

**MAGPLANED**

Programme d'amélioration  
de la planification de l'éducation  
à Madagascar

**EDUCATION PRIMAIRE A MADAGASCAR :  
ADMISSIONS, SCOLARISATION, RETARD, ABANDONS**

**MINISTERE DE L'EDUCATION NATIONALE  
COMMISSION NATIONALE MALGACHE POUR L'UNESCO**

# Préface

Cette étude a été financée dans le cadre du programme, MAGPLANED, Programme d'Amélioration de la Planification de l'Education à Madagascar, et a été coordonnée par l'UNESCO et le Ministère de l'Education nationale de Madagascar. L'étude a été possible grâce à des fonds SAT1 du PNUD permettant, avec d'autres ressources provenant du MAG/91/001 et du 700-MAG-10, d'assurer son financement.

Une partie de l'analyse a été faite par ChangLin Mei et Alberto Ferreira Poblete pendant les mois de juillet, août, et septembre, 1995, au Limburgs Universitair Centrum, Diepenbeek, Belgium, et une autre partie par Patrick Lindsey pendant l'été, 1996.

Les analyses présentées dans ce rapport représentent un travail de plus de six hommes-mois. Un travail supplémentaire de deux mois pour prendre en compte le fait que l'échantillonnage a été fait en grappes, par village et par ménage, n'est pas communiqué dans ce rapport.

Diepenbeek et Liège  
Août, 1996

J.K. Lindsey



# Chapitre 1

## Introduction à l'enquête

### 1.1 Contexte général

L'école primaire à Madagascar comprend cinq années (les classes de 11ème à 7ème, selon le système français), et le secondaire un total de sept années. Dans les années 70, les effectifs scolarisés dans le primaire ont rapidement progressé mais, dans les années 80, les inscriptions ont stagné et ensuite ont commencé à descendre.

Depuis plusieurs années, les effectifs scolarisés dans le primaire à Madagascar tendent à plafonner, avec des transferts importants du secteur public vers le privé. On assiste aussi à une apparente déscolarisation à tous les niveaux, en termes de taux, dont l'une des causes est probablement la crise économique qui affecte le pays. Officiellement, le taux net de scolarisation est de 60%.

En même temps, on voit une préférence croissante pour l'enseignement privé, vraisemblablement parce qu'il donne une image de meilleure qualité, associée à une plus grande discipline. Les événements de 1991 ont par ailleurs sûrement joué un rôle important dans l'accélération des transferts vers le privé, en raison des conséquences des grèves dans l'enseignement public. On peut aussi constater l'état de dénuement des écoles publiques et leur très bas niveau de qualité.

Les finalités et les objectifs de l'enseignement primaire spécifient qu'il doit avant tout être utilitaire, préparant tout Malgache à l'exercice réel d'une profession.

#### 1.1.1 Ecoles

Après avoir augmenté de manière régulière depuis 1975/76, le nombre d'écoles primaires a décliné à partir de 1987/88, mais le nombre d'écoles fonctionnelles avait commencé à diminuer depuis 1982/83. Au total, la baisse atteignait environ 12% en 1992/93 et correspondait à la disparition de près de 1500 écoles. L'infrastructure de l'école, qu'elle soit détruite ou en mauvais état, apparaît comme une cause majeure de fermeture. Le départ des enseignants, qu'ils soient affectés ailleurs et non remplacés, ou qu'ils démissionnent en raison des conditions de travail et de vie, joue lui aussi un rôle très important.

### 1.1.2 Programmes scolaires

Les programmes scolaires de l'enseignement primaire ont fait l'objet de remaniements fréquents. Par exemple, les réaménagements effectués en 1990 et 1992 ont consisté essentiellement à

- éliminer les redondances ;
- enlever les contenus à connotation idéologique ou politique ;
- introduire des thèmes nouveaux relatifs à la protection de l'environnement, à l'hygiène, à l'éducation nutritionnelle, à l'instruction civique, et à la morale ;

Si, auparavant, les programmes étaient exprimés uniquement en termes de contenus et consistaient en une liste de thèmes à étudier par classe et par niveau, à partir de 1985 ils prennent une allure nouvelle avec l'introduction de programmes par objectifs. Les instructions deviennent très détaillées et les directives précises. Il y a des objectifs généraux et spécifiques. Les contenus sont découpés en thèmes répartis sur les 30 semaines de l'année scolaire pour les neuf matières figurant au programme :

1. Malgache,
2. Français,
3. calcul,
4. socio-économie,
5. éducation morale et civique,
6. connaissances usuelles,
7. éducation artistique,
8. activités de production,
9. éducation physique.

En 1991, une réforme proposait de faire du français la langue d'enseignement durant toute la formation de l'élève.

### 1.1.3 Organisation pédagogique

Les horaires d'enseignement sont normalement de 28h30 par semaine, y compris 1h40 de récréation. Mais pour des raisons diverses (effectifs, éloignement de l'école, insuffisance de maîtres, carences nutritionnelles), la majorité des écoles, surtout en milieu rural, ne peuvent fonctionner qu'à mi-temps. Chaque élève ne va alors en classe que pendant une demi-journée. Ainsi, l'horaire hebdomadaire moyen effectif est de 22h.

Trois matières absorbent la moitié de l'horaire : le Malgache, le Français, et le calcul.

Les effectifs par classe sont normalement de 50 au maximum et 25 au minimum. Mais le taux moyen d'encadrement varie entre 20 à Toliara et 69 à Antsiranana. La taille moyenne des écoles varie de 126 élèves à Toliara à 238 élèves à Antananarivo. La faiblesse de cette taille (avec cinq classes) peut s'expliquer par différentes raisons, comme la dispersion de la population.

Les classes multigrades concernent les maîtres qui tiennent, en principe, deux ou trois classes, mais il existe des cas avec jusqu'à cinq classes. En 1992, elles représentaient 17%

de l'ensemble. Les classes fonctionnant à mi-temps se retrouvent surtout dans les écoles à maître unique. Elles représentaient 21% de l'ensemble en 1992.

La plupart des maîtres n'ont pas reçu de formation initiale (ex-volontaires du Service National restés dans l'enseignement) ou en ont reçu des rudiments (formés pendant trois mois dans des centres qui recevaient les titulaires du brevet du 1er cycle de l'enseignement secondaire, BEPC). Statutairement, les enseignants sont tenus d'assurer 27h30 d'enseignement par semaine. Un horaire réduit à 15h30 est proposé aux enseignants dans les écoles fonctionnant à mi-temps. Les directeurs des écoles ayant au moins dix classes sont déchargés de cours.

Chaque mois, beaucoup d'instituteurs doivent s'absenter de trois à dix jours pour rejoindre la ville en taxi-brousse et y percevoir leur salaire.

#### 1.1.4 Elèves

L'enseignement primaire malgache est caractérisé par de faibles taux de promotion et de forts taux de redoublement et d'abandon, surtout dans les zones autres que les grands centres urbains. Les redoublements et les abandons sont très élevés dès la première année.

Il semble que de nombreux élèves, notamment en milieu rural et parmi les enfants originaires des milieux les plus défavorisés, abandonnent peu après le début de l'année scolaire pour des raisons diverses :

- nécessité de participer aux travaux des champs;
- début de la saison des pluies et difficulté d'accès à l'école;
- manque de nourriture pendant la période de soudure (les trois ou quatre mois avant la récolte).

Ces enfants sont souvent inscrits de nouveau à l'école par leurs parents au début de l'année scolaire suivante sous la pression des autorités locales et du chef d'établissement, ou parce que les parents n'ont pas encore perdu l'espoir de les scolariser.

En raison du jeu combiné des abandons et du taux de réussite au CEPE, environ 11% seulement des enfants admis dans le primaire réussissent les cinq années et obtiennent un diplôme (CEPE).

## 1.2 Objectifs de l'étude

L'étude ici décrite a été réalisée pour le Ministère de l'Education nationale durant l'automne de 1994. Elle a impliqué 4012 enfants âgés de six à 16 ans vivant dans 300 villages des six provinces de Madagascar. Les 2514 enfants à l'école primaire en 1993 représentent environ un enfant sur 500 à l'école dans les régions rurales. Un très grand nombre de questions ont été investiguées afin d'inclure des informations sur les enfants, les familles, les villages, et les provinces prélevées.

L'objectif principal de cette étude est d'identifier les déterminants de l'accès, de la fréquentation scolaire, et de l'abandon, en prenant comme variables explicatives les caractéristiques de l'enfant, de la famille, de l'école, et du village.

L'étude repose sur l'hypothèse que les admissions et la fréquentation scolaire sont déterminées par cinq grands groupes de facteurs :

1. des déterminants économiques comme le coût direct de la scolarisation et le coût d'opportunité, considérés du point de vue des possibilités de travail et des revenus;
2. des déterminants socio-culturels, valeurs véhiculées par l'école, la religion, les tabous, ... ;
3. les caractéristiques de l'offre du système éducatif : distance à parcourir et accessibilité physique des écoles environnantes, qualité des services éducatifs, langue d'enseignement, ... ;
4. les résultats scolaires de l'enfant ;
5. certaines conditions locales comme le banditisme qui entraîne la fermeture d'écoles.

Ainsi, dans cette étude, nous avons voulu analyser les différents aspects et caractéristiques de l'éducation primaire à Madagascar. L'étude porte principalement sur l'analyse des facteurs qui influencent la fréquentation scolaire des enfants. Cette étude a quatre objectifs :

1. identifier les facteurs qui influencent l'accès à l'enseignement primaire;
2. expliquer les déterminants de la scolarisation en 1993 ;
3. expliquer les raisons pour lesquelles les enfants commencent l'école avec retard ;
4. identifier les facteurs qui influencent l'abandon de l'école.

Le but de cette étude est non seulement d'identifier les facteurs importants qui influencent l'éducation des enfants, mais aussi de fournir des informations qui pourraient être utiles lors de la prise de décision quant à l'amélioration du système d'éducation de Madagascar.

Des informations ont aussi été collectées sur des facteurs relatifs à la fermeture des écoles. Cet aspect ne sera pas repris dans ce texte.

## **1.3 Description de l'enquête**

Dans cette partie, la méthode utilisée lors de l'enquête est décrite brièvement.

### **1.3.1 Choix de l'échantillon**

Etant donné que l'accessibilité à l'éducation primaire varie entre les zones rurales et urbaines, il aurait fallu mener différentes enquêtes pour prendre en compte les caractéristiques spécifiques de ces deux zones ; il a plutôt été décidé de prendre un échantillon dans les zones rurales où environ 80% de la population habite.

L'unité d'échantillonnage était le ménage, bien que l'analyse elle-même porte sur l'enfant. L'enquête a été réalisée auprès de 1500 familles, incluant ainsi un total de 4012 enfants tous âgés de six à 16 ans. Etant donné qu'un recensement mis à jour de toute cette population ou même des ménages n'était pas disponible au niveau national, il a fallu procéder par grappes. La plus petite unité accessible étant le village, c'est donc elle qui a été choisie comme l'unité de l'échantillonnage primaire.

Afin d'assurer une distribution similaire de l'échantillon à travers l'île, l'échantillon a été stratifié par district (CISCO). Parce que la présence d'une école dans le village est un déterminant important de la fréquentation scolaire dans chaque CISCO, on a choisi de stratifier les villages selon la présence ou l'absence d'une école fonctionnelle.

Des 111 CISCO dans le pays, 104 sont en zone rurale. Une grappe (un village) est composée de cinq familles; 300 villages ont été choisis. Dans les 46 CISCO les plus densément peuplés (c'est-à-dire comprenant plus de 109.000 habitants), quatre villages ont été échantillonnés pour chaque CISCO, dont deux avec et deux sans écoles. Dans les 58 CISCO restant, deux villages ont été sélectionnés pour chacun, un avec et un autre sans école. De cette manière, on obtient l'échantillon composé de 300 villages.

Comme le recensement n'est pas disponible au-dessus du niveau local, les enquêteurs ont du entreprendre un recensement local dans les villages et ont du choisir cinq familles au hasard. Un total de 1500 familles ont donc été obtenues, stratifiées par CISCO et l'existence ou l'absence d'une école.

### **1.3.2 Questionnaires**

Quatre sortes de questionnaires ont été utilisés; ils concernent le village, l'école, la famille, et l'enfant. L'équipe centrale les a préparés et les a testés dans plusieurs villages avant de les traduire du français en malgache.

Ainsi, le questionnaire comporte plusieurs parties. Les renseignements généraux ont été enregistrés dans le questionnaire village. En liaison avec ceci, la description de chaque école fréquentée par les enfants du village était mise dans un questionnaire école (autant que nécessaire par village). Pour autant que possible, des informations sur des écoles fermées dans le village ont aussi été enregistrés. Chaque ménage dans la grappe a du remplir un questionnaire ménage avec sa situation économique, niveau d'éducation, attitude envers l'école, etc. Enfin, un questionnaire enfant a du être rempli pour chaque enfant de six à 16 ans dans chaque ménage de la grappe. On trouve ici les activités spécifiques de l'enfant dans la famille et son parcours scolaire.

### **1.3.3 Période de l'investigation**

Plusieurs contraintes se présentaient pour le choix des dates de l'étude. Dans certaines parties de l'île, la saison des pluies se situe à partir de la mi-novembre. A ce moment, beaucoup de régions deviennent inaccessibles. L'enquête devait être terminée avant cette date.

L'année scolaire commence début octobre et la situation des inscriptions devrait se stabiliser vers la fin du mois. Ainsi, la première possibilité était de mener l'enquête à partir du début octobre. Mais, il ne nous serait resté que quinze jours pour la faire. En plus, à ce moment, le personnel du ministère de l'éducation nationale était fortement occupé par la rentrée et par la carte scolaire.

Enfin, les mois d'octobre et novembre sont aussi la période de semences pour beaucoup de familles qui sont très occupées dans les champs et risquent de ne pas vouloir donner de



temps pour répondre aux questionnaires.

Une deuxième possibilité était de mener l'enquête au mois de septembre, avant la rentrée. Dans ce cas, il fallait enquêter sur la fréquentation pour l'année 93-94, et sur les intentions pour la rentrée. Malgré certains inconvénients (l'oubli du passé), ce choix semblait le meilleur.

### 1.3.4 Enquêteurs

#### Formation

La formation des enquêteurs, assurée par les membres de l'équipe centrale, s'est déroulée dans les six capitales provinciales pendant la semaine du 12 au 16 septembre 1994. Nous avons prévu le recrutement de 300 candidats pour en retenir 150 comme enquêteurs, plus 18 superviseurs (parmi les meilleurs). Dans la mesure du possible, ils venaient de tous les CISCO concernés et avaient de bonnes connaissances de leur région.

Après deux jours de formation en classe, nous avons envoyé les candidats sur le terrain pour deux jours dans des villages. Ainsi, nous avons accueilli les candidats le lundi matin et distribué une copie des questionnaires plus les guides. Pendant la matinée, ils devaient les étudier, pour montrer leurs capacités de responsabilité et de travail individuel. En même temps, l'équipe centrale formait les agents locaux. La formation en salle a eu lieu le lundi après-midi et le mardi. Ensuite, le mercredi et le jeudi, les candidats sont partis dans les villages près des capitales pour apprendre sur le terrain l'administration des questionnaires et pour montrer leurs capacités de contacts avec les villageois.

La sélection des enquêteurs a eu lieu le jeudi soir et le vendredi matin. Les critères étaient

- le travail individuel,
- la participation à la formation en salle,
- l'interaction dans le village, et
- l'état des questionnaires qu'ils avaient remplis.

Enfin, le vendredi après-midi, ils ont reçu les instructions finales, avant de partir dans les villages.

#### Instructions spécifiques

**Les questionnaires** Chaque enquêteur devait visiter deux villages, sauf quelques rares exceptions où les voyages étaient trop longs (200 km à pied) pour tout faire en trois semaines. Chaque enquêteur recevait

1. trois questionnaires village,
2. trois questionnaires école ouverte,
3. deux questionnaires école fermée,
4. 12 questionnaires ménage,
5. 60 questionnaires enfant.

A la réception des questionnaires, chaque enquêteur devait numéroter toutes les feuilles avec les codes CISCO, villages, et ménages, tout en vérifiant que des pages ne manquaient pas. Tout devait être écrit au crayon (par crainte de la pluie qui effacerait l'encre).

**Villages** Les équipes centrales ont apporté des cartes du pays avec elles. Chaque village de l'enquête devait être localisé sur la carte et la carte rapportée à Tananarive pour reconstituer une carte nationale des villages. Si un village s'avérait introuvable, l'équipe devait contacter Tananarive pour le choix d'un village de remplacement. Aucun autre changement de village n'était permis (pas de remplacement des villages isolés!). Heureusement, cette possibilité ne s'est pas présentée.

**Choix des ménages** A l'arrivée dans un village, l'enquêteur devait obtenir la liste des habitants du chef de village ou du président du comité de sécurité (CLS). Il la vérifiait et la complétait si nécessaire. Il numérotait les chefs de ménage, prenait le nombre total, et le divisait par cinq. Ensuite, il choisissait cinq familles sur la liste avec les numéros qui étaient les multiples de ce nombre. Par exemple, avec 34 ménages, on obtient le numéro 6. Ensuite, il prend chaque sixième ménage (dans notre exemple) : 6, 12, 18, 24, 30, pour avoir un total de cinq ménages. Si le ménage choisi n'avait pas d'enfants entre six et 16 ans, il prenait le ménage précédent sur la liste.

**Contrôles** Les enquêteurs recevaient la moitié de leur *per diem* avant de partir dans les villages, et l'autre moitié au retour, après vérification des questionnaires.

L'enquêteur devait obtenir la signature du chef de village à son arrivée et à son départ. Il devait signer chaque questionnaire rempli. Tous les questionnaires étaient contrôlés dans la capitale de la province au retour des enquêteurs.

Nous avons prévu que les 18 superviseurs passeraient dans trois villages chacun, pour un total de 54 villages. Autrement dit, un village sur six devait être contrôlé. Les listes des villages à contrôler dans chaque province ont été choisies à Tananarive et mises dans des enveloppes fermées, à ouvrir après le départ des enquêteurs dans les villages.

## 1.4 Méthodologie

### 1.4.1 Gestion des données

Même s'il semble qu'il y a plus de mille variables dans la banque de données, c'est illusoire. Sous l'influence néfaste de l'école française de statistique, l'équipe locale a scindé beaucoup de variables à plus de deux catégories en groupes de variables binaires, soit dans la questionnaire soit à l'encodage (à la dernière minute, sans consulter l'auteur). Ceci a multiplié inutilement la taille de la banque de données et a introduit beaucoup d'erreurs incontrôlables. (Par exemple, 25 enfants sont classifiés comme venant d'un village où l'activité *principale* est l'agriculture, l'élevage, *et* la pêche.)

Sur un total de 4012 enfants, les observations concernant les enfants de six ans ont été exclues pour plusieurs analyses parce que la moitié n'avait pas commencé à l'école au mo-

ment de l'enquête. (Leur date pour commencer était l'année suivante.) Pour l'ajustement des modèles, les observations avec des valeurs manquantes étaient d'abord éliminées de l'ajustement. Ceci suppose que les valeurs manquent au hasard, mais nous nous sommes rendu compte que ce n'était pas toujours le cas. Ainsi, dans certaines analyses, nous avons du refaire les analyses en ignorant complètement certaines variables avec des réponses manquantes non-aléatoires.

De l'ensemble de variables dans la banque de données, nous avons sélectionné une cinquantaine comme les plus importantes du point de vue de l'influence sur l'éducation des enfants (voir les Annexes A and B pour la description de chaque variable et les codes utilisés dans l'analyse). Si l'on compte les variables utilisées pour créer les variables réponses, presque 100 variables de la banque de données ont été utilisées.

Les variables explicatives se divisent en cinq groupes :

1. variables pour les enfants (âge, sexe, activités à la maison, etc.);
2. caractéristiques de la famille (taille, éducation des parents, taille de la maison, etc.);
3. caractéristiques du village (activité principale, existence d'un marché, etc.);
4. les écoles (type, fermée?);
5. la province.

Dans l'ajustement des modèles, nous avons pris les variables par groupes, en prenant en compte l'importance relative de chaque groupe.

## 1.4.2 Variables réponses

### **Scolarisation**

La scolarisation est définie comme l'inscription effective d'un enfant au cours d'une année scolaire déterminée, dans une école donnée. Elle n'implique pas d'idée de durée ni de fréquentation. Pour une cohorte donnée, on peut définir une probabilité de scolarisation à un âge donné, dans une année d'étude déterminée. Par agrégation sur les années d'étude, on obtient la probabilité de scolarisation à un âge donné. Par agrégation sur les âges, on obtient la probabilité de scolarisation dans une année d'étude. Par la somme sur les années et les âges, on obtient la probabilité globale de scolarisation.

Par contre, on peut définir la fréquentation par le nombre de jours de l'année scolaire pendant lesquels l'enfant a été effectivement présent à l'école rapporté au nombre total de jours de l'année scolaire.

### **Admission**

L'admission est définie comme la première entrée dans l'enseignement primaire. La probabilité d'admission peut être définie par rapports à différents ensembles de base et de différentes manières :

- la probabilité d'admission relative à une cohorte :
  - la probabilité d'admission à un âge donné;
  - la probabilité d'admission au plus tard à tel âge;

- la probabilité globale d’admission d’une cohorte.
- la probabilité d’admission de l’ensemble des enfants, toutes cohortes confondues ;
- la probabilité conditionnelle d’admission, étant donné que l’enfant n’a jamais été inscrit avant.

## **Retard**

Le retard d’admission est défini par rapport à l’âge ‘normal’ que nous prenons comme six ans malgré le fait que certains enfants commencent plus jeunes. Ainsi, par exemple, un enfant qui est admis à l’école primaire pour la première fois à huit ans a deux ans de retard.

## **Redoublement**

Nous ne prenons en compte que le fait de doubler une fois malgré le fait que certains enfants passent plus de deux ans dans la même classe.

## **Abandon**

L’abandon se manifeste par le départ ‘définitif’ d’un élève du système éducatif. Dans la pratique, la difficulté vient du fait que des élèves quittent le système pour une, voire deux années, pour reprendre ensuite leur scolarité. On admettra qu’un élève ayant quitté l’école pour trois ans au minimum aura abandonné.

### **1.4.3 Modèles**

Deux sortes de modèles linéaires généralisés ont été utilisés. Nous avons choisi les modèles logistiques pour l’admission, la scolarisation, le redoublement, et l’abandon parce que les variables réponses sont binaires (les enfants commencent à l’école ou pas, sont scolarisés en 1993 ou pas, redoublent ou pas, abandonnent ou pas).

Quand la variable réponse était un comptage (avec combien d’année de retard l’enfant entre à l’école), nous avons utilisé les modèles log linéaires avec une distribution de Poisson. Ceci explique le logarithme du retard moyen par une fonction linéaire des variables explicatives.

