Université d'Auvergne, Clermont-Ferrand 1
École d'Économie de Clermont-Ferrand
École Doctorale des Sciences Économiques, Juridiques et de Gestion
Centre d'Études et de Recherches sur le Développement International (CERDI)

Essays on food consumption, child malnutrition and school achievement in developing countries

Thèse Nouveau Régime

Présentée et soutenue publiquement le 26 septembre 2014

Pour l'obtention du titre de **Docteur ès Sciences Économiques**

Par

Frédéric Aubery

sous la direction de

Monsieur le Professeur Jean-Louis Arcand

Composition du jury

Vianney Dequiedt	Président	Professeur à l'Université d'Auvergne - CERDI
Mathilde Maurel	Rapporteur	Directrice de recherche CNRS - Université de Paris 1
Martina Viarengo	Rapporteur	Professeur assistant à l'IHEID - Graduate Institute
Jean-Louis Arcand	Directeur	Professeur à l'IHEID - Graduate Institute

Remerciements

Après tant d'années passées sur le terrain à collecter des données et derrière mon ordinateur à les analyser, nombreux sont ceux à qui je pense au moment de rédiger ces quelques lignes. La thèse que vous vous apprêtez à lire s'est nourrie de nombreuses missions qui m'ont amené sur les campus des universités Gaston Berger de Saint-Louis du Sénégal et Cheikh Anta Diop de Dakar, au bord de la Nya Pende dans le sud du Tchad et à Tananarive. A chaque fois, j'ai rencontré des gens formidables qui m'ont fait grandir.

En premier lieu, je tiens à remercier mon directeur de thèse, le professeur Jean-Louis Arcand, pour m'avoir donné ma chance et avoir été le déclencheur de cette longue aventure en me lançant dans le grand bain de la collecte de données. Je le remercie de m'avoir fait confiance et soutenu à chaque fois que cela fut nécessaire.

Je remercie également Mme Mathilde Maurel, Mme Martina Viarengo et M. Vianney Dequiedt qui ont accepté d'être membres de mon jury de thèse et d'assister à ma soutenance.

L'organisation d'enquêtes est un savoir-faire qui se transmet sur le terrain. A ce titre, je remercie chaleureusement Samba Mbaye, de l'Université Gaston-Berger, pour m'avoir montré comment le bon sens et la simplicité doivent guider les enquêtes ainsi qu'Elizabeth Beasley pour m'avoir inculqué la rigueur indispensable à une telle entreprise. Je remercie également Marie-Charlotte Buisson, ma partenaire de terrain et maintenant co-auteur, pour toutes les discussions qui ont animés nos projets communs et son soutient permanent.

Mille mercis à l'ensemble de mes partenaires de galère, ceux qui ont contribué de près à l'élaboration des enquêtes qui sont la bases des travaux présentés dans cette thèse et avec qui j'ai partagé des moments de joies et de doutes: Paul Randrianirina, Olivier Razafindranovona, Jocelyn Ratziferana et Dody Rasoahoby du Ministère de l'Education de la République de Madagascar; Patrick Ramanantoanina, Muna Meky et Maria Herinjara Ranohatra de la Banque Mondiale; Nicolas Babu et Soava Rakotoarisoa du Programme Alimentaire Mondial; Mino Harivelo Ramaroson et Toky Mahefa Rakotomavo de l'ONG HARDI; ainsi que Taro Boel et Ahmat Abderahim Abbo.

Je remercie le professeur David E. Sahn pour m'avoir associé à son projet de recherche sur la transition à l'âge adulte et proposé de collaborer à la rédaction de l'article qui constitue le troisième chapitre de cette thèse.

Merci à ceux qui ont relu mes travaux, répondu à mes questions et dont les discussions ont été une source inépuisable d'idées et de polémiques: Catherine Simonet (pour m'avoir donné un sacré coup de main avec les données de pluie), *le faux* Frédéric Lesné, Maxence Rageade, Mahamane Maliki Amadou, Guo Huanxiu, Emmanuel Bakirdjian, Sébastien Marchand, Karine Marazyan, Kristina Czura, Timothée Demont, Florent Bresson, Francesca Marchetta et Simone Bertoli.

Je n'oublie pas ceux qui n'ont pas relu mes travaux mais qui ont égayé mes années à Clermont. Merci à vous!

Je tiens à remercier également l'ensemble de l'équipe du CERDI, enseignants et personnel administratif, pour son accueil.

Malgré tout le talent de ceux avec qui j'ai collaboré, tout ceci n'aurait été possible sans le soutien infaillible de mes proches. J'ai une pensée émue pour ma mère, mon père et mon frère qui ont toujours cru en moi. Mais par dessus tout, c'est à Stephanie que je veux adresser les plus chaleureux remerciements pour m'avoir soutenu, conseillé et rassuré pendant toutes ces années.

