	0.302	-0.155	0.150
Actsoc	0.176	0.169	0.328	0.209	0.339	0.114
Reason 2	-0.288	0.287	-0.007	0.372	0.393	0.151
Holidays 2	-0.464	0.208	-0.429	0.255	-0.269	0.140
Language 2	0.234	0.247	0.268	0.281	-0.357	0.147
Village						
Herdv	-0.336	0.248	-0.333	0.227	-0.182	0.120
Vsize	0.000	0.001	0.001	0.001	0.001	0.000
Scale	1.0485	0.179	1.797	0.247		
End-points						
End-point 0	0.078	0.024	0.047	0.015		
End-point 1	0.085	0.073	0.000	fixed		
Deviance	1725.0		1828.3		2009.4	
Observations	2227		2227		2216	

Table B.3: Enrolment and village characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	1.602	0.894	1.777	0.790	1.498	0.522
Children						
Age	-0.229	0.029	-0.262	0.033	-0.185	0.023
Sex	-0.305	0.161	0.463	0.185	-0.301	0.131
Biologic	0.412	0.312	0.347	0.376	0.110	0.264
Rice	0.696	0.180	0.904	0.208	0.635	0.137
Clothes	0.258	0.208	0.289	0.240	0.340	0.164
Shopping	0.887	0.168	1.028	0.189	0.590	0.121
Field	-0.060	0.188	-0.133	0.207	-0.202	0.138
Family						
Harea	0.006	0.004	0.012	0.005	0.011	0.003
Famsize	-0.034	0.066	-0.065	0.084	-0.072	0.048
Numbch	-0.004	0.072	0.004	0.093	0.027	0.053
Fathed	0.102	0.077	0.203	0.105	0.154	0.056
Mread	0.461	0.185	0.816	0.248	0.505	0.134
Religion 2	0.008	0.201	-0.307	0.268	-0.203	0.150
Religion 3	-0.733	0.240	-1.277	0.298	-0.883	0.159
Actsoc	0.201	0.173	0.291	0.220	0.267	0.121
Reason 2	-0.173	0.301	0.077	0.401	0.369	0.160
Holidays 2	-0.441	0.210	-0.562	0.271	-0.452	0.148
Language 2	0.216	0.255	0.155	0.302	-0.187	0.156
Village						
Herdv	-0.145	0.288	-0.202	0.233	-0.047	0.129
Vsize	-0.000	0.001	0.001	0.001	0.001	0.000
School						
Numbsch	0.258	0.386	0.594	0.374	0.703	0.238
Typesch 2	1.791	0.656	1.666	0.547	0.716	0.309
Typesch 3	2.045	0.696	2.514	0.643	1.149	0.350
Typesch 4	1.387	1.038	1.373	0.917	0.095	0.554
Typecsch 2	-0.317	0.383	-0.300	0.344	-0.200	0.175
Typecsch 3	-0.373	0.707	-1.256	0.640	-1.189	0.304
Scale	1.230	0.247	1.731	0.222		
End-points						
End-point 0	0.034	0.016	0.017	0.009		
End-point 1	0.001	0.092	0.000	fixed		
Deviance	1654.9		1715.5		1848.5	
Observations	2202		2202		2191	