Bien que le plan de l’étude prévoyait 50% de villages sans école fonctionnelle, selon les informations disponible dans la base d’échantillonnage, comme expliqué dans la Section 1.3.1, ceci s’est avéré incorrect. Seulement 18,4% des enfants dans l’étude viennent de villages sans école ouverte (voir la Section A.4). Ainsi, il paraît qu’il y a beaucoup moins d’écoles fermées que ce que montrent les statistiques officielles. Puisque nous n’avons pas d’idée de la proportion correcte de villages sans école fonctionnelle, nous ne pouvons pas corriger nos données. Théoriquement, les villages sans écoles devraient être sur-représentés. Il s’ensuit que nos résultats *sous-estiment* le succès de l’éducation primaire pour les enfants de Madagascar (voir les tableaux à la Section B.5).

Puisque l'étude concerne surtout les enfants eux-mêmes, nous pensons que leurs variables sont les plus importantes, suivies des variables pour le ménage, le village, etc. Ainsi, nous avons d'abord introduit toutes les variables significatives pour les enfants, avant le deuxième groupe, pour la famille, etc. (Nous gardons le sexe dans le modèle jusqu'à la dernière élimination.) C'est seulement quand les quatre groupes de variables étaient dans le modèle que nous avons retiré les variables qui sont devenues non-significatives entre temps. Dans tous les cas, nous regardons les interactions du sexe et de l'âge avec les autres variables.

Nous considérons qu'une variable est significative si la déviance diminue de deux quand la variable entre dans le modèle. (C'est le critère d'information d'Akaike, AIC, pour la sélection des modèles.) Nous avons laissé les cas ambigus dans le modèle jusqu'à la sélection finale, après la quatrième groupe. A ce point, nous avons eu notre modèle final.

Enfin, nous avons ajouté les différences entre les provinces qui constituent le cinquième groupe.

Bien entendu, nous aurions pu obtenir d'autres résultats avec un autre mode d'entrée des variables dans le modèle. Pourtant, notre système semble logique, parce que nous nous intéressons surtout aux facteurs proches des enfants eux-mêmes. Les essais avec d'autres ordres d'introduction des variables ont donné des résultats semblables.

Pour l'analyse, nous avons utilisé le logiciel GLIM.

#### 1.4.4 Présentation des résultats

Nos données contiennent des informations énormes sur les trajectoires scolaires des enfants. Malheureusement, il n'a pas encore été possible de les exploiter. Nous présentons, quand même, des tableaux au début de chaque chapitre pour montrer comment les variables réponses évoluent à travers les cohortes et les âges dans notre étude.

Les résultats pour les modèles sont présentés dans une série de tableaux pour chaque variable réponse, dans des chapitres séparés. Les quatre groupes principaux de variables explicatives sont présentés d'abord, dans l'ordre, avant d'éliminer les variables devenues non-significatives, et enfin d'ajouter les provinces.

Dans tous les cas, nous avons cherché à repérer et mesurer l'influence des variables qui agissent sur la variable réponse en question. S'il s'agit de variables dichotomiques, les coefficients obtenus ont des valeurs absolues qui sont représentatives de l'influence des variables et comparables. En revanche, lorsqu'il s'agit de variables quantitatives, on regardera surtout le signe du coefficient car il n'est pas possible de comparer les valeurs absolues des coefficients étant donné que les unités des variables ne sont pas comparables.

Dans les tableaux,  $N$  signifie le nombre d'observations utilisées et  $M$  le nombre d'observations perdues à cause des valeurs manquantes. Les variables perdues à cause des valeurs manquantes non-aléatoires sont indiquées par une étoile (\*) si elles sont non-significatives et par un plus (+) si elles sont significatives.

## Chapitre 2

# Scolarisation

Dans ce chapitre, nous nous intéressons à l'analyse des raisons pour lesquelles les enfants vont à l'école. De cette façon, nous avons un survol du problème, avant d'analyser l'admission, le retard, et l'abandon. La scolarisation qui nous intéresse se rapporte au fait qu'un enfant s'inscrive à l'école ou pas; nous ignorons si l'enfant a vraiment assisté à l'école, et pour combien de temps.

Du tableau de la scolarisation par âge à la Section B.2, nous voyons que plus de 75% des enfants entre sept et treize ans étaient inscrits à l'école en 1993. Ainsi, nous voyons qu'une très forte majorité des enfants accèdent actuellement à l'école, à un âge ou à un autre. Les enfants qui n'accèdent pas à l'école appartiennent probablement soit à des familles qui rejettent la scolarisation, soit à des familles qui sont dans le plus grand dénuement, soit à des villages reculés et d'accès difficile, qui ne sont pas couverts par le réseau scolaire, soit à des villages dont l'école est fermée.

### 2.1 Analyse par cohorte

Nous regardons d'abord comment la scolarisation a évolué depuis quelques années, en nous rapportant aux données du Tableau 2.1. Si l'on regarde les colonnes de gauche, on voit que les enfants sont admis de plus en plus jeunes à l'école. Par conséquent, plus les cohortes observées sont récentes, plus les taux de scolarisation à un âge donné sont élevés.

Si l'on procède à une lecture par ligne du Tableau 2.1, on observe pour chaque cohorte une progression du taux de scolarisation avec l'âge des élèves concernés. Pour les plus anciennes cohortes, le taux maximal se situe à 12 ans car, ensuite, une partie des élèves abandonne et une autre partie poursuit ses études dans l'enseignement secondaire. Dans les cohortes plus récentes, le maximum se situe lors de l'année d'observation ce qui signifie que le taux de scolarisation peut encore progresser.

TAB. 2.1: Scolarisation (%) par cohorte et âge.

Cohorte	Age										N
	7	8	9	10	11	12	13	14	15	16	
1977			21,6	28,8	40,2	62,5	62,9	63,3	57,2	42,8	264
1978		8,9	24,9	38,6	61,8	64,5	68,9	63,1	49,5		293
1979	7,3	13,8	25,7	53,5	65,7	73,0	67,8	63,8			370
1980	9,2	21,6	47,6	61,2	74,3	80,2	76,6				338
1981	8,5	26,5	43,7	62,7	69,1	71,9					437
1982	20,1	39,3	64,1	75,1	82,2						354
1983	23,5	45,6	66,0	76,5							430
1984	32,3	56,6	76,8								371
1985	45,6	72,9									432
1986	73,5										404

## 2.2 Scolarisation en 1993

Dans nos modèles, nous n'étudions que les enfants qui étaient inscrits au début de l'année académique 1993. Un modèle plus adéquat prendrait en compte la carrière de scolarisation de chaque enfant, mais ceci demanderait d'avantage de travail. Des 3693 enfants de sept à seize ans, 135 allaient à l'école secondaire à l'époque qui nous concerne.

Parce que la variable 'est-ce que l'enfant est inscrit à l'école' est binaire, nous utilisons un modèle logistique : avec une distribution, binomiale, le logarithme du rapport de la probabilité d'aller à l'école sur la probabilité de ne pas y aller en 1993 dépend d'une régression linéaire pour les variables explicatives. Le formule pour ce modèle est

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_K x_K \quad (2.1)$$

où  $\pi$  est la probabilité d'être scolarisé (et  $1 - \pi$  celle de ne pas l'être), les  $x_k$ ,  $k = 1, \dots, K$  représentent les variables explicatives incluses dans le modèle, et les  $\beta_k$  sont les paramètres du modèle.

L'analyse logistique menée au niveau de la scolarisation a conduit à mettre en évidence l'action de plusieurs variables. On rappellera que de nombreuses combinaisons de variables ont été testées pour arriver finalement à un ensemble de variables qui explique le mieux la situation observée sous l'angle du modèle choisi. En ce qui concerne la scolarisation en 1993, on rencontre des variables explicatives dans les cinq ensembles de variables retenues au départ : variables liées à l'enfant, à la famille, au village, à l'école, et à la province.

### 2.2.1 Enfants

Nous commençons avec 3693 observations (tous les enfants qui avaient plus de six ans) et un modèle nul avec une déviance de 4513,7. Selon notre règle pour l'élimination des variables non-significatives énoncée au premier chapitre, sept des 12 variables dans ce groupe sont retenues. (Pour la définition des variables, voir l'Annexe A.) La déviance du nouveau modèle est 3229,7 avec 3068 degrés de liberté pour 3078 observations. A

TAB. 2.2: Scolarisation avec les variables pour les enfants. ( $N = 3078, M = 615$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,705	0,244	WATER <sup>+</sup>
AGE	-0,140	0,017	BROTHER
SEX	-0,024	0,115	AFOOD
BIOLOGIC	0,349	0,157	ANIMALS
RICE	1,008	0,150	MEALS
CLOTHES	0,598	0,126	
SHOPPING	0,735	0,090	
FIELD	-0,072	0,139	
SEX.RICE	-0,530	0,202	
SEX.FIELD	-0,591	0,208	

cet étape, 615 observations sont éliminées à cause des valeurs manquantes pour une ou plusieurs variables explicatives. Les résultats sont dans le Tableau 2.2.

Une des variables, chercher de l'eau, a des valeurs manquantes qui sont apparemment non-aléatoires. Même si elle est très significative quand on l'ajoute au modèle (liée positivement avec la scolarisation), la variable n'est pas retenue pour cette raison.

A partir des ces résultats, nous voyons que l'âge des enfants est lié négativement avec la probabilité de scolarisation. Ceci veut dire que, plus l'enfant est âgé, plus la probabilité d'être scolarisé est faible. Probablement, les enfants plus âgés doivent rester à la maison pour aider leur famille.

Quant au sexe, il paraît que les filles ont moins de chance d'être scolarisées, mais ceci dépend de la participation pour piler le riz et travailler dans les champs.

Si l'enfant est un membre biologique de la famille (autrement, il peut venir d'autres parents de la famille ou de leurs amis), la probabilité d'être scolarisé est plus élevé que pour un enfant qui ne l'est pas. Ceci est surprenant parce que l'on aurait attendu qu'un enfant serait envoyé dans la famille pour fréquenter l'école.

Quand nous regardons les tâches que l'enfant doit accomplir dans le ménage, piler le riz, faire la lessive, et faire les courses sont significatives. Les enfants qui les font ont une probabilité plus grande d'être scolarisés, mais, pour piler le riz, cette influence est plus forte pour les garçons que pour les filles. Apparemment, ce sont des tâches plus légères, plus convenables pour les écoliers. Par contre, le travail dans les champs est lié d'une façon négative avec la scolarisation, surtout pour les filles. Les variables, chercher la nourriture pour les animaux et garder ces animaux n'entrent pas dans le modèle. La participation aux tâches ménagères est peut-être une participation ponctuelle par opposition à celle des autres enfants qui pourraient ne pas fréquenter l'école parce qu'employés à des tâches continues auprès des parents, travaux agricoles en particulier.



TAB. 2.3: Scolarisation avec les variables pour les enfants et la famille. ( $N = 2400$ ,  $M = 1293$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	4,328	0,830	HWALL
AGE	-0,309	0,063	SPMEAT
SEX	-0,054	0,151	ACTIV
BIOLOGIC	0,222	0,246	FREAD
RICE	0,887	0,184	DISTMIN
CLOTHES	0,408	0,152	DISTWAT
SHOPPING	0,695	0,112	
FIELD	-0,089	0,167	
SEX.RICE	-0,418	0,248	
SEX.FIELD	-0,521	0,248	
HAREA	0,008	0,003	
FAMSIZE	-0,101	0,044	
NUMBCH	-0,108	0,108	
FATHED	0,159	0,050	
MREAD	0,480	0,122	
RELIGION2	-0,734	0,651	
RELIGION3	-2,756	0,618	
SOCACT	0,362	0,110	
REASON2	0,386	0,147	
HOLPRF2	-0,321	0,137	
LANGSC2	-0,336	0,141	
AGE.NUMBCH	0,015	0,008	
AGE.RELIGION2	0,037	0,053	
AGE.RELIGION3	0,148	0,051	

## 2.2.2 Famille

Par la suite, nous gardons ces variables significatives pour les enfants dans le modèle et nous ajoutons les 16 variables pour la famille. Après l'exclusion des variables non-significatives, il nous en reste dix. La nouvelle déviance pour notre modèle est 2140,9 avec 2376 degrés de liberté pour 2400 observations. 678 observations sont perdues à cause des valeurs manquantes, pour un total de 1293. Les résultats sont dans le Tableau 2.3. Avant d'analyser les résultats pour les nouvelles variables, il est important de noter que toutes les variables pour les enfants restent significatives avec des estimations similaires, sauf si l'enfant est un membre biologique de la famille.

On peut prendre la surface de la maison familiale comme une mesure de la richesse de la famille. Ainsi, une telle variable peut synthétiser l'action de plusieurs facteurs : richesse de la famille, davantage de place disponible pour le travail scolaire, peut-être l'activité du chef de famille. La probabilité d'être scolarisé augmente avec la taille de la maison. Par contre, étant donné la présence de cette variable dans le modèle, la probabilité ne dépend pas des deux autres variables similaires, le type de murs de la maison et les dépenses pour la viande. Pourtant, le fait que l'on mange davantage de viande est en partie lié à une plus grande richesse mais n'est pas associé avec la scolarisation. Ne serait-ce pas le travail

des enfants qui permet de manger plus de viande dans certaines familles ? Le lien entre scolarisation et 'richesse' de la famille n'est donc pas évident.

La probabilité d'être scolarisé augmente avec le nombre d'enfants (de moins de 16 ans) dans la famille, ceci étant plus important quand l'enfant est plus âgé, mais, en même temps, diminue avec la taille de la famille (ces deux variables sont fortement corrélées : 0,82). Apparemment, les familles avec beaucoup d'adultes ont moins de probabilité d'envoyer les jeunes enfants à l'école.

Comme on peut s'y attendre, la probabilité d'être scolarisé augmente avec le niveau d'éducation du père et quand la mère peut lire. Etant donné le niveau d'éducation du père, le fait qu'il peut lire n'entre pas dans le modèle.

La religion du chef de famille semble aussi entrer dans l'explication de la probabilité d'être scolarisé. Nous trouvons que les protestants ont la plus grande probabilité de scolarisation, suivis de près par les catholiques ; par contre, le groupe des autres religions (musulmane, anglicane, ou traditionnelle) est assez différent de ces deux-ci. La valeur négative du paramètre indique que, pour ces religions, la probabilité d'être scolarisé est beaucoup plus basse. Mais, les différences entre les religions diminuent quand l'enfant est plus âgé.

La probabilité d'être scolarisé dépend aussi du fait que le père participe aux activités sociales ou religieuses du village.

Quand nous regardons la raison principale pour laquelle les parents décident d'envoyer leurs enfants à l'école, nous trouvons que ceux qui donnent les relations entre le directeur de l'école, les enseignants, et les parents ont une probabilité plus élevée d'être scolarisé. D'autres raisons, comme la sécurité sur le chemin de l'école, la distance à l'école, ou l'horaire des leçons ne jouent pas.

Pour une raison quelconque, la probabilité d'être scolarisé est plus basse dans les familles qui préfèrent l'enseignement en français et en malgache que dans celles qui préfèrent l'un ou l'autre. Elle est aussi plus basse pour les familles qui auraient préféré que les grandes vacances se trouvent pendant le moisson plutôt que pendant la saison des pluies ou inchangées.

Les variables non-significatives sont si le père est agriculteur, la distance de la maison à la source d'eau potable, et le temps pour les enfants d'aller à l'école.

### 2.2.3 Village

Après avoir introduit toutes les variables pour le village, et retiré celle qui sont non-significatives, seulement deux des neuf variables restent. A ce moment, deux autres variables, faire la lessive et le nombre d'enfants dans la famille, deviennent non-significatives. Toutes les autres variables dans le modèle restent significatives, avec le même signe. Ainsi, les mêmes conclusions restent valables pour ces variables.

Le nouveau modèle, avec les variables pour l'enfant et pour la famille, en plus de celles pour le village, a une déviance de 2002,8 avec 2203 degrés de liberté pour 2230 observations. A ce point, 180 observations sont retirées, pour un total de 1463. Les résultats se trouvent

TAB. 2.4: Scolarisation avec les variables pour les enfants, la famille, et le village. ( $N = 2230, M = 1463$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	3,610	0,883	DISTF
AGE	-0,262	0,068	SHOP
SEX	-0,016	0,156	MARKET
BIOLOGIC	0,199	0,252	AGRV
RICE	0,856	0,190	FISHV
CLOTHES	0,295	0,160	CATHV
SHOPPING	0,706	0,116	PROTV
FIELD	-0,023	0,173	
SEX.RICE	-0,381	0,257	
SEX.FIELD	-0,511	0,256	
HAREA	0,009	0,003	
FAMSIZE	-0,082	0,046	
NUMBCH	-0,116	0,111	
FATHED	0,142	0,052	
MREAD	0,460	0,126	
RELIGION2	-0,639	0,660	
RELIGION3	-2,698	0,632	
SOACT	0,362	0,114	
REASON2	0,396	0,150	
HOLPRF2	-0,260	0,140	
LANGSC2	-0,374	0,147	
AGE.NUMBCH	0,014	0,008	
AGE.RELIGION2	0,027	0,053	
AGE.RELIGION3	0,133	0,053	
HERDV	-0,148	0,121	
VSIZE	0,004	0,002	
AGE.VSIZE	-0,000	0,000	

dans le Tableau 2.4.

Seulement les variables pour la taille du village et si l'activité principale est l'élevage entrent dans le modèle ici. Les enfants ont plus de chance d'être scolarisé dans les grands villages, mais l'effet diminue avec l'âge, et là où l'activité principale n'est pas l'élevage.

Les variables non-significatives sont la distance au faritany, la présence d'un marchand dans le village, un marché dans le village, l'activité principale du village (agriculture et pêche), et la religion du village.

## 2.2.4 Ecoles

Nous introduisons maintenant plusieurs variables pour les écoles dans le village. Les types des écoles ouvertes et fermées, ainsi que le nombre d'écoles dans le village sont significatives.

La déviance de notre nouveau modèle est 1817,1 avec 2165 degrés de liberté pour 2202 observations. La variable type d'école a 28 réponses manquantes supplémentaires, pour un total de 1491. Les résultats sont présentés dans le Tableau 2.5. Ici, la préférence des parents pour la langue d'enseignement à l'école et le fait que l'activité du village est l'élevage deviennent non-significatives. Par contre, faire la lessive redevient significatif.

Comme on pourrait s'y attendre, la probabilité d'être scolarisé augmente avec le nombre d'écoles dans le village, mais cet effet diminue avec l'âge de l'enfant. Elle est aussi plus haute s'il y a une école privée (type 3) ouverte dans le village, surtout pour les filles, mais plus basse s'il y a une école privée fermée. La scolarisation est d'autant plus forte qu'il y a une offre d'éducation importante.

### 2.2.5 Elimination des variables

A cette étape, nous éliminons les variables qui sont devenues non-significatives, en commençant par celles qui sont entrées le plus récemment. Nous enlevons aussi les variables suivantes : le fait que l'activité principale du village est l'élevage, le choix de la langue à l'école, le nombre d'enfants dans la famille, si l'enfant est un membre biologique de la famille, et plusieurs interactions.

Cette procédure diminue le nombre de valeurs manquantes à 1292, en nous laissant 2401 observations. La déviance est 2018,8 avec 2370 degrés de liberté. Les résultats sont dans le Tableau 2.6.

### 2.2.6 Différences régionales

Quand nous introduisons la variable permettant de prendre en compte les caractéristiques des régions du pays, nous trouvons que la scolarisation varie significativement selon la région. La déviance de notre modèle final est 1951,4 avec 2365 degrés de liberté, toujours avec 2401 observations. Les résultats sont dans le Tableau 2.7.

L'analyse de l'influence des variables géographiques est intéressante. En effet, on a pu constater que la localisation de la famille dans une province plutôt que dans une autre pouvait, parfois, avoir une incidence significative sur la scolarisation. Cette variable prend en compte des spécificités provinciales non explicitées par ailleurs et qui ont parfois des effets significatifs. C'est, par exemple, l'attitude défavorable de notables provinciaux vis à vis de la scolarisation des enfants.

Dans deux des provinces, Fianarantsoa et Antisranana, nous trouvons que, par rapport à Antananarivo (le point de référence), la probabilité d'être scolarisé est significativement plus basse. Par contre, pour Mahajanga et Toliara, elle est plus haute. La probabilité est la même pour Toamasina et Antananarivo.

Ces résultats sont en opposition avec les données brutes dans la Section B.6. Là, la scolarisation à Fianarantsoa, Antisranana, et Toamasina est considérablement inférieure à celle des trois autres provinces. Ainsi, les variables dans le modèle peuvent expliquer la faible scolarisation à Toamasina mais pas dans ces deux autres provinces.