Enfin, je remercie tous les élèves à qui j'ai fait passé des tests qui n'étaient pas prévu au programme, les enfants que j'ai fait pleurer sous la torture de la toise, et tout ceux à qui j'ai fait subir des questionnaires trop longs et des questions trop intrusives. La recherche en micro-économie du développement repose sur le temps que vous nous accordez.

Merci à tous.

Ce manuscrit est dédié à tout ceux qui m'ont accordé un peu de leur temps et offert quelques poignées d'arachides.

Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

BAZ	Body Mass Index for Age Z-score
BMI	Body Mass Index
CAR	Central African Republic
CERDI	Centre d'Etudes et de Recherches sur le Développement International
CISCO	Circonscription Scolaire
CONFEMEN	Conférence des ministres francophones de l'éducation nationale des pays ayant le français en partage
CM	Cohort Member
EPSPAM	Progression Through School and Academic Performance in Madagascar Study
HAZ	Height for Age Z-score
IDP	Internally Displaced People
IGA	Income Generating Activities
IHEID	Institut de Hautes Etudes Internationales et du Développement
IV	Instrumental Variable
IWMI	International Water Management Institute
MEN	Ministry of Education
MUAC	Mid-Upper Arm Circumference
MUAC/A	Mid-Upper Arm Circumference for age z-score
MUAC/H	Mid-Upper Arm Circumference for height z-score
OLS	Ordinary Least Square
PASEC	Programme d'Analyse des Systèmes Educatifs de la CONFEMEN
SD	Standard Deviation
SFP	School Feeding Program
UN	United Nations
UNHCR	United Nations High Commissioner for Refugees
VAM	Value-Added Model
WAZ	Weight for Age Z-score
WFP	World Food Program
WHZ	Weight for Height Z-score
ZAP	Zone d'Administration Pédagogique

Figure 1: An enumerator gives the instruction before the test session - Madagascar

General introduction

The intergenerational transmission of human capital is the crux of the fight against extreme poverty in the developing world. Poor health and low cognitive skills prevent individuals from seizing the economic opportunities that would help them to escape poverty. Foster and Rosenzweig (1995) showed how farmers with no education may not be able to take full advantage of new agricultural technologies and may benefit less from their own experience of using this technology. Similarly, the lack of education may hold individuals in the uncertain informal sector. The economic development of low income countries will be effective and sustainable when policies succeed in breaking this transmission of poverty from one generation to another by focusing on early childhood nutrition as well as on the production of cognitive skills.

Malnutrition is one of the most common feature of extreme poverty and the root of its intergenerational transmission. Growth is a cumulative process that is hampered by each episode of food shortage experienced by children. The first three years of life are especially crucial because this is the period when growth velocity is the greatest (Maccini and Yang (2009)). Several studies in developing countries have reported that catch-up growth is unlikely and only partial (Hoddinott and Kinsey (2001)), suggesting that any episode of nutritional deficits can have long-term consequences and that the impact of nutritional programs is very sensitive to the timing of the interventions (Schroeder et al. (1995)).

Low school achievement is one of the outcomes of chronic malnutrition: stunted children usually enroll at school later, reach lower grades and end up with low academic skills (Alderman et al. (2001), Alderman et al. (2006)). Because of poor infrastructures and weak teacher's qualification, the rise in primary school enrollment experienced by developing countries over the past fifteen years was not always followed by an increase in the average knowledge as measured by test scores (Jones et al. (2014)). In other words, the value-added of an additional year of schooling remain low, which questions the potential return to schooling on the job market (Boissiere et al. (1985)), one of the main incentives for schooling, as well as its impact on the economic growth (Hanushek and Woessmann (2008)). There is no doubt that the benefits from improving knowledge acquisition are intergenerational since the level of schooling of parents and teachers are strong predictor of children's success at school.

Moreover, Marchetta and Sahn (2012) have shown how education can reduce the risk of early childbearing.

The widespread prevalence of malnutrition observed in developing countries and its consequence on school achievement are two illustrations of the extreme vulnerability of the poor. There is substantial evidence in the literature that unexpected variation in a household's resources can have detrimental effect on its members' welfare. Drops in the agricultural production due to a drought or a locust invasion, unexpected health expenditures or the death of a family member are usual and recurrent shocks in most part of rural Africa. When households are unable to cope with them by smoothing their own consumption, shocks could have a permanent effect on child's growth (Hoddinott and Kinsey (2001)) or schooling (Gubert and Robilliard (2008)). In other words, parents' vulnerability is transmitted to their children. One of the key roles of the research in development economics is to assist decision-makers by leading policy-oriented research and disentangling what leverage should governments operate to get a healthier and more educated population.