Table B.4: Enrolment and school characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	1.753	0.668	1.980	0.646	1.317	0.376
Children						
Age	-0.223	0.027	-0.266	0.032	-0.186	0.022
Sex	-0.338	0.153	-0.508	0.179	-0.342	0.124
Rice	0.717	0.171	0.916	0.202	0.564	0.128
Clothes	0.285	0.200	0.344	0.232	0.423	0.154
Shopping	0.877	0.160	0.997	0.184	0.566	0.114
Field	-0.109	0.177	-0.216	0.201	-0.262	0.132
Family						
Harea	0.006	0.004	0.012	0.005	0.011	0.003
Famsize	-0.037	0.031	-0.043	0.041	-0.039	0.021
Fathed	0.098	0.073	0.257	0.099	0.183	0.052
Mread	0.053	0.174	0.798	0.239	0.517	0.125
Religion 2	0.040	0.190	-0.356	0.264	-0.229	0.142
Religion 3	-0.720	0.229	-1.318	0.288	-0.875	0.150
Actsoc	0.207	0.164	0.272	0.216	0.250	0.114
Reason 2	-0.158	0.288	0.051	0.397	0.346	0.150
Holidays 2	-0.375	0.204	-0.546	0.271	-0.424	0.142
Village						
Vsize	0.000	0.001	0.001	0.001	0.001	0.000
School						
Numbsch	0.322	0.350	0.592	0.364	0.686	0.221
Typesch 2	1.709	0.604	1.589	0.532	0.708	0.289
Typesch 3	1.890	0.688	2.320	0.629	1.044	0.328
Typesch 4	1.247	0.937	1.224	0.898	0.076	0.521
Typecsch 2	-0.299	0.381	-0.483	0.331	-0.243	0.162
Typecsch 3	-0.230	0.701	-1.139	0.633	-1.108	0.299
Scale	1.153	0.187	1.850	0.235		
End-points						
End-point 0	0.045	0.018	0.015	0.010		
End-point 1	0.034	0.069	0.000	fixed		
Deviance	1813.1		1894.0		2041.6	
Observations	2401		2401		2390	

Table B.5: Enrolment — variable removal

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	2.401	0.615	2.593	0.718	1.695	0.411
Children						
Age	-0.220	0.026	-0.262	0.032	-0.190	0.022
Sex	-0.368	0.150	-0.558	0.176	-0.389	0.126
Rice	0.668	0.169	0.859	0.200	0.535	0.133
Clothes	0.350	0.197	0.385	0.231	0.446	0.158
Shopping	0.913	0.159	0.923	0.182	0.518	0.118
Field	-0.114	0.171	-0.406	0.201	-0.350	0.135
Family						
Harea	0.006	0.004	0.015	0.005	0.014	0.003
Famsize	-0.025	0.030	-0.051	0.042	-0.053	0.022
Fathed	0.125	0.071	0.277	0.095	0.206	0.053
Mread	0.487	0.170	0.679	0.245	0.419	0.130
Religion 2	0.013	0.188	-0.449	0.255	-0.263	0.147
Religion 3	-0.589	0.226	-1.159	0.301	-0.738	0.161
Reason 2	-0.242	0.274	-0.109	0.415	0.376	0.160
Holidays 2	-0.362	0.200	-0.696	0.274	-0.479	0.146
Village						
...						
School						
Numbsch	0.143	0.319	0.413	0.360	0.664	0.229
Typesch 2	1.467	0.523	1.661	0.518	0.943	0.295
Typesch 3	1.790	0.632	2.316	0.614	1.191	0.338
Typesch 4	1.168	0.852	1.366	0.901	0.458	0.535
Typecsch 2	-0.304	0.357	-0.352	0.311	0.005	0.168
Typecsch 3	-0.420	0.623	-1.445	0.682	-1.219	0.319
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	-0.662	0.338	-0.918	0.329	-0.580	0.182
Mahajanga	0.572	0.372	-0.885	0.382	0.459	0.221
Antisranana	-0.786	0.469	-1.098	0.413	-0.815	0.221
Toamasina	-0.479	0.391	-0.081	0.385	0.010	0.216
Toliara	0.324	0.454	0.972	0.474	0.902	0.301
Scale	1.224	0.150	1.924	0.216		
End-points						
End-point 0	0.038	0.017	0.000	fixed		
End-point 1	0.000	fixed	0.047	0.044		
Deviance	1881.9		1950.0		1982.6	
Observations	2515		2515		2390	

Table B.6: Enrolment and provinces

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	4.953	0.548	6.107	0.860	3.111	0.335
Children						
Age	-0.566	0.054	-0.728	0.090	-0.413	0.034
Sex	-0.123	0.231	-0.402	0.314	-0.048	0.156
Water	1.7043	0.285	2.433	0.425	1.237	0.168
Meals	0.205	0.392	-0.229	0.623	-0.586	0.247
Shopping	0.701	0.258	1.413	0.340	0.667	0.153
Field	0.335	0.342	-0.153	0.464	-0.393	0.211
Scale	1.307	0.223	2.221	0.603		
End-points						
End-point 0	0.151	0.041	0.124	0.069		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	839.3		946.6		1040.2	
Observations	1009		1009		1009	