TAB. 2.5: Scolarisation avec les variables pour les enfants, la famille, le village, et les écoles.  
( $N = 2202, M = 1491$ )

Variables significatives	Estimation	Erreur standard
Constante	1,827	1,008
AGE	-0,235	0,077
SEX	-0,050	0,300
BIOLOGIC	0,176	0,265
RICE	0,803	0,198
CLOTHES	0,466	0,169
SHOPPING	0,606	0,122
FIELD	0,047	0,181
SEX.RICE	-0,279	0,272
SEX.FIELD	-0,565	0,270
HAREA	0,010	0,003
FAMSIZE	-0,057	0,048
NUMBCH	-0,111	0,116
FATHED	0,149	0,056
MREAD	0,499	0,136
RELIGION2	-1,024	0,687
RELIGION3	-2,701	0,673
SOACT	0,282	0,122
REASON2	0,358	0,160
HOLPRF2	-0,406	0,151
LANGSC2	-0,226	0,157
AGE.NUMBCH	0,011	0,009
AGE.RELIGION2	0,066	0,055
AGE.RELIGION3	0,154	0,056
HERDV	-0,022	0,131
VSIZE	0,004	0,002
AGE.VSIZE	-0,000	0,000
NUMBSCH	1,487	0,561
TYPESCH2	0,804	0,343
TYPESCH3	0,639	0,411
TYPESCH4	1,063	0,653
TYPECSCH2	-0,170	0,176
TYPECSCH3	-1,167	0,305
SEX.TYPESCH2	-0,166	0,307
SEX.TYPESCH3	1,165	0,496
SEX.TYPESCH4	-1,505	0,568
AGE.NUMBSCH	-0,067	0,040

TAB. 2.6: Scolarisation avec les variables pour les enfants, la famille, le village, et les écoles, après l'élimination des variables non-significatives. ( $N = 2401$ ,  $M = 1292$ )

Variables significatives	Estimation	Erreur standard
Constante	0,666	0,726
AGE	-0,142	0,058
SEX	-0,070	0,270
RICE	0,577	0,130
CLOTHES	0,524	0,158
SHOPPING	0,576	0,115
FIELD	-0,039	0,169
SEX.FIELD	-0,517	0,247
HAREA	0,010	0,003
FAMSIZE	-0,031	0,021
FATHED	0,178	0,052
MREAD	0,520	0,126
RELIGION2	-0,616	0,649
RELIGION3	-2,344	0,635
SOACT	0,265	0,115
REASON2	0,324	0,150
HOLPRF2	-0,384	0,143
AGE.RELIGION2	0,029	0,052
AGE.RELIGION3	0,124	0,053
VSIZE	0,005	0,002
AGE.VSIZE	-0,000	0,000
NUMBSCH	1,438	0,533
TYPESCH2	0,799	0,322
TYPESCH3	0,610	0,389
TYPESCH4	0,928	0,614
TYPECSCH2	-0,231	0,163
TYPECSCH3	-1,081	0,301
SEX.TYPESCH2	-0,211	0,291
SEX.TYPESCH3	0,928	0,457
SEX.TYPESCH4	-1,355	0,545
AGE.NUMBSCH	-0,065	0,039

TAB. 2.7: Scolarisation avec les variables pour les enfants, la famille, le village, les écoles, et les provinces. ( $N = 2401, M = 1292$ )

Variables significatives	Estimation	Erreur standard
Constante	0,928	0,758
AGE	-0,135	0,059
SEX	-0,117	0,276
RICE	0,556	0,135
CLOTHES	0,559	0,163
SHOPPING	0,532	0,119
FIELD	-0,045	0,173
SEX.FIELD	-0,715	0,253
HAREA	0,013	0,003
FAMSIZE	-0,047	0,022
FATHED	0,195	0,053
MREAD	0,419	0,131
RELIGION2	-0,698	0,659
RELIGION3	-2,026	0,648
SOACT	0,184	0,119
REASON2	0,365	0,160
HOLPRF2	-0,457	0,148
AGE.RELIGION2	0,035	0,053
AGE.RELIGION3	0,112	0,054
VSIZE	0,004	0,002
AGE.VSIZE	-0,000	0,000
NUMBSCH	1,475	0,545
TYPESCH2	0,937	0,330
TYPESCH3	0,647	0,400
TYPESCH4	1,216	0,631
TYPECSCH2	-0,008	0,172
TYPECSCH3	-1,168	0,322
SEX.TYPESCH2	-0,158	0,298
SEX.TYPESCH3	1,053	0,463
SEX.TYPESCH4	-1,365	0,559
AGE.NUMBSCH	-0,070	0,039
Fianarantsoa	-0,699	0,186
Mahajanga	0,327	0,229
Antisranana	-0,938	0,225
Toamasina	-0,073	0,223
Toliara	0,807	0,315

## 2.3 Conclusions

A partir de notre modèle final, nous pouvons voir que les groupes principaux de variables qui influencent la scolarisation sont les caractéristiques des enfants et de leurs familles. On peut noter spécialement les variables pour le travail des enfants. Les enfants des familles éduquées et riches ont aussi plus de chance d'être scolarisés. Les variables principales qui diminuent la probabilité sont la religion et la préférence pour le moment des vacances, en plus du sexe et l'âge. Les plus âgés et les filles ont moins de probabilité d'être scolarisés que les jeunes et les garçons.

Aucune variable du village des enfants n'est significative dans le modèle final. Par contre, la probabilité augmente avec la disponibilité des écoles dans le village.

Enfin, nous notons que le fait d'exclure les enfants qui sont à l'école secondaire ne change pas les variables dans le modèle. Ce résultat est raisonnable puisque très peu d'enfants fréquentent l'école secondaire dans notre échantillon. Il serait intéressant d'étudier les facteurs qui influencent cette scolarisation, mais nous avons trop peu d'enfants, surtout à cause des valeurs manquantes.





## Chapitre 3

# Admissions à l'école

Dans ce chapitre, nous allons analyser les facteurs qui influencent l'admission des enfants à l'école pour la première fois.

Les admissions globales (enseignement public et privé) qui avaient augmenté depuis le début des années 80, ont chuté brutalement en 1989/90 pour ne redémarrer véritablement qu'en 1993. Cette évolution des admissions globales, entre 1987 et 1992, pourrait s'expliquer par les conséquences de la crise économique sur la demande d'éducation, notamment en raison du coût direct et indirect de la scolarisation. Elle pourrait aussi résulter des événements politiques de 1991 qui auraient découragé certains parents d'envoyer leurs enfants à l'école pendant les grèves, puis de leur faire reprendre le chemin de l'école, par la suite.

Depuis 1987, les nouvelles admissions dans l'enseignement public baissent régulièrement, mais cette tendance a été contre-balançée par une élévation des inscriptions dans les écoles privées. Le taux de participation du secteur privé, dans les admissions, est ainsi passé de 17% en 1987 à 25% en 1992. Cette situation résulte probablement en premier de la faible qualité de l'enseignement dans les écoles publiques et des événements de 1991 qui ont poussé des parents à transférer leurs enfants dans le privé pour des raisons de grève et de sécurité.

### 3.1 Analyse par cohorte

Les indicateurs d'admission que l'on peut normalement calculer à partir des statistiques globales dérivées des fiches de recensement annuel des écoles et des estimations de population (extrapolation à partir des données du recensement de population de 1978) concernent le taux brut d'admission (TBA : admissions nouvelles divisées par le nombre d'enfants âgés de six ans) et le taux net d'admission (TNA : admissions d'enfants âgés de six ans divisées par le nombre d'enfants âgés de six ans). Si le TBA donne une idée de l'ampleur des admissions annuelles par rapport à l'effectif des enfants en âge d'être admis et de la capacité d'accueil du système, il ne donne aucune information, en revanche, sur le sort d'une cohorte en matière d'admission. Pour sa part, le TNA indique seulement la proportion des enfants d'une cohorte qui entrent à l'école primaire à l'âge normal d'admission.

TAB. 3.1: TBA officiel, avec le nombre d'admissions et d'enfants d'âge six et l'estimation du TBA, avec l'intervalle de confiance à 95% (IC).

	1990	1991	1992	1993
TBA (Ministère)	104,7	103,5	100,5	100,5
Admissions	346	441	449	520
Enfants âge 6	430	371	432	404
TBA (Etude)	80,5	118,9	103,9	128,7
IC	(69,6, 93,0)	(103,2, 136,9)	(90,8, 118,9)	(112,7, 147,0)

TAB. 3.2: Admissions à l'école (%) par cohorte et âge.

Cohorte	Age										
	6	7	8	9	10	11	12	13	14	15	16
1977					7,6	11,0	9,8	5,3	3,0	2,3	0,4
1978				10,2	10,6	14,7	6,5	5,1	0,3	0,3	
1979			6,5	10,0	21,1	13,5	6,8	2,7	2,7		
1980		5,9	12,4	22,8	14,8	14,2	7,4	1,5			
1981	2,5	4,3	16,9	17,8	19,2	9,8	7,1				
1982	4,2	14,1	20,1	25,1	13,0	8,2					
1983	7,7	14,9	22,3	22,1	14,4						
1984	10,2	20,5	26,1	20,2							
1985	11,6	29,2	28,9								
1986	20,0	44,8									
1987	30,1										

A Madagascar, le TBA a eu une tendance à la baisse, puis à la stabilisation, avec un maintien au-dessus de 100%. Les estimations du TBA à partir des données de l'enquête et dérivées des statistiques de recensement, qui concerne l'ensemble du pays, y compris le milieu urbain, se trouvent dans le Tableau 3.1. Les deux séries de résultats sont donc parfaitement compatibles, étant donné les erreurs d'échantillonnage dans le numérateur et dans le dénominateur.

A partir des données de l'enquête, il est possible, en revanche, d'estimer des indicateurs d'admission beaucoup plus précis et pertinents :

- le taux d'admission par âge pour une cohorte déterminée (admission d'enfants de la cohorte à un âge donné divisées par l'effectif de la cohorte) ;
- le taux cumulé d'admission au plus tard à un âge donné, par cohorte ;
- le taux d'admission de cohortes (obtenus en additionnant les taux par âge d'une cohorte, pour tous les âges possibles d'admission).

A défaut de collecter de données précises et fiables sur la distribution par âge des nouveaux admis, dans le cadre du recensement annuel des écoles, l'enquête constitue donc le seul instrument permettant de suivre le sort d'une génération en matière d'admission.

Les taux d'admission par âge pour chaque cohorte dans l'enquête sont présentés dans le Tableau 3.2 et ceux cumulés pour donner l'admission au plus tard à un âge donné, par cohorte, dans le Tableau 3.3. Les estimations figurant dans ces tableaux montrent que, au cours des dernières années, les enfants issus des cohortes successives ont eu tendance à entrer à un âge de plus en plus jeune à l'école primaire. Cette tendance est même frappante par sa régularité et par l'ampleur des progrès enregistrés. Elle éclaire d'un jour nouveau

TAB. 3.3: Admissions cumulées à l'école (%) par cohorte et âge.

Cohorte	Age										
	6	7	8	9	10	11	12	13	14	15	16
1977					7,6	18,6	28,4	33,7	36,7	39,0	39,4
1978				10,2	20,8	35,5	42,0	47,1	47,4	47,7	
1979			6,5	16,5	37,6	51,1	57,9	60,6	63,3		
1980		5,9	18,3	41,1	55,9	70,1	77,5	79,0			
1981	2,5	6,8	23,7	41,5	60,7	70,5	77,6				
1982	4,2	18,3	38,4	63,5	76,5	84,7					
1983	7,7	22,6	44,9	67,0	81,4						
1984	10,2	30,7	56,8	77,0							
1985	11,6	40,8	69,7								
1986	20,0	64,8									
1987	30,1										

l'évolution du TBA constatée au cours des dernières années : dans une large mesure, cette évolution résulterait d'un recul des admissions tardives. Le déficit d'admissions correspondant à des enfants âgés serait en voie de disparition. En d'autres termes, la baisse du TBA ne résulterait pas d'une baisse de la demande, mais ce point reste à confirmer en suivant les admission futures des cohortes qui viennent seulement d'accéder à l'école.

Nous voyons, par exemple, que la cohorte de 1982 a un taux d'admission global de 85% à onze ans, malgré le fait qu'il n'y avait que 18% à sept ans. Par contre, la cohorte de 1986 a déjà un taux de 65% à sept ans. Ainsi, on peut prévoir qu'au moins 90% des enfants de ces cohortes récentes ont été ou seront admis à l'école, à un moment ou à un autre. En d'autres termes, on serait proche de la généralisation des admissions, en milieu rural, et donc probablement aussi dans l'ensemble du pays.

Le niveau de TBA, supérieur à 100%, s'expliquerait alors par le rattrapage d'admissions à un âge avancé, étant donné que le taux d'admission par cohorte est de l'ordre de 90% ou un peu supérieur. Etant donné que, dans les analyses des données de cette enquête, on a enregistré pour chaque enfant une admission et une seule (celle de l'année scolaire où les parents ont déclaré qu'il avait commencé sa scolarité), il est évidemment impossible que les données admission incluent des réinscriptions après abandon, alors que cette situation pourrait très bien se produire à partir des statistiques de recensement, compte tenu de la manière dont elles sont collectées. On ne peut pas non plus trouver de biais résultant de l'utilisation de données démographiques estimées, comme pour le calcul des TBA et TNA de recensement, car les données de scolarisation et de population sont issues de la même base d'information : les ménages. Le seul risque de biais, dans cette enquête, concernerait une surdéclaration de scolarisation des enfants par le ménage (ou une sousdéclaration des enfants non scolarisés).

Nous utiliserons les taux d'admission par âge dans le Tableau 3.2 dans le Chapitre 4 pour étudier le retard d'inscription à l'école.

TAB. 3.4: Admissions à l'école (%), pour ceux qui ne la fréquentaient pas, par cohorte et âge.

Cohorte	Age									
	7	8	9	10	11	12	13	14	15	16
1977				10,0	15,8	21,3	15,1	9,0	6,3	0,9
1978			12,1	14,8	28,7	17,9	15,3	1,2	1,0	
1979		7,1	11,9	31,5	29,9	21,0	10,1	8,8		
1980	6,1	13,7	30,4	28,9	36,9	31,3	7,9			
1981	1,1	18,9	24,6	34,9	27,2	23,8				
1982	21,4	25,4	41,6	37,4	33,7					
1983	16,4	29,4	41,3	43,1						
1984	23,3	39,1	49,3							
1985	35,6	53,6								
1986	62,0									

TAB. 3.5: Admission à l'école avec les variables pour les enfants. ( $N = 1009$ ,  $M = 403$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	3,352	0,374	BIOLOGIC
AGE	-0,438	0,038	RICE
SEX	-0,062	0,157	BROTHER
WATER	1,264	0,171	CLOTHES
MEALS	-2,049	0,917	ANIMALS
SHOPPING	0,641	0,155	AFOOD
FIELD	-0,415	0,212	
AGE.MEALS	0,141	0,085	

## 3.2 Admissions en 1993

Pour calculer des probabilités de première admission utilisables dans un modèle, nous devons prendre, à un point donné dans le temps, seulement les enfants qui n'ont jamais été scolarisés auparavant. Les données, par cohorte, se trouvent dans le Tableau 3.4. Ici, nous n'étudions que les admissions en 1993 ; celles-ci correspondent à la diagonale inférieure du tableau.

Puisque la variable 'si l'enfant entre à l'école ou pas' est binaire, nous allons ajuster des modèles logistiques, de l'Equation (2.1) où  $\pi$  est ici la probabilité d'admission.

### 3.2.1 Enfants

Nous commençons avec 1412 observations (tous les enfants qui avait plus de six ans et qui n'étaient pas à l'école avant cette date) et un modèle nul avec une déviance de 1858,3. Selon notre règle pour l'élimination des variables non-significatives, six des 12 variables dans ce groupe sont retenues. La déviance du nouveau modèle est 1037,5 avec 1001 degrés de liberté pour 1009 observations. Ici, 403 observations sont éliminées à cause des valeurs manquantes dans les variables explicatives. Les résultats sont dans le Tableau 3.5.

TAB. 3.6: Admission à l'école avec les variables pour les enfants et la famille. ( $N = 872$ ,  $M = 540$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	3,525	0,445	HAREA
AGE	-0,469	0,044	HWALL
SEX	-0,167	0,187	SPMEAT
WATER	1,104	0,193	ACTIV
MEALS	-1,267	1,040	FAMSIZE
SHOPPING	0,469	0,176	NUMBCH
FIELD	-0,424	0,238	FATHED
AGE.MEALS	0,066	0,098	FREAD
SOACT	0,382	0,174	DISTMIN
MREAD	0,876	0,181	REASON
DISTWAT	0,153	0,101	LANGSC
RELOTHER	-0,826	0,183	HOLPRF
SEX.DISTWAT	0,298	0,144	

La probabilité de s'inscrire à l'école diminue avec l'âge, comme on pourrait s'y attendre. Elle est plus grande pour ceux qui portent de l'eau et ceux qui font les courses, mais assez bizarrement, plus basse pour les enfants qui préparent les repas (voir le tableau dans la Section B.2). Elle est aussi plus petite pour les enfants qui travaillent aux champs. Il est intéressant de constater qu'il n'y a pas de différence significative entre les deux sexes, mais nous gardons cette variable dans le modèle pour le moment. Apparemment, il n'y a pas de discrimination à l'encontre des filles.

Le fait pour l'enfant d'aller chercher de l'eau constitue le facteur dont l'effet positif sur la probabilité d'admission est le plus marqué et le plus significatif. Peut-être les parents profitent-ils du trajet (bi)quotidien de l'enfant scolarisé vers l'école pour lui demander de rapporter de l'eau. Cela apparaîtrait assez logique, car les enfants non encore scolarisés sont peut-être trop jeunes pour le faire, tandis que presque tous les enfants ayant dépassé l'âge de huit à neuf ans sont soit scolarisés, soit engagés dans des activités productives. Cette variable est peut-être plus importante parce que chercher de l'eau est une tâche pour les enfants des deux sexes, par opposition à la préparation des repas et les courses. En contraste à chercher de l'eau et faire les courses, l'effet de préparer le repas est négatif, mais il diminue avec l'âge.

### 3.2.2 Famille

Nous gardons ces variables pour les enfants dans le modèle et nous ajoutons les 16 variables explicatives qui se réfèrent aux caractéristiques de la famille. Après exclusion, une à une, des variables non-significatives, il nous en reste quatre. La nouvelle déviance est 829.0 avec 859 degrés de liberté pour 872 observations. Nous perdons 137 observations avec les valeurs manquantes, pour un total de 540. Les résultats se trouvent dans le Tableau 3.6.

La probabilité d'admission est plus élevée si le père participe aux activités sociales du

TAB. 3.7: Admission à l'école avec les variables pour les enfants, la famille, et le village. ( $N = 834, M = 578$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	2,799	0,634	AGRV
AGE	-0,411	0,064	HERDV
SEX	-0,196	0,192	FISHV
WATER	1,008	0,199	MARKET
MEALS	-1,538	1,102	SHOP
SHOPPING	0,445	0,182	DISTF
FIELD	-0,312	0,244	CATHV
AGE.MEALS	0,086	0,104	PROTV
SOACT	0,449	0,181	
MREAD	0,786	0,186	
DISTWAT	0,158	0,101	
RELOTHER	-0,909	0,188	
SEX.DISTWAT	0,312	0,146	
VSIZE	0,006	0,003	
AGE.VSIZE	-0,000	0,000	

village et si la mère peut lire. D'une façon inattendue, elle augmente aussi avec la distance de la source d'eau (voir le tableau dans la Section B.3), surtout pour les filles. Elle est plus petite pour les religions autre que catholique et protestante.

L'aptitude à lire de la mère a donc une influence positive sur la probabilité d'admission. En revanche, le nombre d'années d'étude ne constitue pas une variable explicative significative : ce sont les aptitudes réelles et non pas la scolarisation formelle qui comptent vraiment. Cela tend à confirmer que quelques années de scolarisation ne se traduisent pas nécessairement par l'acquisition des apprentissages de base (ou que ceux-ci tendent à disparaître avec le temps) ; cela pourrait signifier aussi qu'il existe une forte variance des apprentissages entre adultes ayant suivi le même nombre d'années de scolarisation.

La participation du père à des activités sociales (association de parents d'élèves, association religieuse, etc.) a un effet positif sur la probabilité d'admission. Cet indicateur traduit le degré d'ouverture du père à l'égard de la communauté et sa motivation concernant les activités sociales, culturelles, et éducatives de son environnement.

### 3.2.3 Village

Après l'introduction des variables pour le village, et l'élimination de celles qui ne sont pas significatives, il ne nous reste qu'une variable, la taille du village. Le nouveau modèle, qui contient des variables pour les caractéristiques de l'enfant et de sa famille en plus de celle pour les différences entre les villages, a une déviance de 784,8 avec 819 degrés de liberté pour 834 observations. A ce point, 38 observations supplémentaires sont éliminées, pour un total de 578. Les résultats sont dans le Tableau 3.7.

La probabilité de commencer à l'école augmente avec la taille du village, surtout pour

TAB. 3.8: Admission à l'école avec les variables pour les enfants, la famille, le village, et les écoles. ( $N = 834, M = 578$ )

Variables significatives	Estimation	Erreur standard
Constante	1,363	0,690
AGE	-0,408	0,066
SEX	-0,269	0,199
WATER	1,131	0,206
MEALS	-0,296	0,254
SHOPPING	0,258	0,190
FIELD	-0,296	0,254
AGE.MEALS	0,053	0,109
SOACT	0,376	0,187
MREAD	0,837	0,194
DISTWAT	0,124	0,101
RELOTHER	-0,653	0,199
SEX.DISTWAT	0,318	0,147
VSIZE	0,007	0,003
AGE.VSIZE	-0,001	0,000
NUMBSCH	0,473	0,278
OPSCH	1,206	0,399

les plus jeunes enfants.

### 3.2.4 Ecoles

Maintenant, nous ajoutons les variables pour les écoles dans le village. Les variables pour les types d'écoles ne sont pas retenues. La déviance du modèle est 736,3 avec 817 degrés de liberté pour 834 observations. Les deux variables, école ouverte et nombre d'écoles, n'ont pas de valeurs manquantes; le total reste 578. Les résultats sont présentés dans le Tableau 3.8.

La probabilité d'admission augmente avec le nombre d'écoles dans le village et s'il y a une école ouverte.

Comme on pouvait s'y attendre, l'offre d'éducation a une influence significative et très forte sur la probabilité d'admission. Cette relation est très marquée lorsqu'il existe une école fonctionnelle par rapport au cas où il n'y en a pas (ou une école fermée). Mais, le nombre d'écoles joue aussi : il ne suffit pas de la présence d'une école pour absorber toute la demande d'admission du village.

### 3.2.5 Elimination des variables

Ici, nous éliminons les variables qui sont devenues non-significatives, d'abord celles qui sont entrées en dernier lieu. Ainsi, nous retirons si l'enfant aide à faire les courses et à travailler aux champs et l'interaction entre l'âge et préparer les repas.



TAB. 3.9: Admission à l'école avec les variables pour les enfants, la famille, le village, et les écoles, après l'élimination des variables non-significatives. ( $N = 856, M = 556$ )

Variables significatives	Estimation	Erreur standard
Constante	1,377	0,655
AGE	-0,413	0,062
SEX	-0,249	0,195
WATER	1,166	0,199
MEALS	-0,470	0,283
SOCACT	0,391	0,184
MREAD	0,824	0,191
DISTWAT	0,141	0,098
RELOTHER	-0,658	0,196
SEX.DISTWAT	0,357	0,141
VSIZE	0,007	0,003
AGE.VSIZE	-0,001	0,000
NUMBSCH	0,412	0,271
OPSCH	1,380	0,389

Après cette procédure, le nombre de valeurs manquantes est 556 ; il nous reste 856 observations. La déviance est 751,7 avec 842 degrés de liberté. Les résultats sont dans le Tableau 3.9.

### 3.2.6 Différences régionales

Après l'introduction de la variable pour les provinces, nous découvrons des différences significatives entre les régions. La déviance est 727,1 avec 837 degrés de liberté pour 856 observations. Les résultats sont dans le Tableau 3.10.

Dans deux des provinces, Fianarantsoa et Antisranana, nous découvrons que, par rapport à Antananarivo, la probabilité d'admission est plus basse. Par contre, pour Mahajanga et Toliara, elle est plus élevée. Dans la dernière province, Toamasina, la probabilité est la même que dans la région d'Antananarivo.

Comme pour la scolarisation, ces résultats sont en opposition avec les données brutes dans la Section B.6. Là, les taux d'admissions à Fianarantsoa, Antisranana, et Toamasina sont à peu près la moitié de ceux des trois autres provinces. Ainsi, les variables dans le modèle peuvent expliquer le faible taux d'admission à Toamasina mais pas dans ces deux autres provinces. Le modèle indique aussi que la province d'Antananarivo a un taux d'admission trop élevé, par rapport aux valeurs des variables utilisées.