Table B.7: Admission and children characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.044	0.635	5.511	0.919	3.328	0.410
Children						
Age	-0.590	0.058	-0.709	0.087	-0.452	0.039
Sex	-0.122	0.245	-0.343	0.318	-0.049	0.178
Water	1.453	0.287	1.882	0.408	1.059	0.190
Meals	0.482	0.428	-0.522	0.479	-0.494	0.267
Shopping	0.561	0.268	0.874	0.329	0.471	0.174
Field	0.289	0.358	-0.368	0.431	-0.439	0.237
Family						
Actsoc	-0.014	0.266	0.558	0.353	0.383	0.173
Mread	0.785	0.279	1.472	0.386	0.858	0.180
Distwt	0.000	0.000	0.001	0.000	0.000	0.000
Religion 3	-0.712	0.321	-1.485	0.395	-0.826	0.182
Scale	1.198	0.263	2.204	0.451		
End-points						
End-point 0	0.114	0.037	0.042	0.051		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	726.2		778.1		834.3	
Observations	872		872		872	

Table B.8: Admission and family characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	4.993	0.644	5.245	0.946	3.222	0.423
Children						
Age	-0.584	0.058	-0.695	0.086	-0.463	0.041
Sex	-0.155	0.247	-0.364	0.335	-0.080	0.183
Water	1.335	0.288	1.689	0.404	0.934	0.195
Meals	0.450	0.435	-0.622	0.488	-0.526	0.278
Shopping	0.569	0.270	0.845	0.330	0.439	0.181
Field	0.297	0.357	-0.242	0.426	-0.319	0.242
Family						
Actsoc	0.078	0.268	0.610	0.388	0.440	0.180
Mread	0.637	0.280	1.249	0.385	0.781	0.185
Distwt	0.000	0.000	0.001	0.000	0.000	0.000
Religion 3	-0.824	0.322	-1.561	0.399	-0.903	0.187
Village						
Vsize	0.000	0.001	0.002	0.001	0.002	0.001
Scale	1.125	0.263	2.022	0.551		
End-points						
End-point 0	0.119	0.039	0.054	0.083		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	697.3		743.6		788.6	
Observations	834		834		831	

Table B.9: Admission and village characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	4.895	0.853	5.152	1.069	2.045	0.468
Children						
Age	-0.590	0.060	-0.725	0.090	-0.486	0.043
Sex	-0.284	0.260	-0.694	0.343	-0.145	0.190
Water	1.370	0.296	1.881	0.399	1.051	0.203
Meals	0.345	0.442	-0.271	0.511	-0.323	0.291
Shopping	0.368	0.280	0.580	0.332	0.242	0.188
Field	0.111	0.357	-0.218	0.429	-0.304	0.251
Family						
Actsoc	0.051	0.276	0.308	0.374	0.371	0.186
Mread	0.852	0.290	1.330	0.388	0.824	0.193
Distwt	0.000	0.000	0.001	0.000	0.000	0.000
Religion 3	-0.772	0.327	-1.087	0.388	-0.642	0.198
Village						
Vsize	0.000	0.001	0.002	0.002	0.002	0.001
School						
Numbsch	0.371	0.466	0.961	0.568	0.439	0.269
Opsch					1.232	0.391
Typesch 1	-2.121	0.722	-1.8475	0.796		
Scale	1.364	0.289	1.612	0.434		
End-points						
End-point 0	0.051	0.032	0.094	0.039		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	660.2		684.7		740.4	
Observations	824		824		831	

Table B.10: Admission and school characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.475	0.639	6.180	0.856	2.209	0.442
Children						
Age	-0.604	0.059	-0.733	0.086	2.209	0.442
Water	1.426	0.291	1.751	0.374	1.043	0.197
Field	0.117	0.352	-0.198	0.402	-0.346	0.238
Family						
Actsoc	0.080	0.274	0.483	0.351	0.360	0.183
Mread	0.872	0.289	1.353	0.366	0.835	0.189
Distwt	0.000	0.000	0.000	0.000	0.000	0.000
Religion 3	-0.780	0.328	-1.157	0.373	-0.671	0.195
Village						
Vsize	0.000	0.001	0.001	0.001	0.001	0.001
School						
Opsch					1.817	0.257
Typesch 1	-2.668	0.497	-3.015	0.538		
Scale	1.455	0.292	1.918	0.466		
End-points						
End-point 0	0.042	0.032	0.039	0.046		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	673.1		705.9		757.4	
Observations	841		841		848	

Table B.11: Admission — variable removal

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	5.590	0.702	6.255	0.884	2.502	0.502
Children						
Age	-0.604	0.058	-00.725	0.085	-0.496	0.041
Water	1.552	0.304	1.910	0.392	1.227	0.206
Field	0.048	0.356	-0.258	0.406	-0.408	0.242
Family						
Actsoc	0.024	0.274	0.424	0.333	0.307	0.184
Mread	0.983	0.288	1.4166	0.378	0.829	0.193
Distwt	0.000	0.000	0.000	0.000	0.000	0.000
Religion 3	-0.459	0.358	-0.757	0.396	-0.403	0.210
Village						
...						
School						
Opsch					1.824	0.258
Typesch 1	-2.485	0.505	-2.920	0.531		
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	-0.796	0.647	-0.702	0.544	-0.588	0.294
Mahajanga	0.665	0.655	0.923	0.694	0.437	0.363
Antisranana	-1.4517	0.710	-1.310	0.631	-1.044	0.350
Toamasina	-0.149	0.611	0.049	0.583	0.071	0.329
Toliara	0.348	0.773	0.858	0.852	0.465	0.495
Scale	1.627	0.323	2.049	0.396		
End-points						
End-point 0	0.014	0.034	0.016	0.032		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	688.9		718.6		771.5	
Observations	876		876		889	

Table B.12: Admission and provinces

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	1.116	0.045	1.115	0.044	0.789	0.032
Children						
Year	-0.113	0.011	-0.113	0.012	-0.057	0.010
Sex	-0.121	0.030	-0.136	0.033	-0.133	0.028
Clothes	0.228	0.038	0.238	0.040	0.231	0.032
Scale	0.360	0.020	0.424	0.022		
End-points						
End-point 0	0.000	fixed	0.017	0.008		
Deviance	3160.2		3163.9		3492.7	
Observations	2232		2232		2232	

Table B.13: Delay in 1989—1993 and children characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	1.201	0.081	1.275	0.089	0.980	0.084
Children						
Year	-0.122	0.013	-0.114	0.013	-0.064	0.011
Sex	-0.143	0.035	-0.141	0.038	-0.142	0.033
Clothes	0.215	0.044	0.217	0.046	0.214	0.038
Family						
Harea	-0.002	0.001	-0.002	0.001	-0.002	0.001
Numbch	0.030	0.008	0.020	0.009	0.026	0.007
Fathed	-0.080	0.016	-0.112	0.019	-0.095	0.014
Mntdis	0.002	0.001	0.002	0.001	0.002	0.001
Language 3	-0.088	0.050	-0.097	0.054	-0.222	0.073
Scale	0.352	0.023	0.386	0.026		
End-point						
End-point 0	0.000	fixed	0.014	0.008		
Deviance	2394.1		2412.5		2600.8	
Observations	1725		1725		1730	