## 3.3 Conclusions

A partir de notre modèle final, nous pouvons voir que les groupes principaux de variables qui influencent l'admission à l'école sont les caractéristiques des enfants eux-mêmes

TAB. 3.10: Admission à l'école avec les variables pour les enfants, la famille, le village, les écoles, et les provinces. ( $N = 856, M = 556$ )

Variables significatives	Estimation	Erreur standard
Constante	1,483	0,710
WATER	1,303	0,215
MEALS	-0,524	0,289
SOACT	0,379	0,192
MREAD	0,732	0,200
DISTWAT	0,091	0,101
RELOTHER	-0,443	0,216
SEX.DISTWAT	0,368	0,142
VSIZE	0,007	0,003
AGE.VSIZE	-0,001	0,000
NUMBSCH	0,509	0,294
OPSCH	1,402	0,403
Fianarantsoa	-0,606	0,308
Mahajanga	0,419	0,394
Antisranana	-0,914	0,368
Toamasina	0,111	0,353
Toliara	0,580	0,558

et de leur famille. Plusieurs des variables pour le travail domestiques sont significatives. Les enfants des familles plus actives et plus éduquées ont plus de chance d'entrer à l'école. En fait, ce sont les variables qui changent la probabilité de commencer (à part l'âge — les plus âgés ont moins de chance d'admission).

Aucune variable pour le village ne reste dans le modèle final. Par contre, la probabilité d'admission augmente considérablement avec la disponibilité d'écoles dans le village.



# Chapitre 4

## Retard d'admission à l'école

Dans ce chapitre, nous analysons les explications pour le retard d'admission des enfants à l'école. A Madagascar, un enfant devrait commencer à l'école à six ans. Pourtant, ce n'est pas toujours le cas ; dans la période de 1989 à 1993, certains enfants étaient plus de cinq ans en retard et le retard moyen était de 2,54 ans. Nous ne prenons en compte que les années à partir de 1989 parce que le début de la carrière des enfants plus âgés n'est pas disponible.

### 4.1 Analyse par cohorte

Les retards d'admission pour les différentes cohortes sont dans le Tableau 3.2 (il suffit de soustraire six de l'âge d'admission). On peut aussi représenter les données de ce tableau par année au lieu de cohorte, comme dans le Tableau 4.1. Ce tableau donne la distribution des âges des enfants qui commencent à l'école pour la première fois chaque année. Mais, les pourcentages pour les plus petits retards dans les premières années sont surestimés à cause des valeurs qui manquent dans le coin supérieur du tableau.

Les raisons pour lesquelles les enfants ont du retard ont peut-être changé dans le temps avec le développement du pays depuis quelques années. Ainsi, nous avons décidé

TAB. 4.1: Retard d'admission à l'école (% et nombre) par année.

Année	Retard										
	0	1	2	3	4	5	6	7	8	9	10
1989	12,3	16,2	24,0	25,3	25,3	14,0	8,1				
	38	50	74	78	78	43	25				
1990	12,8	15,8	17,8	20,0	12,5	13,0	4,8	3,5			
	51	63	71	79	50	52	19	14			
1991	14,6	14,8	18,7	17,1	16,1	9,3	4,9	2,9	1,6		
	75	76	96	88	83	48	25	15	8		
1992	17,9	23,2	17,9	17,3	8,3	7,9	4,6	1,7	0,2	0,9	
	97	126	97	94	45	43	25	9	1	5	
1993	16,0	30,0	20,6	11,7	9,7	4,8	4,3	0,8	1,7	0,2	0,2
	97	182	125	71	59	29	26	5	10	1	1

de construire, d'abord, un modèle plus général pour tous les enfants qui ont commencé à l'école entre 1989 et 1993, basé sur les données de ce tableau. Ensuite, nous regardons le cas spécifique des enfants qui ont commencé en 1993 (la dernière ligne du tableau). Le retard moyen a diminué depuis quelques années, avec un retard moyen de 2,17 ans en 1993.

L'enfant peut avoir un retard de zero ans (s'il commence à six ans) jusqu'à dix ans. Nous allons ajuster un modèle linéaire généralisé avec une distribution de Poisson. Dans ce modèle, le logarithme du retard moyen a une relation linéaire avec les variables explicatives :

$$\log(\mu) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_K x_K$$

où  $\mu$  est le retard moyen d'entrée à l'école,  $\beta_k$   $k = 1, \dots, K$  sont les coefficients de régression, et  $x_k$  sont les variables explicatives.

Les variables explicatives, et la méthode d'introduction dans le modèle, sont les mêmes que dans les chapitres précédents. La méthode de sélection des variables reste inchangée. Par contre, dans ce cas-ci, les enfants qui avait six ans et ont commencé à l'école sont retenus.

## 4.2 Retard, 1989–1993

Dans cette section, nous étudions le retard observé pour tous les enfants qui ont commencé à l'école entre 1989 et 1993. Parmi les données se trouvent des enfants qui, apparemment, ont commencé avant d'avoir six ans. Ceci pourrait être le résultat d'erreurs d'encodage. Ou les parents ont emmené leurs enfants à l'école en avance. Il y a 65 enfants qui ont commencé à cinq ans ; nous les incluons dans le modèle comme si ils n'ont pas de retard. Par contre, ceux qui avaient deux, trois, ou quatre ans, selon les données (en tout, 20 enfants), sont exclus ; s'ils ne sont pas des erreurs d'encodage, ils sont probablement des enfants d'institutrices.

Pour prendre en compte la lente évolution dans le retard à travers les années, nous introduisons dans notre modèle une tendance linéaire dans le temps.

### 4.2.1 Tendance linéaire dans le temps

Nous nous intéressons d'abord à la façon dont le retard moyen change pendant la période d'observation. Nous créons une nouvelle variable (YEAR) avec les valeurs de un à cinq pour les enfants qui ont commencé de 1989 à 1993.

Si nous regardons le Tableau 4.2 pour le retard moyen pendant les années de l'étude, nous trouvons une diminution graduelle de 1993 à 1989. (Rappelons, du Tableau 4.1, que les premières années sont sous-estimées.) Pour cette raison, nous avons inclus la variable pour l'année dans le modèle pour prendre en compte cette tendance.

Quand nous ne prenons que cette variable explicative (Tableau 4.3), nous trouvons qu'elle est très significatives. La déviance pour ce modèle est 3913,1 avec 2445 degrés de

TAB. 4.2: Retard moyen chaque année de l'étude (1989–1993).

Année	1993	1992	1991	1990	1989
Retard moyen	2,17	2,33	2,77	2,79	2,87

TAB. 4.3: Retard d'admission avec seulement la tendance dans le temps comme variable explicative. ( $N = 2447$ )

Variables	Estimation	Erreur standard
Constante	1,096	0,023
YEAR	-0,075	0,009

liberté pour 2447 observations, par rapport à une déviance de 3982.4 avec 2446 degrés de liberté pour le modèle nul.

Par la suite, nous suivrons la procédure habituelle pour l'addition des variables pour l'enfant, sa famille, le village, l'école, et la région.

#### 4.2.2 Enfants

La plupart des variables pour les enfants sont significatives ici. Des 15 variables disponibles, dix sont retenues. Malheureusement, il existe un problème majeur avec des réponses non-aléatoires. Ainsi, nous ne pouvons retenir que trois variables en fin de compte. La déviance est 3457,5 avec 2226 degrés de liberté pour 2232 observations. Les résultats sont présentés dans le Tableau 4.4.

A première vue, les filles sont un peu moins en retard que les garçons, étant donné les autres variables dans le modèle. L'estimation du rapport du retard moyen des filles par rapport à celui des garçons est

$$\begin{aligned}\hat{\mu} &= \frac{\hat{\mu}_f}{\hat{\mu}_g} \\ &= \exp(-0.133)\end{aligned}$$

TAB. 4.4: Retard d'admission avec les variables pour les enfants. ( $N = 2232$ ,  $M = 215$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,156	0,031	BIOLOGIC
YEAR	-0,088	0,011	RICE <sup>+</sup>
SEX	-0,162	0,032	SHOPPING <sup>+</sup>
CLOTHES	-0,093	0,070	WATER <sup>+</sup>
SEX.CLOTHES	0,161	0,068	MEALS <sup>+</sup>
YEAR.CLOTHES	0,116	0,021	BROTHER*
			FIELD <sup>+</sup>
			AFOOD*
			ANIMALS <sup>+</sup>

où  $\mu_f$  et  $\mu_g$  sont les retards moyens pour les filles et pour les garçons,  $\mu$  est le rapport des deux, et le ‘chapeau’ indique qu’elles sont des valeurs estimées dans notre modèle. Ainsi, le retard des filles est, en moyenne, 0,88 fois celui des garçons. On peut faire ce type de calcul de la même façon pour toutes les variables dans notre modèle.

Il n’y a pas de différence significative de retard entre les enfants qui sont les membres biologiques de la famille et les autres.

Parmi les tâches ayant une relation significative avec le retard, nous trouvons : chercher de l’eau, piler le riz, faire la lessive, faire les courses, préparer les repas, s’occuper des animaux, et travailler dans les champs. Toutes, sauf faire les courses, ont un rapport positif avec le retard. Ceci veut dire que les enfants qui font ces tâches commencent plus tard. Malheureusement, toutes ces variables, sauf faire la lessive, ont des valeurs manquantes non-aléatoires. Pour cette raison, nous ne les gardons pas dans le modèle. Pour les tâches domestiques, les valeurs manquantes sont surtout pour les enfants avec les plus courts retards. Par contre, pour les tâches agricoles, elles sont pour les retards grands. Deux autres variables, s’occuper des frères et soeurs et chercher de la nourriture pour les animaux, sont non-significatives.

Faire la lessive a une influence négative sur le retard pour les filles ; son influence a augmenté à travers les années.

### 4.2.3 Famille

L’introduction de ce groupe de variables ne change en rien les conclusions de la section précédente. Toutes les variables restent significatives et les coefficients ne changent pas de signes. Seulement cinq des 18 variables dans ce groupe sont significatives. La déviance diminue à 2568,4 avec 1717 degrés de liberté pour 1731 observations. A cause des cinq variables ajoutées, 501 observations ont dues être éliminées. Les résultats se trouvent dans le Tableau 4.5.

Le surface de la maison montre un rapport négatif avec le retard ; le retard moyen diminue quand la maison est plus grande. Ceci est logique parce qu’une maison plus grande indique souvent une famille plus riche et ces familles s’occupent souvent plus de l’éducation de leurs enfants. Par contre, pour les familles plus pauvres, les frais de scolarisation peuvent poser des problèmes pour la famille, ce qui peut expliquer le plus grand retard.

Pour les familles avec beaucoup d’enfants, le retard est plus grand, surtout pour les garçons. Ceci est aussi, peut-être, expliqué par les finances des familles. A la place, elles font peut-être travailler leurs enfants à la maison.

Plus le père est éduqué, plus le retard est court pour l’admission de ses enfants à l’école. On le voit par le signe négatif du coefficient. Un tel père peut s’occuper plus de l’éducation de ses enfants parce qu’il comprend mieux l’importance d’une éducation.

Un résultat important tiré de notre modèle est l’augmentation du retard avec la distance à l’école (mesurée en minutes), et ceci augmente avec les années. Evidemment, les parents se soucient de la sécurité de leurs enfants ; ils ne veulent pas que les jeunes enfants

TAB. 4.5: Retard d'admission avec les variables pour les enfants et pour la famille. ( $N = 1731, M = 716$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,469	0,147	HWALL
YEAR	-0,180	0,052	SPMEAT*
SEX	0,020	0,082	ACTIV
CLOTHES	-0,084	0,082	FAMSIZE
SEX.CLOTHES	0,175	0,081	SOCACT
YEAR.CLOTHES	0,095	0,025	MREAD
HAREA	-0,001	0,001	FREAD
NUMBCH	0,043	0,009	DISTWAT
FATHED	-0,093	0,013	REASON
DISTMIN	0,000	0,001	HOLPRF <sup>+</sup>
LANGSC3	-0,393	0,136	RELIGION*
SEX.NUMBCH	-0,034	0,013	
YEAR.DISTMIN	0,001	0,000	
YEAR.LANGSC3	0,077	0,052	

marchent loin pour aller à l'école. Ainsi, quand l'école est loin, ils retardent l'entrée de leurs enfants à l'école plutôt que de les laisser commencer à six ans.

Pour une raison quelconque, le retard est moindre pour les familles qui souhaiteraient que l'enseignement se fasse en français et en malgache, mais cet effet diminue avec les années.

Beaucoup de variables dans ce groupe sont non-significatives. Certaines sont peut-être liées avec des variables déjà dans le modèle (par exemple, la taille de la famille et le nombre d'enfants; si chaque parent peut lire et le niveau d'éducation du père, etc.). Quand une de ces variables est significative, les autres ont des relations plus faibles avec le retard.

#### 4.2.4 Village

Ici, seulement quatre des dix nouvelles variables sont significatives. Toutes les variables déjà dans le modèle restent significatives, sans changement de signe. Ainsi, les conclusions restent inchangées pour ces variables. La nouvelle déviance est 1928,2 avec 1286 degrés de liberté pour 1305 observations. A la suite de l'introduction de ces variables, 426 observations disparaissent. Presque la moitié des observations sont maintenant manquantes. Les résultats sont dans le Tableau 4.6.

Pour les villages où l'activité principale est l'élevage, où la majorité des habitants sont catholiques, ou ceux qui sont le plus loin de la capitale, les enfants ont plus de retard que dans les autres villages. Par contre, les garçons dans les villages avec un magasin ont un retard plus petit.

Des variables non-significatives sont la taille du village, si l'activité est surtout de l'agriculture et de la pêche, et s'il y a un marché ou pas.



TAB. 4.6: Retard d'admission avec les variables pour les enfants, la famille, et le village. ( $N = 1305, M = 1142$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,260	0,201	AGRV <sup>+</sup>
YEAR	-0,068	0,071	FISHV
SEX	-0,107	0,105	MARKET
CLOTHES	0,141	0,096	VSIZE
SEX.CLOTHES	0,213	0,096	PROTV*
YEAR.CLOTHES	0,088	0,030	
HAREA	-0,001	0,001	
NUMBCH	0,041	0,010	
FATHED	-0,087	0,016	
DISTMIN	0,000	0,001	
LANGSC3	-0,170	0,187	
SEX.NUMBCH	-0,034	0,015	
YEAR.DISTMIN	0,001	0,000	
YEAR.LANGSC3	0,033	0,071	
HERDV	0,083	0,039	
SHOP	-0,162	0,051	
DISTF	0,000	0,000	
CATHV	0,098	0,037	
SEX.SHOP	0,166	0,075	

#### 4.2.5 Ecoles

Aucune variable de ce groupe n'est significative. Ainsi, le modèle ne change pas par rapport à la section précédente.

#### 4.2.6 Différences entre les régions

Pour étudier comment le retard varie entre les provinces, nous introduisons la variable pour les provinces dans le modèle, avec Antananarivo comme point de référence, et retirons l'interaction entre la distance à l'eau et l'année. Les résultats sont dans le Tableau 4.7. La déviance pour ce modèle est 1916,1 avec 1282 degrés de liberté pour 1305 observations.

A partir du tableau, nous voyons que les enfants qui habitent dans la province d'Antananarivo ont un retard plus grand que les autres, après avoir pris en compte les autres variables dans le modèle. Les provinces d'Antisranana et de Toliara ont moins de retard. Pour les trois autres provinces, le retard est similaire.

### 4.3 Retard en 1993

Maintenant, nous étudions seulement les enfants qui ont commencé l'école en 1993, un cas spécial du modèle précédent ; les enfants qui ont commencé plus tôt ne sont pas repris.

TAB. 4.7: Retard d'admission avec les variables pour les cinq groupes de variables. ( $N = 1305$ ,  $M = 1142$ )

Variables significatives	Estimation	Erreur standard
Constante	1,361	0,126
YEAR	-0,103	0,019
SEX	-0,117	0,106
CLOTHES	-0,113	0,097
SEX.CLOTHES	0,194	0,096
YEAR.CLOTHES	0,088	0,030
HAREA	-0,001	0,001
NUMBCH	0,038	0,010
FATHED	-0,091	0,016
DISTMIN	0,000	0,001
LANGSC3	-0,212	0,094
SEX.NUMBCH	-0,032	0,015
YEAR.DISTMIN	0,001	0,000
HERDV	0,062	0,041
SHOP	-0,162	0,053
DISTF	0,000	0,000
CATHV	0,097	0,039
SEX.SHOP	0,155	0,075
Fianarantsoa	-0,054	0,056
Mahajanga	-0,047	0,060
Antisranana	-0,151	0,071
Toamasina	-0,060	0,069
Toliara	-0,243	0,079

TAB. 4.8: Retard d'admission en 1993 avec les variables pour les enfants. ( $N = 528, M = 78$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	0,339	0,134	BROTHER*
SEX	0,054	0,101	
BIOLOGIC	0,159	0,112	
WATER	0,185	0,070	
RICE	0,165	0,068	
CLOTHES	-0,295	0,175	
MEALS	0,282	0,099	
SHOPPING	0,043	0,086	
ANIMALS	0,179	0,078	
FIELD	0,329	0,079	
AFOOD	-0,190	0,170	
SEX.CLOTHES	0,624	0,193	
SEX.SHOPPING	-0,254	0,121	
SEX.AFOOD	-0,618	0,400	

Comme avant, nous éliminons les observations avec des valeurs manquantes pour des variables explicatives dans le modèle. Nous commençons avec 606 enfants et une déviance de 1008,2.

#### 4.3.1 Enfants

Nous introduisons d'abord les caractéristiques des enfants. Presque toutes les variables significatives se rapportent au tâches des enfants chez eux. La déviance de notre modèle est 766,0 avec 514 degrés de liberté pour 528 observations. Les résultats sont dans le Tableau 4.8.

Si l'on compare ce modèle avec celui prenant en compte tous les enfants qui avaient commencé entre 1989 et 1993, nous retrouvons les mêmes variables, à part celles pour la tendance linéaire dans le temps. Une différence importante ici est qu'il n'y a plus d'indication que les valeurs manquantes sont non-aléatoires.

Les tâches que les enfants doivent faire, domestiques (porter l'eau, piler le riz, faire la lessive, préparer le repas, faire les courses) ou agricoles (s'occuper des animaux, travailler aux champs, chercher de la nourriture pour les animaux) sont toutes, sauf trois, des facteurs qui influencent positivement le retard : elles impliquent plus de retard.

Les filles, mais pas les garçons, qui font les courses ont moins de retard. Les filles qui font la lessive ont plus de retard, mais les garçons qui la font en ont moins. Le fait de chercher de la nourriture pour les animaux diminue plus le retard pour les filles que pour les garçons.

Parmi les autres variables, il paraît que le retard est un peu plus grand pour les garçons que pour les filles, mais ce résultat n'est pas significatif. Si l'enfant est un membre

TAB. 4.9: Retard d'admission en 1993 avec les variables pour les enfants et la famille.  
( $N = 492, M = 114$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	0,073	0,195	HWALL*
SEX	0,119	0,205	SPMEAT*
BIOLOGIC	0,163	0,132	ACTIV
WATER	0,196	0,074	FAMSIZE
RICE	0,154	0,070	RELIGION*
CLOTHES	-0,125	0,186	SOCACT
MEALS	0,208	0,105	FATHED <sup>+</sup>
SHOPPING	0,014	0,091	FREAD <sup>+</sup>
ANIMALS	0,202	0,081	DISTWAT
FIELD	0,291	0,085	DISTMIN <sup>+</sup>
AFOOD	-0,219	0,176	LANGSC
SEX.CLOTHES	0,434	0,209	HOLPRF <sup>+</sup>
SEX.SHOPPING	-0,257	0,128	
SEX.AFOOD	-0,624	0,406	
HAREA	-0,007	0,002	
NUMBCH	0,084	0,018	
MREAD	-0,204	0,093	
REASON2	0,203	0,069	
SEX.HAREA	0,005	0,003	
SEX.NUMBCH	-0,063	0,025	
SEX.MREAD	0,233	0,132	

biologique de la famille, il y a plus de chance qu'il commence en retard.

On n'a pas pu introduire la plupart des variables qui se rapportent aux tâches domestiques et agricoles dans le modèle général avec tous les enfants, mais, de toute façon, elles n'étaient pas toutes significatives. Par contre, ici, elles sont toutes significatives. Ceci pourrait nous conduire à penser que les familles ont de plus en plus besoin des enfants à la maison, peut-être parce que le niveau de vie diminue.

Dans ce groupe, il n'y a qu'une seule variable non-significative : si l'enfant s'occupe de ses frères et sœurs. En plus, cette variable a des valeurs manquantes non-aléatoires.

### 4.3.2 Famille

Pour ce groupe, il y a quatre variables dans le modèle, contre deux dans le modèle pour l'ensemble des enfants (la surface de la maison et le nombre d'enfants) ; elles sont toujours significatives et ont le même signe. Les deux autres sont si la mère peut lire et la raison principale pour envoyer les enfants à l'école (une école pas trop loin, les livres gratuits, et les rapports entre le directeur et les enseignants). La déviance est 681,8 avec 471 degrés de liberté pour 492 observations. Les résultats sont dans le Tableau 4.9.

Comme nous avons vu plus haut, la surface de la maison a une relation négative avec

le retard (les enfants dans une famille avec une grande maison commencent plus tôt), mais ici l'effet est surtout pour les garçons. Nous voyons aussi que, plus grand est le nombre d'enfants dans la famille, plus grand est le retard, encore une fois, surtout pour les garçons. Comme on pourrait s'y attendre, le retard diminue lorsque la mère peut lire, mais, de nouveau, seulement pour les garçons.

Plusieurs des raisons que les parents donnent pour envoyer leurs enfants à l'école sont significatives : si l'école n'est pas trop loin (REASON1, pour les numéros, voir le tableau dans la Section B.3), si les livres sont gratuits (REASON3), et si les relations entre le directeur et les enseignants sont bonnes (REASON9) s'opposent à toutes les autres. Nous les comparons simultanément aux autres (celles-ci sont le point de référence). Les enfants des parents qui donnent les premières raisons ont plus de retard (ainsi, si les parents se soucient plus de ces raisons, ils gardent leurs enfants plus longtemps à la maison).

Enfin, il reste quatre autres variables qui sont significatives mais qui ont des valeurs manquantes non-aléatoires et qui sont exclues du modèle. Les deux premières se rapportent au niveau d'éducation du père. Les deux autres sont le temps pour aller à l'école et la préférence pour le moment des vacances scolaires.

### 4.3.3 Village

Quatre autres variables entrent dans le modèle maintenant, mais seulement deux étaient dans celui avec tous les enfants (la distance à la capitale et si la majorité des habitants du village sont catholiques). Elles sont significatives mais le signe de la première a changé. Les deux autres variables qui entrent sont la présence d'un marché dans le village et si la majorité des habitants du village sont protestants. La déviance est 552,3 avec 391 degrés de liberté pour 416 observations. Les résultats sont dans le Tableau 4.10.

Comme on pourrait s'y attendre, la présence d'un marché dans le village influence négativement le retard; la même chose se passe pour la distance à la capitale : le retard diminue quand la distance est plus petite.

Dans les villages où la majorité des habitants sont protestants, les filles ont un plus petit retard que dans les autres villages.

### 4.3.4 Ecoles

Deux des quatre variables pour l'école sont significatives et entrent dans le modèle. Ce sont la présence d'une école privée et d'une école publique dans le village et s'il y a une école fermée dans le village. La déviance est 539,2 avec 381 degrés de liberté pour 408 observations. Les résultats se trouvent dans le Tableau 4.11.

Pour les deux variables, le coefficient est positif. Ce résultat pour le type d'école est surprenant parce que la présence de plusieurs écoles implique un plus grand retard (voir le tableau dans la Section B.5). Par contre, comme on pourrait s'y attendre, il y a un plus grand retard quand il y a une école, publique ou privée, fermée dans le village.

TAB. 4.10: Retard d'admission en 1993 avec les variables pour les enfants, la famille, et le village. ( $N = 416, M = 190$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	0,176	0,234	AGRV*
SEX	0,339	0,231	HERDV*
BIOLOGIC	0,166	0,160	FISHV*
WATER	0,190	0,083	SHOP
RICE	0,200	0,079	VSIZE
CLOTHES	-0,044	0,207	
MEALS	0,247	0,121	
SHOPPING	0,103	0,101	
ANIMALS	0,239	0,091	
FIELD	0,301	0,095	
AFOOD	-0,144	0,207	
SEX.CLOTHES	0,293	0,230	
SEX.SHOPPING	-0,377	0,139	
SEX.AFOOD	-0,377	0,503	
HAREA	-0,007	0,003	
NUMBCH	0,087	0,021	
MREAD	-0,282	0,102	
REASON2	0,239	0,077	
SEX.HAREA	0,005	0,003	
SEX.NUMBCH	-0,077	0,029	
SEX.MREAD	0,271	0,143	
MARKET	-0,302	0,121	
DISTF	-0,001	0,000	
PROTV	-0,030	0,126	
SEX.PROTV	-0,350	0,182	

TAB. 4.11: Retard d'admission en 1993 avec les variables pour les enfants, la famille, le village, et les écoles. ( $N = 408, M = 198$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	0,187	0,236	NUMBSCH
SEX	0,331	0,235	
BIOLOGIC	0,140	0,161	
WATER	0,213	0,085	
RICE	0,190	0,080	
CLOTHES	-0,072	0,207	
MEALS	0,271	0,122	
SHOPPING	0,122	0,102	
ANIMALS	0,238	0,092	
FIELD	0,300	0,097	
AFOOD	-0,113	0,207	
SEX.CLOTHES	0,362	0,232	
SEX.SHOPPING	-0,397	0,142	
SEX.AFOOD	-0,359	0,503	
HAREA	-0,007	0,003	
NUMBCH	0,088	0,021	
MREAD	-0,295	0,104	
REASON2	0,214	0,079	
SEX.HAREA	0,005	0,003	
SEX.NUMBCH	-0,086	0,029	
SEX.MREAD	0,301	0,145	
MARKET	-0,351	0,123	
DISTF	-0,001	0,000	
PROTV	-0,080	0,132	
SEX.PROTV	-0,299	0,187	
TYPESCH2	0,320	0,134	
TYPECSCH2	0,511	0,265	

TAB. 4.12: Retard d'admission en 1993 avec les variables pour les enfants, la famille, le village, et l'école, après l'élimination des variables non-significatives. ( $N = 408$ ,  $M = 198$ )

Variables significatives	Estimation	Erreur standard
Constante	0,324	0,191
SEX	0,329	0,234
WATER	0,211	0,085
RICE	0,199	0,080
CLOTHES	-0,069	0,207
MEALS	0,280	0,122
FIELD	0,293	0,096
SEX.CLOTHES	0,356	0,232
SEX.SHOPPING	-0,402	0,142
HAREA	-0,008	0,003
NUMBCH	0,088	0,021
MREAD	-0,287	0,103
REASON2	0,216	0,079
SEX.HAREA	0,006	0,003
SEX.NUMBCH	-0,087	0,029
SEX.MREAD	0,294	0,145
MARKET	-0,351	0,123
DISTF	-0,001	0,000
PROTV	-0,084	0,132
SEX.PROTV	-0,313	0,187
TYPESCH2	0,327	0,134
TYPECSCH2	0,530	0,264

#### 4.3.5 Elimination des variables

Maintenant, nous éliminons les variables qui sont devenues non-significatives, en commençant par celles qui sont entrées dans le modèle en dernier lieu. Ainsi, nous retirons si l'enfant prépare les repas, va chercher la nourriture pour les animaux ou donne à manger aux animaux, et s'il est un membre biologique de la famille.

Cette procédure laisse le nombre de réponses manquantes à 198, avec 408 observations. La déviance est 541,5 avec 384 degrés de liberté. Les résultats sont dans le Tableau 4.12.

#### 4.3.6 Différences régionales

Quand nous introduisons dans le modèle la variable pour les différences entre provinces, nous trouvons les résultats dans le Tableau 4.13. Pour les deux provinces, Antisranana et Toliara, le retard moyen est plus petit que dans les autres, comme dans le modèle plus général avec plus d'années. La déviance est 524,5 avec 380 degrés de liberté pour 408 observations.



TAB. 4.13: Retard d'admission en 1993 avec les cinq groupes de variables. ( $N = 408$ ,  $M = 198$ )

Variables significatives	Estimation	Erreur standard
Constante	0,339	0,216
SEX	0,255	0,232
WATER	0,263	0,087
RICE	0,248	0,082
CLOTHES	0,206	0,108
MEALS	0,266	0,123
FIELD	0,255	0,097
SEX.SHOPPING	-0,326	0,143
HAREA	-0,008	0,003
NUMBCH	0,085	0,021
MREAD	-0,308	0,106
REASON2	0,275	0,080
SEX.HAREA	0,005	0,003
SEX.NUMBCH	-0,082	0,029
SEX.MREAD	0,337	0,145
MARKET	-0,374	0,125
DISTF	-0,001	0,000
PROTV	-0,067	0,139
SEX.PROTV	-0,290	0,187
TYPESCH2	0,221	0,137
TYPECSCH2	0,429	0,265
Fianarantsoa	0,125	0,118
Mahajanga	-0,023	0,123
Antisranana	-0,385	0,152
Toamasina	0,116	0,148
Toliara	-0,284	0,174

#### 4.4 Changement du retard, 1989–1993

Si nous comparons le modèle pour les années 1989 à 1993 avec celui pour le retard en 1993, nous trouvons quelques modifications dans les variables explicatives. Le groupe de variables le plus important se réfère aux tâches des enfants chez eux, mais, dans le premier cas, il y a beaucoup de réponses manquantes non-aléatoires. Ces variables ont une influence positive sur le retard (elles l'augmentent).

En 1993, les variables pour la famille et le village sont légèrement différentes du modèle général. Ainsi, nous pouvons conclure que la plus grande influence vient des variables des enfants (ou plutôt des décisions des parents à leur égard). Puisqu'il y a plus de tâches domestiques qui augmentent le retard en 1993, on pourrait supposer que l'avis des parents a évolué. Ils ont peut-être plus besoin de l'aide de leurs enfants qu'avant.

Nous pouvons conclure que la baisse récente des admissions nouvelles dans la première année de l'école primaire semble largement expliquée par une régularisation des âges à l'admission, les enfants tendant de plus en plus à entrer à l'école autour de six à sept ans.



# Chapitre 5

## Redoublement

Dans ce chapitre, nous allons analyser les facteurs qui influencent les redoublements. Ceci n'est pas la même chose que le taux d'échec parce qu'il ne prend pas en compte les enfants qui abandonnent après un échec.

### 5.1 Analyse par cohorte

Les taux de redoublements pour chaque cohorte par âge sont présentés dans le Tableau 5.1. Etant donné la variation d'échantillonnage, ces valeurs restent assez constantes à travers les cohortes. Elles sont d'environ 30% jusqu'à neuf ou dix ans, descendant un peu pour les plus âgés. Cette baisse peut s'expliquer en partie par l'abandon des enfants qui ont déjà doublé plusieurs fois. Bien entendu, certains des enfants plus âgés sont à l'école secondaire où le taux d'échec paraît moins élevé. Nous pouvons ainsi conclure que le taux de redoublement n'a pas vraiment changé à travers les années.

TAB. 5.1: Doublants (%) par cohorte et l'âge quand ils ont eu l'échec.

Cohorte	Age								
	7	8	9	10	11	12	13	14	15
1977			33,3	24,0	24,0	17,7	18,2	19,8	24,0
1978		26,7	31,1	20,2	17,2	27,0	19,3	13,6	
1979	11,1	29,4	29,5	26,5	21,8	21,5	24,7		
1980	36,7	27,8	22,4	20,8	24,0	28,3			
1981	27,0	30,4	26,8	21,7	25,6				
1982	27,1	29,7	20,7	32,5					
1983	30,6	28,6	31,5						
1984	30,3	31,8							
1985	34,7								

TAB. 5.2: Doublants avec les variables pour les enfants. ( $N = 1853, M = 262$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	0,829	0,401	BIOLOGIC
AGE	-0,225	0,037	MEALS
SEX	-1,462	0,539	BROTHER
WATER	0,306	0,128	CLOTHES
RICE	0,278	0,112	ANIMALS
SHOPPING	0,169	0,106	AFOOD
CLASS	0,197	0,045	FIELD
SEX.AGE	0,110	0,045	

## 5.2 Doublants, 1992–1993

Nous étudierons les enfants qui étaient dans la même classe à l'école en 1992 et 1993. Ceci correspond à la diagonale inférieure du Tableau 5.1. Des 2115 enfants pour qui nous avons des données pour ces deux années, quelques uns (57) étaient déjà dans la même classe l'année précédente.

### 5.2.1 Enfants

Nous commençons avec 2115 enfants : ceux qui avaient au moins sept ans en 1992 et qui étaient à l'école en 1992 et en 1993, pour lesquels nous avons des informations. La déviance nulle est 2524,6. Six variables sont significatives, avec une déviance de 2158,2 avec 1845 degrés de libertés pour 1853 observations. Il y a 262 valeurs manquantes. Les résultats sont dans le Tableaux 5.2.

La probabilité de doubler est inférieure pour les filles et diminue avec l'âge, mais moins vite pour les filles. Par contre, étant donné les autres variables dans le modèle, elle augmente avec la classe à l'école.

Un enfant a plus de chance de doubler s'il doit chercher de l'eau, piler le riz, ou faire les courses. Ceci est en contraste avec l'effet de ces variables sur les autres variables réponses où elles ont une influence favorable. C'est peut-être dans les familles les plus défavorisées, mais motivées pour envoyer les enfants à l'école que les enfants doivent accomplir ces tâches. Notons qu'aucune des tâches agricoles n'entre dans le modèle.

### 5.2.2 Famille

Après vérification des variables pour la famille, nous découvrons qu'un grand nombre (7) sont significatives. (Même si les erreurs standards pour le choix de langue sont grandes, la déviance augmente de 5,4 si l'on l'élimine.) Mais 338 observations sont perdues à cause des valeurs manquantes. La nouvelle déviance est 1757,6 avec 1499 degrés de libertés pour 1515 observations. Les résultats sont dans le Tableau 5.3. Parmi les variables déjà dans le modèle, faire les courses n'est pas significatif.

TAB. 5.3: Doublants avec les variables pour les enfants et la famille. ( $N = 1515$ ,  $M = 600$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,168	0,626	RELIGION
AGE	-0,289	0,042	REASON
SEX	-1,465	0,595	DISTMIN
WATER	0,334	0,143	DISTWAT
RICE	0,278	0,126	ACTSOC
SHOPPING	0,090	0,117	SPMEAT
CLASS	0,305	0,056	NUMBCH
SEX.AGE	0,102	0,050	HOLPRF
HAREA	-0,005	0,003	HWALL
ACTIV	-0,298	0,154	
FAMSIZE	0,073	0,024	
FATHED	-0,239	0,063	
MREAD	-0,295	0,138	
FREAD	0,451	0,192	
LANGSC2	0,004	0,391	
LANGSC3	0,369	0,357	

La probabilité de doubler diminue avec la richesse de la famille, mesurée par la taille de la maison, et quand le père est actif dans le village. Elle diminue aussi avec la taille de la famille.

Sans surprise, toutes les variables concernant l'éducation entrent dans le modèle. La probabilité de doubler est plus faible dans les familles où le père a plus d'éducation et si la mère peut lire. Mais, étant donné les autres variables, elle est plus élevée si le père peut lire. Elle est aussi plus élevée si la famille préfère que l'enfant reçoive l'enseignement en malgache et en français.

### 5.2.3 Village

Seulement deux des variables pour le village entrent dans le modèle. La déviance est 1656,5 avec 1429 degrés de liberté pour 1448 observations. Ici, 66 observations de plus sont perdues à cause des valeurs manquantes. Les résultats sont dans le Tableau 5.4.

Le résultat le plus frappant est que le redoublement est beaucoup moins fréquent dans les villages à prédominance d'élevage (voir le tableau dans la Section B.4), surtout pour les plus jeunes enfants. Par contre, la probabilité de doubler est plus haute dans les villages protestants que dans les autres.

Aucune autre variable, surtout celles concernant la taille et la richesse du village, n'entre dans le modèle.

TAB. 5.4: Doublants avec les variables pour les enfants, la famille, et le village. ( $N = 1448, M = 667$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,722	0,680	AGRV
AGE	-0,326	0,047	FISHV
SEX	-1,548	0,615	MARKET
WATER	0,302	0,148	SHOP
RICE	0,245	0,130	VSIZE
SHOPPING	0,060	0,122	DISTF
CLASS	0,294	0,058	CATHV
SEX.AGE	0,111	0,052	
HAREA	-0,004	0,003	
ACTIV	-0,214	0,159	
FAMSIZE	0,065	0,025	
FATHED	-0,290	0,067	
MREAD	-0,238	0,144	
FREAD	0,463	0,201	
LANGSC2	0,166	0,398	
LANGSC3	0,419	0,360	
HERDV	-2,033	0,670	
PROTV	0,268	0,133	
AGE.HERDV	0,129	0,056	

#### 5.2.4 Ecoles

La seule variable significative à cette étape est le type d'école. La déviance est 1597,6 avec 1392 degrés de libertés et 1414 observations (34 valeurs manquantes supplémentaires). Les résultats se trouvent dans le Tableau 5.5.

Le résultat important ici est que les écoles privées ont un taux de redoublement considérablement plus élevé que les écoles publiques (voir le tableau dans la Section B.5).

Comme on pourrait s'y attendre, ni le nombre d'écoles ni le fait qu'il y a une école fermée dans le village n'influencent le redoublement.

#### 5.2.5 Elimination des variables

Nous pouvons maintenant éliminer trois variables, comme non-significatives, tout en gagnant 40 observations. Ce sont faire les courses, si le père participe dans les activités du village, et le choix de langue d'enseignement. La déviance est 1647,6 avec 1436 degrés de libertés pour 1454 observations. Les résultats sont présentés dans le Tableau 5.6.

Aucun des rapports entre les variables présentées plus haut ne change beaucoup.

TAB. 5.5: Doublants avec les variables pour les enfants, la famille, le village, et les écoles.  
 ( $N = 1414, M = 701$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	1,545	0,714	NUMBSCH
AGE	-0,337	0,048	TYPECSCH
SEX	-1,837	0,630	
WATER	0,327	0,151	
RICE	0,224	0,133	
SHOPPING	0,114	0,125	
CLASS	0,301	0,060	
SEX.AGE	0,134	0,053	
HAREA	-0,004	0,003	
ACTIV	-0,203	0,164	
FAMSIZE	0,068	0,025	
FATHED	-0,283	0,068	
MREAD	-0,289	0,148	
FREAD	0,474	0,207	
LANGSC2	0,123	0,399	
LANGSC3	0,362	0,361	
HERDV	-2,190	0,684	
PROTV	0,232	0,139	
AGE.HERDV	0,143	0,057	
TYPESCH2	0,213	0,219	
TYPESCH3	0,577	0,276	
TYPESCH4	0,465	0,282	



TAB. 5.6: Doublants avec les variables pour les enfants, la famille, le village, et les écoles, après l'élimination des variables non-significatives. ( $N = 1454$ ,  $M = 661$ )

Variables significatives	Estimation	Erreur standard
Constante	1,969	0,594
AGE	-0,354	0,048
SEX	-1,935	0,621
WATER	0,346	0,149
RICE	0,182	0,129
CLASS	0,322	0,058
SEX.AGE	0,141	0,052
HAREA	-0,004	0,003
FAMSIZE	0,061	0,025
FATHED	-0,267	0,065
MREAD	-0,280	0,145
FREAD	0,491	0,202
HERDV	-2,489	0,673
PROTV	0,267	0,135
AGE.HERDV	0,165	0,056
TYPESCH2	0,176	0,217
TYPESCH3	0,557	0,269
TYPESCH4	0,453	0,280

### 5.2.6 Différences regionales

Après l'introduction des différences entre les provinces, deux variables peuvent être éliminées : la taille de la maison et piler le riz (en gagnant 52 observations). Apparemment, ces deux variables varient énormément entre les provinces. La nouvelle déviance est 1650,2 avec 1485 degrés de liberté pour 1506 observations.

Le taux de redoublement est beaucoup plus élevé dans les provinces de Mahajanga et Antsiranana et il est plus bas dans celle de Fianarantsoa que dans les trois autres provinces (voir le tableau dans la Section B.6). Les variables socio-économiques dans le modèle ne peuvent pas expliquer ces différences.

## 5.3 Conclusions

Comme on peut s'y attendre, le niveau d'éducation des parents joue un rôle central dans la détermination des redoublements des enfants. Pourtant, l'effet positif de l'alphabétisme du père, augmentant la probabilité de doubler, est surprenant.

Les autres résultats importants sont le rôle des tâches domestiques, associées avec plus de redoublements, ce qui contraste avec les résultats des autres chapitres, et la bonne performance des villages d'élevage.

TAB. 5.7: Doublants avec les variables pour les enfants, la famille, le village, les écoles, et les provinces. ( $N = 1506, M = 609$ )

Variables significatives	Estimation	Erreur standard
Constante	0,880	0,624
AGE	-0,313	0,047
SEX	-1,642	0,619
WATER	0,330	0,147
CLASS	0,319	0,058
SEX.AGE	0,122	0,052
FAMSIZE	0,069	0,025
FATHED	-0,289	0,066
MREAD	-0,242	0,149
FREAD	0,486	0,208
HERDV	-2,056	0,672
PROTV	0,276	0,143
AGE.HERDV	0,132	0,056
TYPESCH2	0,396	0,224
TYPESCH3	1,111	0,285
TYPESCH4	0,826	0,297
Fianarantsoa	-0,356	0,211
Mahajanga	0,956	0,195
Antisranana	0,864	0,236
Toamasina	-0,083	0,246
Toliara	0,081	0,234



# Chapitre 6

## Abandons

Il est intéressant, et important, de comprendre les raisons pour lesquelles les enfants ne terminent pas leur éducation (surtout primaire) après qu'ils l'ont commencée. Dans ce chapitre, nous allons ajuster des modèles logistiques pour analyser les raisons pour lesquelles les enfants abandonnent l'école.

Nous allons utiliser la définition suivante de l'abandon : si l'enfant était à l'école en 1990, et s'il s'est arrêté pendant les trois années suivantes, nous le prenons comme un abandon. Nous avons choisi une période de trois ans pour éviter d'inclure des enfants qui abandonnent momentanément pendant un an ou deux.

### 6.1 Analyse par cohorte

Le taux d'abandon a été calculé par âge et par cohorte en rapportant aux enfants scolarisés (de l'âge et de la cohorte étudiés) le nombre de départs constatés, sous réserve qu'il n'y ait pas de retour dans les trois années suivant l'observation. En d'autres termes, lorsqu'un enfant n'est plus repéré dans le système éducatif pendant trois années, on considère qu'il a abandonné.

Compte tenu de ce délai de trois ans, on n'a pu observer qu'un nombre restreint de cohortes : on ne peut aller au-delà de la cohorte de 1983 dont les enfants avaient sept ans en 1990 et pour lesquels il a fallu attendre 1993 pour mesurer le nombre réel d'abandons, à l'âge de sept ans selon la définition donnée. Dans chaque cellule du Tableau 6.1, on a rapporté le nombre d'abandons constaté à l'effectif scolarisé cette année-là.

Le taux d'abandon n'est jamais très élevé. Il semble un peu plus grand pour les cohortes les plus anciennes. Un enfant qui abandonne avant 12 ans semble le faire plutôt au début de sa carrière, c'est-à-dire vers sept ans. Mais le nombre d'élèves est trop faible pour tirer des conclusions définitives.

TAB. 6.1: Abandons (%) par cohorte et âge.

Cohorte	Age						
	7	8	9	10	11	12	13
1977			8,8	5,3	3,8	9,7	9,0
1978		6,5	2,7	4,4	8,8	2,1	
1979	11,1	0,0	4,2	4,5	2,1		
1980	6,5	1,4	3,1	1,9			
1981	8,1	3,4	3,1				
1982	5,6	1,4					
1983	4,0						

TAB. 6.2: Abandon de l'école avec les variables pour les enfants. ( $N = 1041$ ,  $M = 195$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	-9,815	9,252	RICE
AGE	0,115	0,143	BROTHER
SEX	-4,772	2,953	ANIMALS
BIOLOGIC	5,737	9,118	FIELD
WATER	-0,678	0,411	SHOPPING
CLOTHES	-1,068	0,607	AFOOD
MEALS	-1,046	0,672	
SEX.AGE	0,470	0,261	

## 6.2 Abandons après 1990

Du nombre total de 4012 enfants, nous avons exclu ceux qui n'avaient pas six ans en 1990. Parmi les autres, nous avons aussi éliminé tous les enfants qui n'étaient pas à l'école en 1990, parce qu'ils ne peuvent pas abandonner par la suite. En fin de compte, il nous reste 1236 enfants inscrits en 1990, parmi lesquels seulement 40 ont abandonné l'année suivante (pour au moins trois ans) et 1196 qui sont restés à l'école pendant au moins une partie de cette période.