Table B.14: Delay in 1989—1993 and family characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	1.208	0.116	1.261	0.120	0.983	0.106
Children						
Year	-0.119	0.015	-0.110	0.016	-0.059	0.013
Sex	-0.136	0.041	-0.150	0.043	-0.150	0.038
Clothes	0.149	0.052	0.169	0.054	0.167	0.045
Family						
Harea	-0.002	0.001	-0.002	0.001	-0.002	0.001
Numbch	0.023	0.009	0.017	0.011	0.024	0.008
Fathed	-0.083	0.019	-0.103	0.022	-0.086	0.016
Mntdis	0.003	0.001	0.001	0.001	0.002	0.001
Language 3	-0.149	0.062	-0.138	0.065	-0.248	0.092
Village						
Herdv	0.107	0.072	0.100	0.054	0.092	0.038
Merch	-0.108	0.070	-0.104	0.054	-0.086	0.038
Distcf	0.000	0.000	0.000	0.000	0.000	0.000
Cathv	0.153	0.072	0.131	0.052	0.091	0.037
Scale	0.352	0.029	0.380	0.030		
End-points						
End-point 0	0.000	fixed	0.014	0.009		
Deviance	1800.2		1821.9		1953.8	
Observations	1304		1304		1304	

Table B.15: Delay in 1989—1993 and village characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	1.216	0.130	1.316	0.127	0.988	0.114
Children						
Year	-0.120	0.015	-0.111	0.016	-0.061	0.013
Sex	-0.139	0.041	-0.155	0.043	-0.158	0.038
Clothes	0.156	0.052	0.180	0.054	0.184	0.045
Family						
Harea	-0.002	0.001	-0.002	0.001	-0.002	0.001
Numbch	0.024	0.009	0.015	0.011	0.022	0.008
Fathed	-0.085	0.019	-0.108	0.022	-0.090	0.016
Mntdis	0.003	0.001	0.002	0.001	0.002	0.001
Language 3	-0.151	0.062	-0.135	0.067	-0.211	0.094
Village						
Herdv	0.084	0.074	0.075	0.057	0.071	0.041
Merch	-0.103	0.072	-0.100	0.056	-0.090	0.040
Distcf	0.001	0.000	0.000	0.000	0.000	0.000
Cathv	0.153	0.070	0.129	0.053	0.093	0.039
School						
...						
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	0.016	0.097	-0.038	0.077	-0.050	0.056
Mahajanga	-0.004	0.108	-0.051	0.081	-0.045	0.060
Antisranana	-0.123	0.126	-0.146	0.098	-0.144	0.071
Toamasina	-0.132	0.116	-0.720	0.096	-0.066	0.069
Toliara	-0.220	0.128	-0.282	0.106	-0.269	0.079
Scale	0.350	0.029	0.371	0.030		
End-point						
End-point 0	0.000	fixed	0.015	0.009		
Deviance	1795.0		1813.4		1939.2	
Observations	1304		1304		1304	

Table B.16: Delay in 1989—1993 and provinces

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.277	0.151	0.359	0.152	0.409	0.129
Children						
Sex	-0.073	0.075	-0.067	0.078	-0.043	0.066
Biologic	0.113	0.130	0.115	0.133	0.168	0.112
Water	0.235	0.082	0.177	0.082	0.179	0.097
Rice	0.295	0.083	0.196	0.081	0.167	0.068
Clothes	0.179	0.102	0.183	0.109	0.160	0.092
Meals	0.211	0.119	0.188	0.120	0.271	0.098
Shopping	-0.098	0.077	-0.127	0.073	-0.101	0.061
Animals	0.149	0.091	0.155	0.094	0.160	0.078
Field	0.293	0.094	0.305	0.096	0.289	0.080
Afood	-0.174	0.179	-0.282	0.180	-0.355	0.156
Scale	0.437	0.050	0.424	0.044		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	715.7		730.1		783.6	
Observations	528		528		528	