A la suite de ces remarques, nous ajustons des modèles logistiques pour trouver les raisons importantes de l'abandon. Nous utiliserons le modèle de l'Equation (2.1) où  $\pi$  est ici la probabilité d'abandon. La méthode d'introduction des variables et de choix des variables significatives reste la même.

### 6.2.1 Enfants

Parmi les 16 variables dans le groupe pour les enfants, six sont significatives. Les résultats sont dans le Tableau 6.2. La déviance du modèle est 217,8 avec 1033 degrés de liberté pour 1041 observations. Nous perdons 195 observations à cause des valeurs manquantes.

Pour les trois variables qui décrivent l'enfant (son âge, son sexe, et s'il est un membre biologique de la famille), nous apprenons que l'âge est lié positivement avec l'abandon :

TAB. 6.3: Abandon de l'école avec les variables pour les enfants et la famille. ( $N = 927, M = 309$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	19,890	41,64	HAREA
AGE	-4,319	4,667	HWALL
SEX	-4,457	2,973	SPMEAT
BIOLOGIC	7,690	25,52	FAMSIZE
WATER	-0,576	0,431	NUMBCH
CLOTHES	-1,105	0,622	SOCACT
MEALS	-0,991	0,704	FATHED
SEX.AGE	0,455	0,263	MREAD
ACTIV	-33,030	32,95	FREAD
RELPROT	1,002	0,429	DISTWAT
AGE.ACTIV	4,534	4,671	DISTMIN
			REASON
			LANGSC
			HOLPRF

plus âgé l'enfant est, surtout pour les filles, plus grande est la probabilité d'abandon. Les membres biologiques de la famille ont plus de probabilité d'abandon. La variable sexe n'est pas significative, mais nous la retenons dans le modèle jusqu'à l'élimination finale des variables.

Les trois autres variables dans le modèle se rapportent aux tâches domestiques des enfants (porter de l'eau, faire la lessive, et aider à la préparation des repas). Pour ces variables, nous constatons que les enfants qui les font ont moins de chance d'abandonner. Les autres variables pour le travail des enfants ne sont pas significatives.

## 6.2.2 Famille

Après l'introduction et la vérification des variables pour la famille, seulement deux restent dans le modèle. Elles se rapportent à l'activité du père dans le village et si sa religion est protestante. Les résultats se trouvent dans le Tableau 6.3. La déviance du modèle est 185,8 avec 916 degrés de liberté pour 927 observations. Avec l'addition de ces trois variables, 114 observations ont du être éliminées.

Quand le père est un agriculteur, la probabilité d'abandon augmente avec l'âge de l'enfant.

La religion du chef de famille a un effet significatif sur la probabilité d'abandon. Nous trouvons que les protestants ont une probabilité plus grande de voir leurs enfants quitter l'école. La valeur positive du coefficient signifie que, quand le père appartient à une autre religion que protestante, les enfants ont une probabilité plus faible d'abandonner.

Il est intéressant de remarquer que le niveau d'éducation des parents (FATHED, FREAD, et MREAD) n'est pas lié avec l'abandon.

TAB. 6.4: Abandon de l'école avec les variables pour les enfants, la famille, et le village. ( $N = 830, M = 406$ )

Variables significatives	Estimation	Erreur standard	Variables omises
Constante	19,420	45,33	AGRV
AGE	-4,278	4,932	HERDV
SEX	-5,618	3,311	FISHV
BIOLOGIC	7,601	29,01	MARKET
WATER	-0,748	0,469	DISTF
CLOTHES	-1,029	0,680	CATHV
MEALS	-0,513	0,751	PROTV
SEX.AGE	0,551	0,292	VSIZE
ACTIV	-33,960	34,79	
RELPROT	0,986	0,467	
AGE.ACTIV	4,682	4,932	
SHOP	3,790	3,324	
AGE.SHOP	-0,436	0,294	

Ici, toutes les variables du groupe précédent sont toujours significatives. La plupart des coefficients restent inchangés et les mêmes conclusions sont valables après l'introduction des variables pour la famille.

### 6.2.3 Villages

Quand nous ajoutons les dix variables qui décrivent le village, il n'y en a qu'une qui est significative. C'est la variable pour l'existence d'un magasin dans le village. Maintenant, le modèle a une déviance de 157,5 avec 817 degrés de liberté pour 830 observations. 97 observations sont éliminées à cette étape. Les résultats se trouvent dans le Tableau 6.4.

L'existence d'un magasin dans le village a un fort rapport négatif avec la probabilité d'abandon, surtout pour les enfants les plus âgés : s'il y a au moins un magasin dans le village, les enfants (âgés) abandonnent moins. Cette conclusion est raisonnable parce que les villages avec un magasin sont plus prospères.

Les autres variables, comme la taille du village, si l'activité principale est l'agriculture, l'élevage, ou la pêche, et la religion principale des habitants, n'ont pas de rapports significatifs avec l'abandon.

### 6.2.4 Ecoles

Quand nous introduisons les variables pour les écoles dans le modèle, nous trouvons qu'aucune n'est significative. Ceci n'est peut-être pas surprenant parce que nous nous occupons ici de la rétention des enfants déjà à l'école.

TAB. 6.5: Abandon de l'école avec les variables pour les enfants, la famille, et le village, après l'élimination des variables non-significatives. ( $N = 841$ ,  $M = 395$ )

Variables significatives	Estimation	Erreur standard
Constante	27,520	34,56
AGE	-4,344	4,895
SEX	-5,619	3,185
WATER	-0,789	0,458
CLOTHES	-1,337	0,649
SEX.AGE	0,554	0,283
ACTIV	-34,410	34,53
RELPROT	1,090	0,458
AGE.ACTIV	4,741	4,894
SHOP	3,755	3,204
AGE.SHOP	-0,447	0,284

### 6.2.5 Elimination des variables

Avec notre modèle complet pour les quatre groupes de variables, nous pouvons maintenant entreprendre l'élimination des variables devenues non-significatives. Les seules que l'on peut retirer se rapportent aux enfants. Nous éliminons si l'enfant est un membre biologique de la famille, et si l'enfant aide à préparer les repas. La nouvelle déviance est 164,2 avec 830 degrés de liberté pour 841 observations. Les résultats sont dans le Tableau 6.5.

Aucune des relations ne change beaucoup par rapport aux modèles précédents.

### 6.2.6 Différences régionales

Pour voir s'il y a des différences dans le taux d'abandons entre les provinces, nous ajoutons cette variable au modèle. Le résultat est dans le Tableau 6.6. La déviance du modèle diminue à 149,7 avec 825 degrés de liberté; il nous reste toujours 841 observations.

Avec Antananarivo comme point de référence, nous voyons que les enfants qui habitent dans les provinces de Fianarantsoa, et peut-être Antsiranana et Toamasina, abandonnent l'école plus que ceux qui habitent dans la province d'Antananarivo. Toliara n'a pas d'abandon dans l'échantillon.

Pour vérifier nos conclusions, nous présentons aussi un tableau des fréquences et pourcentages des abandons (Tableau 6.7). Le tableau a deux parties : la première donne toutes les observations (les enfants qui avaient plus de six ans en 1990 et qui étaient à l'école cette année-là) et l'autre seulement les observations utilisées dans le modèle final (en retirant les observations avec des valeurs manquantes pour les variables explicatives dans le modèle).

A partir de ce tableau, nous voyons que Fianarantsoa et Toamasina ont les plus grands pourcentages d'abandon. Les pourcentages pour Antananarivo et Antsiranana sont pres-



TAB. 6.6: Abandon de l'école avec les cinq groupes de variables. ( $N = 841, M = 395$ )

Variables significatives	Estimation	Erreur standard
Constante	26,590	33,20
AGE	-4,313	4,692
SEX	-6,844	3,483
WATER	-0,659	0,475
CLOTHES	-1,252	0,660
SEX.AGE	0,677	0,309
ACTIV	-35,550	33,16
RELPROT	1,434	0,500
AGE.ACTIV	4,791	4,693
SHOP	5,328	3,493
AGE.SHOP	-0,567	0,308
Fianarantsoa	1,618	0,611
Mahajanga	-0,503	1,129
Antsiranana	1,044	0,911
Toamasina	0,844	0,823
Toliara	-6,974	16,07

TAB. 6.7: Comparaison des abandons dans les provinces.

Provinces	Abandons (avec toutes les observations au début)			Abandons (avec les observations du modèle final)		
	0	1	Pourcentage	0	1	Pourcentage
Antananarivo	349	8	2,2%	260	6	2,3%
Fianarantsoa	242	15	4,1%	152	11	7,2%
Mahajanga	210	2	0,9%	145	1	0,7%
Antsiranana	130	4	3,0%	62	2	3,2%
Toamasina	145	11	7,1%	87	3	3,4%
Toliara	120	0	0,0%	112	0	0,0%

qu'identiques, comme ceux pour Mahajanga et Toliara. Mais, ces pourcentages ne prennent pas en compte les différences entre les régions expliquées par les autres variables dans le modèle.

Il faut noter comment le pourcentage d'abandon pour Toamasina diminue et celui pour Fianarantsoa augmente quand nous éliminons les observations avec des valeurs manquantes dans les variables explicatives. Ceci est une indication que les valeurs manquantes ne sont pas aléatoires. Par exemple, pour Toamasina, il y a proportionnellement plus de valeurs manquantes pour les enfants qui ont abandonné que pour les autres.

### **6.3 Conclusions**

Plusieurs facteurs ont des influences significatives sur l'abandon : l'âge de l'enfant ; ses tâches domestiques ; la position de la famille dans le village ; la prospérité du village avec, comme indication, l'existence d'un magasin. Comme dans tous les chapitres, il faut souligner que les conclusions doivent être interprétées avec précaution. Il y a beaucoup de réponses manquantes dans les données, et souvent elles ne sont pas aléatoires.

En contraste avec ce que l'on pense habituellement, le taux d'abandon à Madagascar est assez bas. Ce résultat concorde avec le niveau élevé de scolarisation jusqu'à treize ans.



# Chapitre 7

## Discussion

Pour conclure ce rapport, nous allons faire quelques commentaires sur les données, les variables et les modèles ajustés, et les résultats obtenus. Nous indiquerons les problèmes que nous avons rencontrés à plusieurs étapes de l'analyse.

### 7.1 Données

Comme nous avons dit dans le Chapitre 1, nous avons obtenu les données en questionnant chaque famille au sujet de ses enfants. Ainsi, comme pour la plupart des données obtenues de cette façon, il y a souvent beaucoup de valeurs manquantes, surtout dans une investigation de la taille de la nôtre et avec un travail de terrain aussi compliqué.

Dans nos données, la plupart des variables explicatives contenaient des valeurs manquantes. Quelques unes en avaient même plus de 500. La combinaison des valeurs manquantes des différentes variables a fait que le nombre d'observations disponible pour le modèle final a été parfois réduit à près de la moitié de la taille de l'ensemble de départ pour le modèle initial. En plus, certaines des variables avec beaucoup de valeurs manquantes étaient très importantes pour la scolarisation et pour le retard.

Un problème très important que nous avons trouvé avec les valeurs manquantes était qu'elles n'étaient pas aléatoires. En général, il y avait plus de valeurs manquantes pour les enfants qui n'allaient pas à l'école que pour ceux qui y allaient.

Le fait que les valeurs manquantes n'étaient pas aléatoires a eu des effets pervers sur la variable réponse : dans le cas des modèles logistiques, la combinaison des valeurs manquantes à travers les variables explicatives dans le modèle pouvait faire disparaître trop d'observations d'une des catégories de la variable réponse (ou trop peu de zéros ou trop peu de uns). Par exemple, pour la scolarisation en 1993, avant de commencer l'ajustement des variables, nous avions 1109 enfants qui n'étaient pas à l'école et 2584 qui étaient scolarisés ; mais, dans le modèle final, nous n'avions eu que 564 et 1826 enfants, respectivement. Ceci veut dire que 49% des observations pour les enfants qui n'étaient pas à l'école ont été éliminées mais seulement 29% pour ceux qui étaient scolarisés. Ceci s'est passé malgré le fait que nous n'avons pas utilisé plusieurs variables avec beaucoup de valeurs manquantes parce que nous croyons qu'elles ne sont pas aléatoires. Il y avait aussi

des problèmes semblables avec les autres variables réponses.

Les effets pervers se présentent aussi pour les variables explicatives, changeant parfois la direction de rapports. Par exemple, pour la distance à la source d'eau potable, si l'on n'exclut que les valeurs manquantes pour cette variable, la distance moyenne, pour les enfants à l'école en 1993, est 246,8 mètres et, pour ceux qui n'y étaient pas, 304,9 mètres. Quand nous prenons en compte les valeurs manquantes pour les autres variables explicatives dans le modèle pour la scolarisation, les distances moyennes se réduisent à 236,5 et 222,7 respectivement. Autrement dit, la combinaison des valeurs manquantes change la taille moyenne de la variable explicative pour chaque niveau des variables réponses d'une façon différente. Nous n'avons pas utilisé cette variable dans les modèles finaux pour cette raison.

Au lieu d'éliminer les réponses manquantes, nous pourrions utiliser d'autres techniques, par exemple un modèle de régression pour les prédire. Mais, étant donné la taille de la banque de données et le grand nombre de valeurs manquantes, ceci demanderait beaucoup plus de travail.

Nous avons aussi trouvé plusieurs problèmes avec l'encodage des variables. Ou l'encodage officiel n'a pas été suivi ou les valeurs mesurées n'étaient appropriées. Nous avons écarté certaines variables pour cette raison.

## 7.2 Variables

Il y a plus de 1000 variables dans la banque de données. C'est un très grand nombre par rapport au nombre d'observations. Ainsi, nous avons dû sélectionner certaines variables pour garder un rapport raisonnable entre variables et observations.

Nous avons choisi approximativement 50 variables qui, à notre avis, pourraient avoir un effet important sur les variables réponses. Bien entendu, il se peut qu'il y ait d'autres variables que nous n'avons pas utilisées ce qui ont un effet explicatif dans les modèles retenus.

## 7.3 Modèles

Une façon habituelle pour vérifier si les modèles s'ajustent bien est l'inspection des graphiques Q-Q (une graphique des résidus ordonnés et standardisés contre des quantiles d'une distribution normale). Si le modèle s'ajuste bien, ces résidus devraient s'aligner à 45 degrés. En réalité, c'est rarement le cas.

Nous avons tracé des graphiques Q-Q pour beaucoup des modèles logistiques ajustés. Souvent, ils étaient proches de 45 degrés.

Pour les modèles log linéaire pour le retard avec une distribution de Poisson, les graphiques Q-Q sont aussi assez près de 45 degrés, sauf pour les petits résidus. Les lignes sont un peu courbées, peut-être à cause d'une tendance dans les données (une variable importante oubliée). Pourtant, étant donné le grand nombre d'observations et les variables

dans nos modèles, nous pensons que les graphiques ne sont pas trop mauvais. Ainsi les combinaisons de variables que nous avons utilisées sont peut-être raisonnable.

## 7.4 Résultats

### 7.4.1 Scolarisation

La probabilité d'être scolarisé semblait avoir comme empêchement le plus important les activités des enfants chez eux (travail dans les champs, recherche de la nourriture pour les animaux, etc.) mais malheureusement elles avaient toutes des valeurs manquantes non-aléatoires. Pour cette raison, les variables pour la famille ont pris plus d'importance qu'autrement. Par contre, les variables pour le village ne paraissent pas du tout. Mais, la présence d'une école et son type ont un impact majeur (qui doit être nuancer à cause des problèmes avec les variables pour les enfants). Apparemment, les villages plus développés donnent plus de chances aux enfants d'être scolarisés.

Les tâches domestiques légères sont liées avec une plus grande probabilité de scolarisation, alors que les tâches agricoles vont dans le sens contraire. Les enfants qui vont à l'école prennent en charge les tâches légères tandis que ceux qui n'y vont pas ou plus assurent les tâches plus lourdes qui demandent une implication à temps plein et/ou avec des horaires incompatibles avec ceux de l'école. Nous ne savons pas si cette répartition se fait à l'intérieur du foyer ou entre familles.

Le niveau d'éducation et le patrimoine (surface de l'habitation) des parents, de même que la participation du père à des activités sociales, jouent en faveur de la scolarisation.

L'existence d'une école publique, comme une école privée, agit sur la scolarisation. Il est frappant que les filles ont plus de chances d'être scolarisées dans les villages avec seulement une école privée. Le fait d'avoir une seule école n'épuise pas nécessairement les effets de l'offre sur la demande. Pour autant que le nombre d'écoles n'est pas confondu avec la taille du village, il semble qu'il y a place à la fois pour les écoles publiques et privées, du point de vue du comportement effectif des familles. En sens inverse, l'existence d'une école fermée, qu'elle soit publique ou privée, tend à diminuer la probabilité de scolarisation.

### 7.4.2 Admission

La probabilité d'admission ne dépend que d'une des variables pour les tâches domestiques et d'une pour les tâches agricoles, positivement avec la première et négativement avec la deuxième.

Plusieurs variables des parents ont aussi un effet, mais aucune du village. Le niveau d'éducation de la mère (évalué par son aptitude à lire, plutôt que par le niveau d'études atteint) et le fait d'être catholique ou protestant jouent favorablement.

La taille du village pourrait englober l'impact positif des effets moyens de revenu, d'ouverture sur l'environnement, et de prise de conscience des effets de l'éducation.

Mais, le facteur le plus important est l'existence d'une école ouverte dans le village. Il

### 7.4.3 Retard

Le retard pour commencer à l'école a diminué pendant les années étudiées. Nous avons regardé les années de 1989 à 1993. Encore une fois, à cause des valeurs manquantes non-aléatoires pour beaucoup des variables pour les enfants, que nous avons dû écarter du modèle, les variables pour la famille ont pris plus de place (ceci est confirmé par les résultats pour la seule année 1993). Plusieurs des variables pour le village sont aussi significatives.

Le retard de scolarisation en 1993 était positivement influencé par plusieurs des tâches des enfants à la maison (porter de l'eau, faire la lessive, etc.). Ici, pour une fois, elles ont la même relation que les tâches agricoles, toutes les deux empêchant la scolarisation. Plusieurs des variables pour la famille et le village montrent des relations importantes. Des aspects divers de la famille ont eu un effet significatif sur le retard. La plupart se rapportent aux différences d'éducation et de richesse : les enfants dans les familles avec moins d'enfants, les plus grandes maisons, et une mère qui sait lire ont moins de retard. Les enfants dans les villages à majorité protestante avaient un retard moyen plus petit aussi.

Plus le niveau d'éducation des parents est élevé, plus l'enfant tend à entrer tôt à l'école. Pour l'ensemble des admissions, c'est le niveau d'éducation du père qui intervient, alors que c'est l'aptitude de la mère qui joue pour les admissions de la seule année 1993. Plus la surface de l'habitation est grande, plus l'enfant entre tôt à l'école ; plus le nombre d'enfants est élevé, plus le retard à l'admission augmente.

Alors que, pour l'ensemble des admissions, ni l'existence d'une école, ni la distance à l'école n'ont d'effet sur le retard, la seule présence d'une école privée tend à accroître le retard, pour les admissions pour l'année 1993. Ce résultat semble paradoxal, si l'on se réfère à la forte croissance des admissions dans le secteur privé constatée à partir de 1991, mais ce mouvement a concerné beaucoup plus les villes que les campagnes. Une explication possible serait que la baisse du pouvoir d'achat qui s'est manifestée depuis plusieurs années a affecté les ménages ruraux à un point tel qu'ils préfèrent désormais, dans leur majorité, une école publique pour scolariser leurs enfants, alors que ce n'était pas le cas auparavant.

### 7.4.4 Redoublement

Comme on pourrait s'y attendre, l'éducation des parents joue un rôle primordial pour déterminer si leurs enfants doublent. Vu que le fait que les enfants doivent participer aux tâches domestiques augmente le risque de doubler, en contraste avec l'effet des ces facteurs pour les autres variables réponses, il paraît que les enfants dans les familles les plus près du seuil de survie ont plus de chances de doubler.

### 7.4.5 Abandon

Pour les abandons, les caractéristiques personnelles des enfants (l'âge, le sexe, un membre biologique de la famille) jouent un plus grand rôle dans les premières étapes

de la construction du modèle. Mais, à la fin, seulement quelques variables restent dans le modèle. Aucune se rapporte aux écoles dans le village. On voit surtout que la probabilité d'abandon augmente rapidement avec l'âge de l'enfant quand le père est agriculteur.

## 7.5 Conclusions

Sur la base des données de l'étude, il apparaît clairement que la politique de l'éducation, en matière d'admission et de rétention dans l'enseignement primaire, devrait se concentrer sur quatre objectifs complémentaires :

1. réduire le nombre de doublants ;
2. concentrer les actions en faveur d'une meilleure rétention dans l'enseignement primaire ;
3. favoriser la poursuite du mouvement d'admission à l'âge normal de six ans ;
4. agir en faveur de la scolarisation des 5% à 7% d'enfants qui n'entrent pas à l'école à l'heure actuelle.

En même temps, il existe une grande variabilité entre les régions. Les provinces de Fianarantsoa, Antisranana, et Toamasina sont généralement loin derrière les autres. Ceci ne peut être expliqué qu'en partie par des facteurs socio-économique, surtout pour Toamasina.

Un grand nombre d'analyses restent à faire, à partir des données collectées : analyse de la durée de la fréquentation scolaire, selon l'âge d'admission ; étude du patrimoine et des dépenses des ménages ; analyse des dépenses d'éducation des familles, en relation avec les dépenses totales et le nombre d'enfants ; choix en faveur des écoles publiques ou privées ; relation entre les opinions exprimées par les parents et leur comportement réel, etc.

Une étude de cette taille en Europe demanderait une équipe de cinq à dix statisticiens travaillant pendant au moins un an. Beaucoup reste à faire pour obtenir une analyse réaliste et effective des données disponibles.

Cette étude fait clairement apparaître le poids déterminant et l'interaction complexe de la présence d'une école, de l'organisation de l'économie familiale, et du niveau d'éducation et de revenu des parents dans les décisions concernant l'admission, la scolarisation, et l'abandon des enfants en milieu rural. Dans cet ensemble de facteurs qui influent sur les décisions des ménages, le poids relatif des variables concernant la famille et l'enfant est très élevé : le Ministère de l'Education nationale n'a évidemment aucune possibilité d'action dans ce domaine. Son seul domaine d'intervention concerne l'existence et la proximité des écoles, mais l'effet potentiel de la qualité de l'enseignement et d'autres variables scolaires sur la probabilité de scolarisation et d'abandon n'avait pas été pris en compte dans cette étude. Ces résultats devraient peut-être amener le Ministère à réévaluer sa marge d'action face à l'objectif de généralisation des admissions.





# Annexe A

## Liste des variables explicatives

### A.1 Enfants

SEX : le sexe de l'enfant.

Garçons	Filles
2051	1961

AGE : l'âge de l'enfant.

6	7	8	9	10	11	12	13	14	15	16
319	404	432	371	430	354	437	338	370	293	264

BIOLOGIC : si l'enfant est né dans la famille (il habite peut-être avec des parents de sa famille).

Non	Oui
313	3699

BROTHER : si l'enfant s'occupe de ses frères et soeurs.

Non	Oui	VM
2755	756	501

CLOTHES : si l'enfant fait la lessive.

Non	Oui	VM
2706	820	486

WATER : si l'enfant apporte de l'eau.

Non	Oui	VM
1270	2162	580

RICE : si l'enfant pile le riz.

Non	Oui	VM
2095	1423	494

MEALS : si l'enfant aide à préparer les repas.

Non	Oui	VM
2811	702	499

SHOPPING : si l'enfant fait les courses.

Non	Oui	VM
1756	1813	443

FIELD : si l'enfant travaille dans les champs.

Non	Oui	VM
2606	929	477

ANIMALS : si l'enfants s'occupe des animaux.

Non	Oui	VM
2749	789	474

AFOOD : si l'enfant va chercher la nourriture pour les animaux.

Non	Oui	VM
3226	282	504

CLASS93 : L'année d'étude des enfants inscrits à l'école en 1993.

1	2	3	4	5	6	7	8	9	10	11
844	640	486	292	252	59	54	20	9	5	3

## A.2 Famille

NUMBCH : le nombre d'enfants dans la famille de six à 16 ans (moyenne = 5,56).

1	2	3	4	5	6	7	8	9	10	11	12
112	243	409	610	652	619	518	294	200	102	107	19
13	14	VM									
14	5	108									

FAMSIZE : la taille de la famille (moyenne = 8,08).

ACTIV : si le père est agriculteur.

Autre	Agriculteur	VM
718	3102	192

HAREA : le surface de la maison (moyenne = 30,79 m<sup>2</sup>).

HWALL : la matière de construction des murs de la maison.

Autre	Dur	VM
3671	309	32

SPMEAT : le montant d'argent (moyenne = 109.111 francs malgaches) dépensé pour l'achat de viande par la famille par an.

DISTWAT : la distance de la maison à la source d'eau la plus près (moyenne = 0,392 km).

DISTMIN : le temps en minutes pour aller de la maison à l'école (moyenne = 23,67).

RELIGION : la religion du chef de la famille.

Protestante	Catholique	Autre	VM
1316	1307	1123	266

RELOTHER : si le chef de famille a une religion autre que catholique ou protestant (modification de la variable précédente).

FATHED : le niveau d'éducation du père (moyenne = 1,52).

0	1	2	3	4	5	6	VM
1120	755	1229	516	264	19	5	104

FREAD : si le père peut lire.

Non	Oui	VM
964	2890	158

MREAD : si la mère peut lire.

Non	Oui	VM
1532	2436	44

SOCACT : si le père participe dans des activités sociales ou religieuses du village.

Non	Oui
1911	2101

REASON (1 à 10) : la raison la plus importante pour laquelle les parents ont décidé d'envoyer l'enfant à l'école (pour l'explication, voir l'Annexe B).

1	2	3	4	5	6	7	8	9	10	11	VM
1001	280	217	604	142	26	24	32	131	556	754	245

HOLPREF : quand les parents préfèrent avoir les grandes vacances.

Saison des pluies	Moisson	Pas de changement	VM
1389	683	1716	224

LANGSCH : la langue d'enseignement préférée pour les enfants.

Malgache	Français	Les deux	VM
156	540	3055	261

### A.3 Village

DISTF : la distance à la faritany (capitale du district, moyenne = 234,0 km).

VSIZE : la taille du village (moyenne = 177,8 habitations).

AGRV : si le village est surtout agricole.

Non	Oui	VM
333	3650	29

HERDV : si le village fait surtout de l'élevage.

Non	Oui	VM
2556	1359	97

FISHV : si c'est un village de pêche.

Non	Oui	VM
3278	137	597

SHOP : s'il y a un magasin dans le village.

Non	Oui	VM
1099	2371	542

MARKET : si'il y a un marché dans le village.

Non	Oui	VM
3408	472	132

CATHV : si la plupart des habitants sont catholiques.

Non	Oui	VM
2520	1327	165

PROTV : si la plupart des habitants sont protestants.

Non	Oui	VM
2857	955	200

## A.4 Ecoles

NUMBSCH : le nombre d'écoles dans le village (moyenne = 0,953).

0	1	2	3
806	2670	455	81

OPSCH : s'il y a une école ouverte dans le village (modification de la variable précédente).

TYPESCH : les types d'écoles ouvertes.

Aucune	Publique	Privée	Les deux	VM
739	2543	381	282	67

TYPECSCH : les types d'écoles fermées.

Aucune	Publique	Privée
3063	848	101

## A.5 Régions

PROV : la province où l'enfant habite.

Antananarivo	Fianarantsoa	Mahajanga
862	826	608
Antsiranana	Toamasina	Toliara
557	817	342

## A.6 Tendance dans le temps

YEAR : l'année d'admission à l'école.

# Annexe B

## Statistiques descriptives

Dans cette annexe, nous présentons trois tableaux pour chacune des variables explicatives dans nos modèles. Pour les variables continues, nous donnons un tableau des valeurs moyennes de la variable pour la scolarisation, l'admission, et l'abandon. Pour les autres variables, nous donnons des tableaux à double entrée pour les fréquences marginales et les pourcentages pour la scolarisation en 1993, les admissions en 1993, et les abandons en 1990, en plus de l'âge moyen pour commencer l'école en 1993 (retirer six pour obtenir le retard). Dans quelques cas, par exemple des variables avec quelques catégories mais utilisés comme variables continues (l'âge), nous donnons un tableau avec les fréquences marginales pour chaque catégorie en plus du tableau des moyennes.

Les seules observations omises des tableaux sont pour les enfants de six ans, puisqu'ils ne sont pas dans les modèles, sauf pour le retard.

$N$  indique le nombre de réponses,  $M$  le nombre de valeurs manquantes, et VM les valeurs manquantes. Pour les variables binaires, la première catégorie est codée zéro et la deuxième un.

### B.1 Variables réponses

**ATTENDANCE : Les enfants qui sont scolarisés en 1993.**  $N = 3693$

	Age	Scolarisation		
	moyen	0	1	%
Total	8,172	1109	2584	70,0
Total (sans l'école secondaire)		1098	2460	69,1

**ADMISSION : Les enfants admis en 1993.**  $N = 1412$

	Admissions		
	0	1	%
Total	892	520	36,8

**AGE89 : L'âge des enfants à l'admission de 1989 à 1993.**  $N = 2447$

AGE89	6	7	8	9	10	11	12	13	14	15	16
	358	497	463	410	315	215	120	43	19	6	1

**AGE93 : L'âge des enfants à l'admission en 1993.  $N = 606$**

AGE93	6	7	8	9	10	11	12	13	14	15	16
	97	182	125	71	59	29	26	5	10	1	1

**REPEAT : Enfants doublant en 1993.  $N = 2115$**

Doublants			
	0	1	%
Total	1514	601	28,4

**DROPOUT : Les enfants fréquentant l'école in 1990, avec ceux qui ont abandonné pour les trois années suivantes.  $N = 1236$**

Abandons			
	0	1	%
Total	1196	40	3,2

## B.2 Enfants

**SEX des enfants.**

$N = 3693, M = 0$				
SEX	Age moyen	Scolarisation		
		0	1	%
Garçons	8,242	567	1334	70,2
Filles	8,105	542	1250	69,8

$N = 1412, M = 0$				
Admissions				
SEX	0	1	%	
Garçons	469	262	35,8	
Filles	423	258	37,9	

$N = 2115, M = 0$				
Doublants				
SEX	0	1	%	
Garçons	761	314	29,2	
Filles	753	287	27,6	

$N = 1236, M = 0$				
Abandons				
SEX	0	1	%	
Garçons	586	21	3,5	
Filles	610	19	3,0	

**AGE moyen des enfants.**

$N = 3693, M = 0$		
Scolarisation		
	0	1
AGE moyen	11,88	10,87

## AGE des enfants.

$N = 3693, M = 0$			
AGE	Scolarisation		
	0	1	%
7	107	297	73,5
8	117	315	72,9
9	86	285	76,8
10	101	329	76,5
11	63	291	82,2
12	123	314	71,9
13	79	259	76,6
14	134	236	63,8
15	148	145	49,5
16	151	113	42,8

$N = 1412, M = 0$			
AGE	Admissions		
	0	1	%
7	101	181	64,2
8	108	125	53,6
9	77	75	49,3
10	82	62	43,1
11	57	29	33,7
12	99	31	23,8
13	58	5	7,9
14	103	10	8,8
15	102	1	1,0
16	105	1	0,9

$N = 2115, M = 0$			
AGE	Doublants		
	0	1	%
7	128	68	34,7
8	144	67	31,6
9	191	88	31,5
10	179	86	32,5
11	224	77	25,4
12	195	77	28,3
13	189	62	24,7
14	147	39	21,0
15	117	37	24,0



$N = 1236, M = 0$			
	Abandons		
AGE	0	1	%
7	97	4	4,0
8	137	2	1,4
9	185	6	3,1
10	203	4	1,9
11	238	5	2,1
12	185	4	2,1
13	151	15	9,0

**BIOLOGIC : Si l'enfant est un membre biologique de la famille.**

$N = 3693, M = 0$				
	Age	Scolarisation		
BIOLOGIC	moyen	0	1	%
Non	8,113	102	179	63,7
Oui	8,178	1007	2405	70,5

$N = 1412, M = 0$			
	Admissions		
BIOLOGIC	0	1	%
Non	82	51	38,3
Oui	810	469	36,7

$N = 2115, M = 0$			
	Doublants		
BIOLOGIC	0	1	%
Non	100	40	28,6
Oui	1414	561	28,4

$N = 1236, M = 0$			
	Abandons		
BIOLOGIC	0	1	%
Non	76	3	3,8
Oui	1120	37	3,2

**BROTHER : Si l'enfant s'occupe de ses frères et soeurs.**

$N = 3255, M = 438$				
	Age	Scolarisation		
BROTHER	moyen	0	1	%
Non	8,177	720	1827	71,7
Oui	8,277	141	567	80,1
VM	7,756	248	190	43,4

$N = 1175, M = 237$			
	Admissions		
BROTHER	0	1	%
Non	588	361	38,0
Oui	102	124	54,9
VM	202	35	14,8

$N = 1934, M = 181$

Doublants			
BROTHER	0	1	%
Non	1065	417	28,1
Oui	315	137	30,4
VM	134	47	26,0

  

$N = 1109, M = 127$			
Abandons			
BROTHER	0	1	%
Non	853	23	2,6
Oui	229	4	1,7
VM	114	13	10,2

**CLOTHES : Si l'enfant fait la lessive.**

$N = 3252, M = 441$				
CLOTHES	Age	Scolarisation		
	moyen	0	1	%
Non	8,059	712	1751	71,1
Oui	9,239	158	641	80,2
VM	7,830	239	192	44,5

  

$N = 1178, M = 234$			
Admissions			
CLOTHES	0	1	%
Non	595	413	41,0
Oui	106	64	37,6
VM	191	43	18,4

  

$N = 1941, M = 174$			
Doublants			
CLOTHES	0	1	%
Non	947	390	29,2
Oui	440	164	27,1
VM	127	47	27,0

  

$N = 1118, M = 118$			
Abandons			
CLOTHES	0	1	%
Non	677	25	3,6
Oui	412	4	1,0
VM	107	11	9,3

**WATER : Si l'enfant porte de l'eau.**

$N = 3184, M = 509$				
WATER	Age	Scolarisation		
	moyen	0	1	%
Non	7,774	413	696	62,8
Oui	8,466	354	1721	82,9
VM	7,978	342	167	32,8

  

$N = 1085, M = 327$	
---------------------	--

Admissions			
WATER	0	1	%
Non	355	181	33,8
Oui	249	300	54,6
VM	288	39	11,9

---

$N = 1965, M = 150$			
Doublants			
WATER	0	1	%
Non	397	124	24,7
Oui	1005	439	30,2
VM	112	38	25,3

---

$N = 1103, M = 133$			
Abandons			
WATER	0	1	%
Non	272	13	4,6
Oui	804	14	1,7
VM	120	13	9,8

**RICE : Si l'enfant pile le riz.**

$N = 3258, M = 435$				
RICE	Age	Scolarisation		
	moyen	0	1	%
Non	7,914	601	1283	68,1
Oui	8,735	256	1118	81,4
VM	7,935	252	183	42,1

---

$N = 1154, M = 258$			
Admissions			
RICE	0	1	%
Non	512	312	37,9
Oui	164	166	50,3
VM	216	42	16,3

---

$N = 1959, M = 156$			
Doublants			
RICE	0	1	%
Non	722	247	25,5
Oui	681	309	31,2
VM	111	45	28,8

---

$N = 1117, M = 119$			
Abandons			
RICE	0	1	%
Non	532	20	3,6
Oui	556	9	1,6
VM	108	11	9,2

**MEALS : Si l'enfant aide dans la préparation des repas.**

$N = 3249, M = 444$

MEALS	Age moyen	Scolarisation		
		0	1	%
Non	8,067	716	1850	72,1
Oui	9,373	145	538	78,8
VM	8,000	248	196	44,1

$N = 1170, M = 242$

Admissions			
MEALS	0	1	%
Non	596	425	41,6
Oui	98	51	34,2
VM	198	44	18,2

$N = 1938, M = 177$

Doublants			
MEALS	0	1	%
Non	1010	413	29,0
Oui	369	146	28,3
VM	135	42	23,7

$N = 1111, M = 125$

Abandons			
MEALS	0	1	%
Non	722	25	3,3
Oui	361	3	0,8
VM	113	12	9,6

**SHOPPING : Si l'enfant fait les course.**

$N = 3304, M = 389$

SHOPPING	Age moyen	Scolarisation		
		0	1	%
Non	8,376	576	1077	65,2
Oui	8,060	304	1347	81,6
VM	7,937	229	160	41,1

$N = 1197, M = 215$

Admissions			
SHOPPING	0	1	%
Non	472	197	29,4
Oui	232	296	56,1
VM	188	27	12,6

$N = 1959, M = 156$

Doublants			
SHOPPING	0	1	%
Non	678	241	26,2
Oui	723	317	30,5
VM	113	43	27,6

$N = 1127, M = 109$

Abandons			
SHOPPING	0	1	%
Non	571	21	3,5
Oui	526	9	1,7
VM	99	10	9,2

**FIELD : Si l'enfant travail dans les champs.**

$N = 3271, M = 422$				
FIELD	Age	Scolarisation		
	moyen	0	1	%
Non	8,004	631	1726	73,2
Oui	9,146	250	664	72,6
VM	8,163	228	194	46,0

$N = 1184, M = 228$			
FIELD	Admissions		
	0	1	%
Non	542	400	42,5
Oui	167	75	31,0
VM	183	45	19,7

$N = 1942, M = 173$			
FIELD	Doublants		
	0	1	%
Non	923	369	28,6
Oui	465	185	28,5
VM	126	47	27,2

$N = 1127, M = 109$			
FIELD	Abandons		
	0	1	%
Non	660	16	2,4
Oui	439	12	2,7
VM	97	12	11,0

**ANIMALS : Si l'enfant garde les animaux.**

$N = 3271, M = 422$				
ANIMALS	Age	Scolarisation		
	moyen	0	1	%
Non	8,069	687	1827	72,7
Oui	8,522	191	568	74,8
VM	8,317	231	189	45,0

$N = 1190, M = 222$			
ANIMALS	Admissions		
	0	1	%
Non	560	371	39,8
Oui	147	112	43,2
VM	185	37	16,7

$N = 1936, M = 179$			
ANIMALS	Doublants		
	0	1	%
Non	1044	418	28,6
Oui	345	129	27,2
VM	125	54	30,2

$N = 1116, M = 120$

	Abandons		
ANIMALS	0	1	%
Non	811	20	2,4
Oui	275	10	3,5
VM	110	10	8,3

**AFOOD : Si l'enfant cherche de la nourriture pour les animaux.**

$N = 3246, M = 447$				
AFOOD	Age moyen	Scolarisation		
		0	1	%
Non	8,184	814	2157	72,6
Oui	7,967	70	205	74,5
VM	8,163	225	222	49,7

$N = 1188, M = 224$			
AFOOD	Admissions		
	0	1	%
Non	659	447	40,4
Oui	53	29	35,4
VM	180	44	19,6

$N = 1914, M = 201$			
AFOOD	Doublants		
	0	1	%
Non	1234	495	28,6
Oui	138	47	25,4
VM	142	59	29,4

$N = 1104, M = 132$			
AFOOD	Abandons		
	0	1	%
Non	947	25	2,6
Oui	127	5	3,8
VM	122	10	7,6

**CLASS : La classe de l'enfant à l'école.**

$N = 2115, M = 0$			
Doublants			
CLASS	0	1	%
1	444	171	33,2
2	408	154	27,4
3	290	122	29,6
4	199	53	21,0
5	92	84	47,7
6	48	9	15,8
7	18	6	25,0
8	10	2	16,7
9	4	0	0,0
10	1	0	0,0

  

$N = 1229, M = 7$			
Abandons			
CLASS	0	1	%
1	453	10	2,2
2	347	10	2,8
3	209	8	3,7
4	107	7	3,3
5	49	4	7,5
6	17	0	0,0
7	6	0	0,0
8	0	1	100,0
9	0	0	—
10	1	0	0,0

**B.3 Famille****NUMBCH : Nombre d'enfants dans la famille.**

$N = 3588, M = 105$		
Scolarisation		
	0	1
NUMBCH moyen	5,715	5,511

  

$N = 1373, M = 39$		
Admissions		
	0	1
NUMBCH moyen	5,673	5,426

  

$N = 2053, M = 62$		
Doublants		
	0	1
NUMBCH moyen	5,605	5,663

  

$N = 1190, M = 46$		
Abandons		
	0	1
NUMBCH moyen	5,723	5,590

**FAMSIZE : Taille de la famille.**

$N = 3693, M = 0$		
	Scolarisation	
	0	1
FAMSIZE moyen	8,329	7,975
$N = 1412, M = 0$		
	Admissions	
	0	1
FAMSIZE moyen	8,265	7,977
$N = 2115, M = 0$		
	Doublants	
	0	1
FAMSIZE moyen	7,980	8,213
$N = 1236, M = 0$		
	Abandons	
	0	1
FAMSIZE moyen	8,131	8,075

**ACTIV : Si le père est agriculteur.**

$N = 3511, M = 182$				
ACTIV	Age moyen	Scolarisation		
		0	1	%
Autre	8,293	125	533	81,0
Agriculteur	8,188	971	1882	66,0
VM	7,444	13	169	92,9

  

$N = 1374, M = 38$			
ACTIV		Admissions	
		0	1
Autre	100	83	45,4
Agriculteur	779	412	34,6
VM	13	25	65,8

  

$N = 1982, M = 133$			
ACTIV		Doublants	
		0	1
Autre	310	126	28,9
Agriculteur	1099	447	28,9
VM	105	28	21,1

  

$N = 1153, M = 83$			
ACTIV		Abandons	
		0	1
Autre	278	2	0,7
Agriculteur	836	37	4,2
VM	82	1	1,1

**HAREA : Surface des planchers de la maison.**

$$N = 3653, M = 40$$



Scolarisation		
	0	1
HAREA moyen	26,79	32,60
$N = 1392, M = 20$		
Admissions		
	0	1
HAREA moyen	26,98	28,73
$N = 2097, M = 18$		
Doublants		
	0	1
HAREA moyen	33,66	31,60
$N = 1221, M = 15$		
Abandons		
	0	1
HAREA moyen	34,76	27,05

**HWALL : Si les murs sont construits en dur.**

$N = 3665, M = 28$				
	Age	Scolarisation		
HWALL	moyen	0	1	%
Autre	8,210	1072	2303	68,2
Dur	7,795	35	255	87,9
VM	7,333	2	26	92,9
$N = 1403, M = 9$				
Admissions				
HWALL		0	1	%
Autre	1045	566	35,1	
Dur	35	44	55,7	
VM	3	6	66,7	
$N = 2093, M = 22$				
Doublants				
HWALL		0	1	%
Autre	1337	543	28,9	
Dur	159	54	25,4	
VM	18	4	18,2	
$N = 1221, M = 15$				
Abandons				
HWALL		0	1	%
Autre	1046	36	3,3	
Dur	136	3	2,2	
VM	14	1	6,7	

**SPMEAT : L'argent dépensé pour la viande par la famille (francs malgache).**

$N = 3685, M = 8$		
Scolarisation		
	0	1
SPMEAT moyen	114.215	106.682

$N = 1412, M = 0$		
Admissions		
	0	1
SPMEAT moyen	114.206	97.674
$N = 2107, M = 8$		
Doublants		
	0	1
SPMEAT moyen	106.392	113.064
$N = 1231, M = 5$		
Abandons		
	0	1
SPMEAT moyen	119.751	77.543

**DISTWAT : Distance de la maison à la source d'eau potable (km).**

$N = 3610, M = 83$		
Scolarisation		
	0	1
DISTWAT moyen	0,304	0,246
$N = 1375, M = 37$		
Admissions		
	0	1
DISTWAT moyen	0,325	0,436
$N = 2066, M = 49$		
Doublants		
	0	1
DISTWAT moyen	0,191	0,214
$N = 1199, M = 37$		
Abandons		
	0	1
DISTWAT moyen	0,210	0,191

**DISTMIN : Temps, en minutes, de la maison à l'école.**

$N = 3009, M = 684$		
Scolarisation		
	0	1
DISTMIN moyen	25,40	23,13
$N = 1042, M = 370$		
Admissions		
	0	1
DISTMIN moyen	26,65	23,11
$N = 1822, M = 293$		
Doublants		
	0	1
DISTMIN moyen	22,63	23,64
$N = 1055, M = 181$		

	Abandons	
	0	1
DISTMIN moyen	23,60	21,78

**RELIGION : Religion du chef de famille.**

$N = 3352, M = 341$				
RELIGION	Age moyen	Scolarisation		
		0	1	%
Protestante	7,881	240	988	80,5
Catholique	8,183	264	941	78,1
Autre	8,346	476	543	53,3
VM	9,030	129	112	46,5

$N = 1262, M = 150$			
RELIGION	Admissions		
	0	1	%
Protestante	186	179	49,0
Catholique	155	175	53,0
Autre	428	139	24,5
VM	123	27	18,0

$N = 2032, M = 83$			
RELIGION	Doublants		
	0	1	%
Protestante	589	215	26,7
Catholique	593	223	27,3
Autre	280	132	32,0
VM	52	31	37,3