Table B.17: Delay in 1993 and children characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.030	0.201	0.183	0.206	0.140	0.172
Children						
Sex	-0.070	0.078	-0.076	0.080	-0.081	0.069
Biologic	0.111	0.148	0.152	0.151	0.200	0.131
Water	0.303	0.088	0.216	0.087	0.209	0.074
Rice	0.277	0.087	0.198	0.083	0.160	0.070
Clothes	0.183	0.108	0.167	0.113	0.168	0.096
Meals	0.165	0.131	0.183	0.136	0.250	0.105
Shopping	-0.135	0.081	-0.143	0.077	-0.127	0.064
Animals	0.173	0.095	0.152	0.097	0.166	0.081
Field	0.270	0.097	0.263	0.102	0.238	0.084
Afood	-0.237	0.185	-0.284	0.183	-0.368	0.160
Family						
Harea	-0.004	0.002	-0.003	0.002	-0.004	0.002
Numbch	0.075	0.015	0.054	0.017	0.056	0.013
Mread	-0.078	0.086	-0.090	0.083	-0.118	0.066
Reason 2	0.116	0.137	0.170	0.147	0.211	0.068
Scale	0.423	0.051	0.405	0.046		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	648.5		675.2		707.0	
Observations	492		492		492	

Table B.18: Delay in 1993 and family characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.243	0.244	0.304	0.239	0.308	0.208
Children						
Sex	-0.087	0.083	-0.093	0.084	-0.087	0.074
Biologic	0.193	0.173	0.256	0.177	0.262	0.159
Water	0.283	0.096	0.213	0.093	0.195	0.083
Rice	0.297	0.093	0.217	0.090	0.185	0.078
Clothes	0.226	0.119	0.185	0.121	0.178	0.105
Meals	0.213	0.139	0.286	0.143	0.317	0.119
Shopping	-0.131	0.089	-0.116	0.082	-0.102	0.071
Animals	0.137	0.104	0.103	0.106	0.133	0.090
Field	0.260	0.105	0.254	0.108	0.246	0.093
Afood	-0.150	0.211	-0.183	0.210	-0.209	0.187
Family						
Harea	-0.004	0.002	-0.003	0.002	-0.004	0.002
Numbch	0.065	0.017	0.045	0.018	0.046	0.015
Mread	-0.134	0.091	-0.145	0.088	-0.161	0.073
Reason 2	0.077	0.149	0.129	0.157	0.234	0.077
Village						
Market	-0.410	0.169	-0.354	0.141	-0.304	0.121
Distcf	-0.001	0.000	-0.512	0.003	-0.001	0.000
Cathv	0.121	0.109	0.158	0.089	0.104	0.072
Protv	-0.190	0.135	-0.145	0.111	-0.215	0.093
Scale	0.418	0.055	0.371	0.051		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	535.2		557.4		573.5	
Observations	414		414		414	

Table B.19: Delay in 1993 and village characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.338	0.286	0.407	0.273	0.315	0.210
Children						
Sex	-0.092	0.084	-0.100	0.086	-0.095	0.075
Biologic	0.176	0.174	0.237	0.179	0.242	0.159
Water	0.293	0.097	0.218	0.096	0.216	0.085
Rice	0.296	0.096	0.210	0.094	0.181	0.080
Clothes	0.229	0.121	0.191	0.123	0.195	0.106
Meals	0.232	0.143	0.310	0.147	0.341	0.121
Shopping	-0.115	0.092	-0.103	0.085	-0.096	0.073
Animals	0.130	0.106	0.088	0.108	0.130	0.091
Field	0.265	0.107	0.251	0.111	0.242	0.096
Afood	-0.154	0.212	-0.171	0.212	-0.180	0.187
Family						
Harea	-0.004	0.002	-0.003	0.002	-0.004	0.002
Numbch	0.065	0.017	0.044	0.019	0.043	0.015
Mread	-0.132	0.094	-0.158	0.092	-0.167	0.074
Reason 2	0.065	0.150	0.127	0.159	0.216	0.079
Village						
Market	-0.402	0.170	-0.348	0.143	-0.347	0.124
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Cathv	0.107	0.117	0.156	0.094	0.115	0.074
Protv	-0.189	0.139	-0.145	0.116	-0.234	0.099
School						
Typesch 2	-0.102	0.125	-0.087	0.104	0.264	0.131
Typecsch 2	-0.067	0.152	-0.050	0.119	0.397	0.264
Scale	0.423	0.056	0.383	0.052		
End-point						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	529.7		550.1		563.7	
Observations	406		406		406	

Table B.20: Delay in 1993 and school characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.530	0.203	0.668	0.191	0.557	0.144
Children						
Sex	-0.127	0.079	-0.136	0.081	-0.123	0.071
Water	0.268	0.096	0.193	0.096	0.190	0.083
Rice	0.295	0.094	0.240	0.093	0.208	0.079
Clothes	0.252	0.120	0.222	0.124	0.203	0.106
Meals	0.285	0.136	0.335	0.143	0.386	0.117
Field	0.285	0.107	0.268	0.111	0.259	0.095
Family						
Harea	-0.004	0.002	-0.003	0.002	-0.003	0.002
Numbch	0.066	0.017	0.044	0.019	0.044	0.015
Mread	-0.100	0.094	-0.113	0.091	-0.143	0.073
Reason 2	0.057	0.149	0.110	0.158	0.235	0.077
Village						
Market	-0.418	0.169	-0.369	0.144	-0.344	0.122
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Protv	-0.234	0.131	-0.211	0.112	-0.300	0.094
School						
Typesch 2	-0.119	0.117	-0.113	0.100	0.279	0.130
Scale	0.433	0.054	0.403	0.052		
End-point						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	552.5		574.6		592.7	
Observations	410		410		410	

Table B.21: Delay in 1993 — variable removal

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	0.491	0.232	0.597	0.220	0.510	0.175
Children						
Sex	-0.127	0.079	-0.140	0.081	-0.130	0.071
Water	0.303	0.097	0.255	0.098	0.247	0.086
Rice	0.331	0.095	0.285	0.094	0.261	0.081
Clothes	0.245	0.122	0.216	0.125	0.199	0.107
Meals	0.290	0.137	0.350	0.143	0.381	0.118
Field	0.272	0.108	0.243	0.112	0.247	0.096
Family						
Harea	-0.004	0.002	-0.003	0.002	-0.004	0.002
Numbch	0.062	0.018	0.042	0019	0.042	0.015
Mread	-0.093	0.098	-0.101	0.093	-0.147	0.076
Reason 2	0.067	0.149	0.127	0.157	0.275	0.079
Village						
Market	-0.442	0.176	-0.392	0.145	-0.364	0.124
Distcf	-0.001	0.000	-0.001	0.000	-0.001	0.000
Protv	-0.219	0.142	-0.186	0.122	-0.261	0.103
School						
Typesch 2	-0.046	0.120	-0.047	0.101	0.212	0.136
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	0.075	0.173	0.103	0.144	0.092	0.116
Mahajanga	-0.085	0.174	-0.071	0.149	-0.023	0.121
Antisranana	-0.305	0.215	-0.072	0.179	-0.334	0.145
Toamasina	0.126	0.219	0.135	0.177	0.125	0.145
Toliara	-0.266	0.250	-0.220	0.204	-0.287	0.173
Scale	0.420	0.058	0.386	0.053		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
Deviance	546.3		564.8		575.4	
Observations	410		410		410	

Table B.22: Delay in 1993 and provinces

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	-18.243	1.870	-20.414	3.073	-11.760	9.228
Children						
Age	0.284	0.122	0.322	0.153	0.283	0.120
Sex	0.403	0.457	0.428	0.532	0.411	0.444
Biologic	11.903	AL[E]	11.969	AL[E]	5.851	9.128
Water	-0.698	0.424	-0.716	0.498	-0.684	0.410
Clothes	-1.029	0.615	-1.1052	0.707	-1.024	0.601
Meals	-1.001	0.675	-1.140	0.801	-1.000	0.667
Scale	0.000	2.362	1.740	1.233		
End-points						
End-point 0	0.484	0.898	0.000	fixed		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	220.9		220.0		221.3	
Observations	1041		1041		1041	