$N = 1190, M = 46$			
RELIGION	Abandons		
	0	1	%
Protestante	468	24	4,7
Catholique	472	8	1,7
Autre	210	9	4,1
VM	46	0	0,0

**Niveau moyen d'éducation du père.**

$N = 3595, M = 98$		
	Scolarisation	
	0	1
FATHED moyen	1,062	1,721

**FATHED : Niveau d'éducation du père.**

$$N = 3595, M = 98$$

FATHED	Age	Scolarisation		
	moyen	0	1	%
Jamais à l'école	8,463	487	536	52,4
Primaire				
Premier cycle	8,137	198	489	71,2
Second cycle	8,255	258	900	77,7
Secondaire				
Premier cycle	7,577	101	359	78,0
Second cycle	7,645	30	215	87,8
Supérieure				
Premier cycle	—	1	16	94,1
Second cycle	—	0	5	—
VM	8,250	34	64	65,3

$N = 1374, M = 38$

FATHED	Admissions		
	0	1	%
Jamais à l'école	435	139	24,2
Primaire			
Premier cycle	146	112	43,4
Second cycle	190	172	47,5
Secondaire			
Premier cycle	79	57	41,9
Second cycle	15	28	65,1
Supérieure			
Premier cycle	1	0	0,0
Second cycle	0	0	—
VM	26	12	31,6

$N = 2063, M = 52$

FATHED	Doublants		
	0	1	%
Jamais à l'école	280	134	32,4
Primaire			
Premier cycle	264	130	33,0
Second cycle	543	202	27,1
Secondaire			
Premier cycle	217	85	28,1
Second cycle	153	36	19,0
Supérieure			
Premier cycle	11	4	26,7
Second cycle	4	0	0,0
VM	42	10	19,2

$N = 1206, M = 30$			
	Abandons		
FATHED	0	1	%
Jamais à l'école	227	8	3,4
Primaire			
Premier cycle	205	11	5,1
Second cycle	418	12	2,8
Secondaire			
Premier cycle	194	6	3,0
Second cycle	113	0	0,0
Supérieure			
Premier cycle	11	0	0,0
Second cycle	1	0	0,0
VM	27	3	10,0

**FREAD : Si le père sait lire.**

$N = 3545, M = 148$				
	Age	Scolarisation		
FREAD	moyen	0	1	%
Non	8,660	430	449	51,1
Oui	8,014	634	2032	76,2
VM	8,200	45	103	69,6

$N = 1355, M = 57$			
	Admissions		
FREAD	0	1	%
Non	380	127	25,0
Oui	474	374	44,1
VM	38	19	33,3

$N = 2032, M = 83$			
	Doublants		
FREAD	0	1	%
Non	250	96	27,7
Oui	1200	486	28,8
VM	64	19	22,9

$N = 1192, M = 44$			
	Abandons		
FREAD	0	1	%
Non	197	11	5,3
Oui	955	29	2,9
VM	44	0	0,0

**MREAD : Si la mère sait lire.**

$N = 3651, M = 42$				
	Age	Scolarisation		
MREAD	moyen	0	1	%
Non	8,302	636	765	54,6
Oui	8,111	461	1789	79,5
VM	7,667	12	30	71,4

$N = 1399, M = 13$			
Admissions			
MREAD	0	1	%
Non	550	171	23,7
Oui	332	346	51,0
VM	10	3	23,1

  

$N = 2088, M = 27$			
Doublants			
MREAD	0	1	%
Non	440	193	30,5
Oui	1054	401	27,6
VM	20	7	25,9

  

$N = 1218, M = 18$			
Abandons			
MREAD	0	1	%
Non	354	15	4,1
Oui	824	25	2,9
VM	18	0	0,0

**REASON : La raison la plus importante pour laquelle les parents décident d'envoyer les enfants à l'école.**

$N = 3491, M = 202$				
REASON	Age moyen	Scolarisation		
		0	1	%
L'école n'est pas loin	7,840	318	589	64,9
Sécurité sur le chemin	8,255	58	200	77,5
Livres gratuits	8,045	64	136	68,0
Matériel gratuit	8,229	133	428	76,3
Cantine	8,870	53	77	59,2
Horaire	9,000	3	23	88,5
Heure d'ouverture	7,000	8	14	63,6
(Inconnu)	8,273	8	22	73,3
Relations				
directeur-enseignants	8,167	36	88	71,0
Relations directeur-				
enseignants-parents	8,182	108	403	78,9
Compétence des enseignants	8,426	209	493	70,2
VM	7,765	111	111	50,0

REASON	Admissions		
	0	1	%
L'école n'est pas loin	249	128	34,0
Sécurité sur le chemin	49	43	46,7
Livres gratuits	50	21	29,6
Matériel gratuit	109	82	42,9
Cantine	45	20	30,8
Horaire	3	5	62,5
Heure d'ouverture	7	1	12,5
(Inconnu)	5	10	66,7
Relations			
directeur-enseignants	27	14	34,1
Relations directeur-			
enseignants-parents	86	74	46,3
Compétence des enseignants	178	109	38,0
VM	84	13	13,4

 $N = 1996, M = 119$ 

REASON	Doublants		
	0	1	%
L'école n'est pas loin	361	135	27,2
Sécurité sur le chemin	100	50	33,3
Livres gratuits	78	44	36,1
Matériel gratuit	251	94	27,2
Cantine	48	12	20,0
Horaire	12	6	33,3
Heure d'ouverture	11	2	15,4
(Inconnu)	10	4	28,6
Relations			
directeur-enseignants	53	20	27,4
Relations directeur-			
enseignants-parents	228	91	28,5
Compétence des enseignants	268	118	30,6
VM	94	25	21,0

$N = 1171, M = 65$

REASON	Abandons		
	0	1	%
L'école n'est pas loin	272	17	5,6
Sécurité sur le chemin	91	3	3,2
Livres gratuits	61	1	1,6
Matériel gratuit	207	6	2,8
Cantine	38	2	5,0
Horaire	5	1	16,7
Heure d'ouverture	7	0	0,0
(Inconnu)	5	0	0,0
Relations			
directeur-enseignants	44	2	4,3
Relations directeur-			
enseignants-parents	169	3	1,7
Compétence des enseignants	233	5	2,1
VM	64	1	1,5

**SOACT** : Si le père participe aux activités sociales et religieuses du village.

$N = 3693, M = 0$

SOACT	Age moyen	Scolarisation		
		0	1	%
Non	8,235	590	1163	66,3
Oui	8,114	519	1421	73,2

$N = 1412, M = 0$

SOACT	Admissions		
	0	1	%
Non	484	251	34,1
Oui	408	269	39,7

$N = 2115, M = 0$

SOACT	Doublants		
	0	1	%
Non	686	258	27,3
Oui	828	343	29,3

$N = 1236, M = 0$

SOACT	Abandons		
	0	1	%
Non	524	13	2,4
Oui	672	27	3,9

**HOLPREF** : Quand les parents préfèrent les grandes vacances.

$N = 3489, M = 204$

HOLPREF	Age moyen	Scolarisation		
		0	1	%
La saison des pluies	8,264	364	913	71,5
Le moisson	8,300	197	428	68,5
Pas de changement	8,074	411	1176	74,1
VM	7,937	137	67	32,8



$N = 1268, M = 144$

HOLPREF	Admissions		
	0	1	%
La saison des pluies	280	188	40,2
Le moisson	169	84	33,2
Pas de changement	312	235	43,0
VM	131	13	9,0

$N = 2062, M = 53$

HOLPREF	Doublants		
	0	1	%
La saison des pluies	556	191	25,6
Le moisson	241	101	29,5
Pas de changement	677	296	30,4
VM	40	13	24,5

$N = 1197, M = 39$

HOLPREF	Abandons		
	0	1	%
La saison des pluies	430	9	2,1
La moisson	181	7	3,7
Pas de changement	551	19	3,3
VM	34	5	12,8

**LANGSCH : La langue à l'école préférée par les enfants.**

$N = 3446, M = 247$

LANGSCH	Age moyen	Scolarisation		
		0	1	%
Malgache	8,500	53	92	63,4
Français	8,033	105	392	78,9
Les deux	8,155	841	1963	70,0
VM	8,486	110	137	55,5

$N = 1281, M = 131$

LANGSCH	Admissions		
	0	1	%
Malgache	44	24	35,3
Français	81	76	48,4
Les deux	669	387	36,6
VM	98	33	25,2

$N = 2007, M = 108$

LANGSCH	Doublants		
	0	1	%
Malgache	55	18	24,7
Français	235	67	22,2
Les deux	1140	492	30,1
VM	84	24	22,2

$N = 1169, M = 67$			
Abandons			
LANGSCH	0	1	%
Malgache	37	2	5,1
Français	180	3	1,6
Les deux	915	32	3,4
VM	64	3	4,5

## B.4 Village

**DISTF : Distance à la faritany (capitale du district).**

$N = 3317, M = 376$		
Scolarisation		
	0	1
DISTF moyen	270,4	216,2
$N = 1255, M = 157$		
Admissions		
	0	1
DISTF moyen	265,7	214,6
$N = 1911, M = 204$		
Doublants		
	0	1
DISTF moyen	219,5	224,8
$N = 1097, M = 139$		
Abandons		
	0	1
DISTF moyen	208,9	207,0

**VSIZE : Taille du village (nombre de maisons).**

$N = 3532, M = 161$		
Scolarisation		
	0	1
VSIZE moyen	156,8	188,7
$N = 1364, M = 48$		
Admissions		
	0	1
VSIZE moyen	161,0	170,1
$N = 2013, M = 102$		
Doublants		
	0	1
VSIZE moyen	188,7	191,0
$N = 1168, M = 68$		
Abandons		
	0	1
VSIZE moyen	189,6	246,5

**AGRV : Si le village est principalement agricole.**

$N = 3666, M = 27$				
AGRV	Age	Scolarisation		
	moyen	0	1	%
Non	7,818	44	266	85,8
Oui	8,189	1063	2293	68,3
VM	8,667	2	25	92,6

$N = 1409, M = 3$			
AGRV	Admissions		
	0	1	%
Non	40	31	43,7
Oui	851	487	36,4
VM	1	2	66,7

$N = 2092, M = 23$			
AGRV	Doublants		
	0	1	%
Non	153	61	28,5
Oui	1345	533	28,4
VM	16	7	30,4

$N = 1223, M = 13$			
AGRV	Abandons		
	0	1	%
Non	145	1	0,7
Oui	1038	39	3,6
VM	13	0	0,0

**HERDV : Si l'activité du village est principalement l'élevage.**

$N = 3603, M = 90$				
HERDV	Age	Scolarisation		
	moyen	0	1	%
Non	8,207	704	1674	70,4
Oui	8,039	390	835	68,2
VM	8,684	15	75	83,3

$N = 1387, M = 25$			
HERDV	Admissions		
	0	1	%
Non	580	350	37,6
Oui	302	155	33,9
VM	10	15	60,0

$N = 2055, M = 60$			
HERDV	Doublants		
	0	1	%
Non	913	427	31,9
Oui	559	156	21,8
VM	42	18	30,0

$N = 1199, M = 37$

Abandons			
HERDV	0	1	%
Non	774	28	3,5
Oui	385	12	3,0
VM	37	0	0,0

**FISHV : Si l'activité du village est principalement la pêche.**

$N = 3250, M = 443$				
FISHV	Age moyen	Scolarisation		
		0	1	%
Non	8,201	982	2160	68,7
Oui	7,800	22	86	79,6
VM	8,091	137	414	75,1

$N = 1227, M = 185$			
FISHV	Admissions		
	0	1	%
Non	746	430	36,3
Oui	34	17	33,3
VM	112	73	39,5

$N = 1778, M = 337$			
FISHV	Doublants		
	0	1	%
Non	1208	499	29,2
Oui	57	14	19,2
VM	249	88	26,1

$N = 1017, M = 219$			
FISHV	Abandons		
	0	1	%
Non	941	32	3,3
Oui	44	0	—
VM	211	8	3,7

**SHOP : S'il y a un magasin au village.**

$N = 3197, M = 496$				
SHOP	Age moyen	Scolarisation		
		0	1	%
Non	8,386	383	621	61,9
Oui	8,053	505	1688	77,0
VM	8,261	221	275	55,4

$N = 1133, M = 279$			
SHOP	Admissions		
	0	1	%
Non	300	142	32,1
Oui	392	299	43,3
VM	200	79	28,3

$N = 1918, M = 197$

Doublants			
SHOP	0	1	%
Non	383	140	26,8
Oui	995	400	28,7
VM	136	61	31,0

  

$N = 1103, M = 133$			
Abandons			
SHOP	0	1	%
Non	249	18	6,7
Oui	819	17	2,0
VM	128	5	3,8

**MARKET : S'il y a un marché au village.**

$N = 3575, M = 118$				
MARKET	Age	Scolarisation		
	moyen	0	1	%
Non	8,198	990	2151	68,5
Oui	7,841	91	343	79,0
VM	8,538	28	90	76,3

$N = 1369, M = 43$			
Admissions			
MARKET	0	1	%
Non	798	444	35,7
Oui	74	53	41,7
VM	20	23	53,5

$N = 2046, M = 69$			
Doublants			
MARKET	0	1	%
Non	1267	494	28,0
Oui	198	87	30,5
VM	49	20	29,0

$N = 1199, M = 37$			
Abandons			
MARKET	0	1	%
Non	1003	34	3,3
Oui	158	4	2,5
VM	35	2	5,4

**CATHV : Si la majorité des habitants sont catholiques.**

$N = 3543, M = 150$				
CATHV	Age	Scolarisation		
	moyen	0	1	%
Non	8,082	735	1596	68,5
Oui	8,339	292	920	75,9
VM	7,786	82	68	45,3

$N = 1322, M = 90$

Admissions			
CATHV	0	1	%
Non	622	319	33,9
Oui	192	189	49,6
VM	78	12	13,3
$N = 2060, M = 55$			
Doublants			
CATHV	0	1	%
Non	898	387	30,1
Oui	570	205	26,5
VM	46	9	16,4
$N = 1207, M = 29$			
Abandons			
CATHV	0	1	%
Non	746	26	3,4
Oui	422	13	3,0
VM	28	1	3,4

**PROTV : Si la majorité des habitants sont protestants.**

$N = 3507, M = 186$				
PROTV	Age	Scolarisation		
	moyen	0	1	%
Non	8,220	832	1786	68,2
Oui	8,021	184	705	79,3
VM	8,174	93	93	50,0
$N = 1306, M = 106$				
Admissions				
PROTV	0	1	%	
Non	669	372	35,7	
Oui	138	127	47,9	
VM	85	21	19,8	
$N = 2046, M = 69$				
Doublants				
PROTV	0	1	%	
Non	1057	399	27,4	
Oui	399	191	32,4	
VM	58	11	15,9	
$N = 1201, M = 35$				
Abandons				
PROTV	0	1	%	
Non	805	30	3,6	
Oui	356	10	2,7	
VM	35	0	0,0	

## B.5 Ecoles

**NUMBSCH : Nombre d'écoles dans le village.**

$$N = 3693, M = 0$$

	Scolarisation	
	0	1
NUMBSCH moyen	0,669	1,106
$N = 1412, M = 0$		
	Admissions	
	0	1
NUMBSCH moyen	0,602	1,038
$N = 2115, M = 0$		
	Doublants	
	0	1
NUMBSCH moyen	1,072	1,106
$N = 1236, M = 0$		
	Abandons	
	0	1
NUMBSCH moyen	1,061	0,775

**TYPESCH : Types d'écoles ouvertes.**

$N = 3628, M = 65$				
TYPESCH	Age moyen	Scolarisation		
		0	1	%
Pas d'école	8,135	452	218	32,5
Ecole publique	8,096	530	1812	79,5
Ecole privée	8,354	77	278	69,4
Les deux	8,667	36	225	95,9
VM	8,900	14	51	78,5

  

$N = 1392, M = 20$			
TYPESCH		Admissions	
		0	1
Pas d'école	400	47	10,5
Ecole publique	387	376	49,3
Ecole privée	69	54	43,9
Les deux	26	33	55,9
VM	10	10	50,0

  

$N = 2076, M = 39$			
TYPESCH		Doublants	
		0	1
Pas d'école	165	43	20,7
Ecole publique	1058	404	27,6
Ecole privée	140	73	34,3
Les deux	131	62	32,1
VM	20	19	48,7

  

$N = 1215, M = 21$			
TYPESCH		Abandons	
		0	1
Pas d'école	149	14	8,6
Ecole publique	796	20	2,5
Ecole privée	110	2	1,8
Les deux	120	4	3,2
VM	21	0	0,0

**TYPECSCH : Types d'école fermée.**

$N = 3693, M = 0$				
TYPECSCH	Age moyen	Scolarisation		
		0	1	%
Pas d'école	8,179	680	2151	76,0
Ecole publique	8,139	398	368	48,0
Ecole privée	8,176	31	65	67,7

$N = 1412, M = 0$			
TYPECSCH	Admissions		
	0	1	%
Pas d'école	530	411	43,7
Ecole publique	340	94	21,7
Ecole privée	22	15	40,5

$N = 2115, M = 0$			
TYPECSCH	Doublants		
	0	1	%
Pas d'école	1231	527	30,0
Ecole publique	238	66	21,7
Ecole privée	45	8	15,1

$N = 1236, M = 0$			
TYPECSCH	Abandons		
	0	1	%
Pas d'école	978	29	2,9
Ecole publique	200	11	5,2
Ecole privée	18	0	0,0

**B.6 Provinces****PROV : Province où l'enfant habite.**

$N = 3693, M = 0$				
PROV	Age moyen	Scolarisation		
		0	1	%
Antananarivo	8,132	118	683	85,3
Fianarantsoa	8,248	274	484	63,9
Mahajanga	8,450	108	466	81,2
Antsiranana	8,042	223	273	55,0
Toamasina	8,136	351	393	52,8
Toliara	7,789	35	285	89,1

$N = 1412, M = 0$			
PROV	Admissions		
	0	1	%
Antananarivo	76	117	60,6
Fianarantsoa	207	102	33,0
Mahajanga	73	95	56,5
Antsiranana	207	52	20,1
Toamasina	305	109	26,3
Toliara	24	45	65,2



$N = 2115, M = 0$

PROV	Doublants		
	0	1	%
Antananarivo	412	161	28,1
Fianarantsoa	339	77	18,5
Mahajanga	218	165	43,1
Antsiranana	136	77	36,2
Toamasina	234	62	20,9
Toliara	175	59	25,2

$N = 1236, M = 0$

PROV	Abandons		
	0	1	%
Antananarivo	349	8	2,2
Fianarantsoa	242	15	4,1
Mahajanga	210	2	0,9
Antsiranana	130	4	3,0
Toamasina	145	11	7,1
Toliara	120	0	0,0

# Table des matières

<b>1</b>	<b>Introduction à l'enquête</b>	<b>3</b>
1.1	Contexte général . . . . .	3
1.1.1	Ecoles . . . . .	3
1.1.2	Programmes scolaires . . . . .	4
1.1.3	Organisation pédagogique . . . . .	4
1.1.4	Elèves . . . . .	5
1.2	Objectifs de l'étude . . . . .	5
1.3	Description de l'enquête . . . . .	6
1.3.1	Choix de l'échantillon . . . . .	6
1.3.2	Questionnaires . . . . .	7
1.3.3	Période de l'investigation . . . . .	7
1.3.4	Enquêteurs . . . . .	8
1.4	Méthodologie . . . . .	9
1.4.1	Gestion des données . . . . .	9
1.4.2	Variables réponses . . . . .	10
1.4.3	Modèles . . . . .	11
1.4.4	Présentation des résultats . . . . .	12
<b>2</b>	<b>Scolarisation</b>	<b>13</b>
2.1	Analyse par cohorte . . . . .	13
2.2	Scolarisation en 1993 . . . . .	14
2.2.1	Enfants . . . . .	14
2.2.2	Famille . . . . .	16
2.2.3	Village . . . . .	17
2.2.4	Ecoles . . . . .	18
2.2.5	Elimination des variables . . . . .	19
2.2.6	Différences régionales . . . . .	19
2.3	Conclusions . . . . .	23
<b>3</b>	<b>Admissions à l'école</b>	<b>25</b>
3.1	Analyse par cohorte . . . . .	25
3.2	Admissions en 1993 . . . . .	28
3.2.1	Enfants . . . . .	28
3.2.2	Famille . . . . .	29
3.2.3	Village . . . . .	30
3.2.4	Ecoles . . . . .	31
3.2.5	Elimination des variables . . . . .	31
3.2.6	Différences régionales . . . . .	32
3.3	Conclusions . . . . .	32

<b>4</b>	<b>Retard d'admission à l'école</b>	<b>35</b>
4.1	Analyse par cohorte	35
4.2	Retard, 1989–1993	36
4.2.1	Tendance linéaire dans le temps	36
4.2.2	Enfants	37
4.2.3	Famille	38
4.2.4	Village	39
4.2.5	Ecoles	40
4.2.6	Différences entre les régions	40
4.3	Retard en 1993	40
4.3.1	Enfants	42
4.3.2	Famille	43
4.3.3	Village	44
4.3.4	Ecoles	44
4.3.5	Elimination des variables	47
4.3.6	Différences régionales	47
4.4	Changement du retard, 1989–1993	49
<b>5</b>	<b>Redoublement</b>	<b>51</b>
5.1	Analyse par cohorte	51
5.2	Doublants, 1992–1993	52
5.2.1	Enfants	52
5.2.2	Famille	52
5.2.3	Village	53
5.2.4	Ecoles	54
5.2.5	Elimination des variables	54
5.2.6	Différences regionales	56
5.3	Conclusions	56
<b>6</b>	<b>Abandons</b>	<b>59</b>
6.1	Analyse par cohorte	59
6.2	Abandons après 1990	60
6.2.1	Enfants	60
6.2.2	Famille	61
6.2.3	Villages	62
6.2.4	Ecoles	62
6.2.5	Elimination des variables	63
6.2.6	Différences régionales	63
6.3	Conclusions	65
<b>7</b>	<b>Discussion</b>	<b>67</b>
7.1	Données	67
7.2	Variables	68
7.3	Modèles	68
7.4	Résultats	69
7.4.1	Scolarisation	69
7.4.2	Admission	69
7.4.3	Retard	70
7.4.4	Redoublement	70
7.4.5	Abandon	70

**A Liste des variables explicatives 73**

A.1 Enfants . . . . . 73

A.2 Famille . . . . . 74

A.3 Village . . . . . 75

A.4 Ecoles . . . . . 76

A.5 Régions . . . . . 76

A.6 Tendance dans le temps . . . . . 76

**B Statistiques descriptives 77**

B.1 Variables réponses . . . . . 77

B.2 Enfants . . . . . 78

B.3 Famille . . . . . 86

B.4 Village . . . . . 97

B.5 Ecoles . . . . . 101

B.6 Provinces . . . . . 103