Table B.23: Dropouts and children characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	-19.575	172.87	-21.261	2.274	-14.470	9.756
Children						
Age	0.267	0.116	0.327	0.133	0.318	0.128
Sex	0.376	0.439	0.554	0.489	0.567	0.460
Biologic	10.925	172.86	12.148	AL[E]	5.924	9.598
Water	-0.783	0.408	-0.670	0.449	-0.650	0.424
Clothes	-1.429	0.611	-1.146	0.637	-1.115	0.609
Meals	-0.389	0.590	-0.972	0.711	-0.958	0.685
Family						
Activ	2.163	0.964	2.090	1.043	2.047	1.028
Religion 1	0.927	0.412	1.010	0.459	0.962	0.425
Scale	0.500	0.496	0.007	3.452		
End-points						
End-point 0	0.000	0.355	0.662	1.367		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	197.4		195.1		195.4	
Observations	927		927		927	

Table B.24: Dropouts and family characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	-19.828	2.272	-20.884	2.895	-12.950	10.97
Children						
Age	0.275	0.135	0.297	0.152	0.275	0.134
Sex	0.358	0.535	0.346	0.573	0.387	0.512
Biologic	11.786	AL[E]	11.962	AL[E]	5.804	10.82
Water	-0.870	0.489	-0.901	0.528	-0.834	0.451
Clothes	-0.925	0.657	-0.990	0.736	-0.933	0.644
Meals	-0.653	0.729	-0.703	0.789	-0.677	0.712
Family						
Activ	1.801	1.049	1.894	1.115	1.789	1.040
Religion 1	0.966	0.473	1.030	0.542	0.949	0.455
Village						
Merch	-0.974	0.484	-1.108	0.577	-0.949	0.453
Scale	0.479	1.134	1.300	1.050		
End-points						
End-point 0	0.000	fixed	0.000	fixed		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	170.9		170.3		171.0	
Observations	830		830		830	

Table B.25: Dropouts and village characteristics

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed effects*	*S.E.*
Intercept	-7.701	2.631	-7.851	1.996	-7.037	1.797
Children						
Age	0.278	0.137	0.270	0.121	0.270	0.130
Water	-1.008	0.535	-0.849	0.427	-0.849	0.439
Clothes	-1.114	0.611	-1.076	0.516	-1.076	0.565
Family						
Activ	1.872	1.072	1.831	0.872	1.827	0.995
Religion 1	1.137	0.506	1.024	0.432	1.024	0.446
Village						
Merch	-1.148	0.529	-1.080	0.436	-1.080	0.439
School						
...						
Scale	0.599	1.532	1.000	0.432		
End-points						
End-point 0	0.382	1.226	0.000	0.328		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	181.5		181.5		179.0	
Observations	841		841		841	

Table B.26: Dropouts — variable removal

In model are	*Village r.e.*	*S.E.*	*Family r.e.*	*S.E.*	*Fixed eff.*	*S.E.*
Intercept	-8.482	2.348	-8.378	2.192	-7.536	1.931
Children						
Age	0.277	0.137	0.278	0.129	0.278	0.135
Water	-0.846	0.509	-0.793	0.458	-0.786	0.456
Clothes	-0.947	0.585	-0.958	0.542	-0.951	0.575
Family						
Activ	1.595	0.057	1.571	0.906	1.578	1.045
Religion 1	1.227	0.502	1.184	0.477	1.190	0.472
Village						
Merch	-0.855	0.493	-0.831	0.468	-0.824	0.460
School						
...						
Province						
Antananarivo	0.000	AL[I]	0.000	AL[I]	0.000	AL[I]
Fianarantsoa	1.228	0.597	1.183	0.554	1.189	0.554
Mahajanga	-0.627	1.131	-0.662	0.978	-0.655	1.108
Antisranana	0.669	0.920	0.614	0.851	0.618	0.874
Toamasina	0.274	0.795	0.247	0.728	0.255	0.759
Toliara	-31558	2E+154	-8459.3	2E+154	-6.147	10.56
Scale	0.538	0.889	0.989	0.489		
End-points						
End-point 0	0.000	fixed	0.005	0.366		
End-point 1	0.000	fixed	0.000	fixed		
Deviance	167.7		169.9		167.8	
Observations	841		841		841	

Table B.27: Dropouts and provinces