The aim of this doctoral thesis is to contribute to the literatures on food consumption, child malnutrition and school achievement. The chapters that constitute it are based on three original microeconomic databases that were collected with the purpose of giving clear answers to decision makers. For two of them (first and second chapter), I was personally involved in the design elaboration, the tools creation and the survey implementation. The dataset used in the first chapter was expected to be a baseline for the impact evaluation of a school feeding program, which was not implemented finally because of the putsch of March 2009 in Madagascar. The second chapter is based on data collected for the United Nations High Commissioner for Refugees (UNHCR) and Care International with the aim of assessing refugee households' self-sufficiency. Finally, the third chapter relies on a long-term panel data, which remains uncommon in the developing world. Each of these data collection processes can illustrate some of the challenges that must be faced on the field to obtain unbiased estimates in later analyses.

In the first chapter, I focus on the relationship between malnutrition and the acquisition of basic knowledge in a large sample of Malagasy children enrolled in primary school. The widespread school-based interventions, like school feeding programs, rest on the assumption that malnutrition is a major obstacle to the acquisition of skills at school. As both nutrition and educational outcomes results from parents' choices, the relationship is complex and its estimation challenging. Unobservable school, family and child's characteristics (e.g. parent's preferences or the student's innate ability) are likely to lead to a severe endogeneity issue. Based on a large cross-sectional data set from Madagascar (about 6000 primary school students), this chapter proposes an instrumental variables method to estimate the causal relationship between nutrition and school achievement. The variation in the total amount of rainfall from one year to the next during the first years of a child's life are used as exogenous instruments to predict the long-term nutritional status (height for age z-score). Once instrumented the effect of malnutrition on test scores appears to be larger, suggesting that naive ordinary least squares estimates of this relationship were biased downward. This chapter illustrates how household's vulnerability is transmitted to the next generation. The household's inability to smooth its consumption after an unexpected income shock has a large negative impact in the long run on children's growth and their school achievement.

The second chapter, which was written in collaboration with Marie-Charlotte Buisson (IWMI-CGIAR), is based on data collected in southern Chad among a population of refugee households from Central African Republic. Being a refugee is often a long-term situation where households have to deal with the loss of their lands, cattle and assets with the assistance of relief agencies and international non-governmental organizations. During this uncertain situation, refugee households receive a monthly food aid as well as a piece of land and farming inputs. This food ration protects them from the traditional lean season most households face once a year in a rural area such as southern Chad. In spite of this regular food intake, a large share of children born in camps suffers from long and short-term malnutrition at the time of the survey. With the help of a large household database collected in three refugee camps and an original survey design that allows us to compute an exogenous time interval between the food distribution and each household interview, we estimate the effect of time on food consumption. Results suggest that the more distant the food distribution, the less important is household's consumption of cereals. The effect is sufficiently large to impact children short-term nutritional status: the average mid-upper arm circumference over a sample of children aged 6 to 59 months, an indicator of short-term malnutrition, significantly decreases as the time since the previous distribution goes by. This finding suggests that the rhythm of the distribution has an impact on the way households use their ration. In other words, despite the regular food distribution, household are unable to smooth their consumption within the time interval between two food distributions, which creates a monthly episodes of food scarcity.

The third chapter was written in collaboration with the Professor David E. Sahn (Cornell University). The objective is to assess the role of schooling in the formation of cognitive skills among a cohort of Malagasy young adults. The estimation of an education production function requires a complete history of family and school inputs. Some of these factors, like parents' preferences or individual innate abilities, remain unobservable. Others, like the timing of schooling or attendance, result from a choice made by parents. Therefore, the challenge of this literature is to deal with unobserved heterogeneity that could bias the estimation. The panel dimension of our dataset allows us to estimate a value-added model in which we assess the determinants of final achievement at adulthood, as measured by a test score, by controlling for prior achievement, measured during teenage years. We combine this approach with an instrumental variables method: the availability of a secondary school in the commune and proxies for the presence of a credit market are used as exogenous variables to predict grade progression and prior achievement. Results indicate a strong effect of additional schooling. Secondly, parents' education and household's wealth, two of the main family inputs associated with grade progression, have a very limited direct impact on learning. Thirdly we observe acceleration in learning: adolescent who had better results at the baseline tests have acquired higher knowledge at adulthood.

Bibliography

Alderman, H., Behrman, J., Lavy, V., and Menon, R. 2001. Child health and school enrollment: A longitudinal analysis. *Journal of Human Resources* 26:185–205.

Alderman, H., Hoddinott, J., and Kinsey, B. 2006. Long term consequences of early childhood malnutrition. *Oxford Economic Papers* 58:450–474.

Boissiere, M., Knight, J. B., and Sabot, R. H. 1985. Earnings, schooling, ability, and cognitive skills. *The American Economic Review* pp. 1016–1030.

Foster, A. D. and Rosenzweig, M. R. 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy* pp. 1176–1209.

Gubert, F. and Robilliard, A. S. 2008. Risk and schooling decisions in rural Madagascar: A panel data-analysis. *Journal of African Economies* 17:207–238.

Hanushek, E. A. and Woessmann, L. 2008. The role of cognitive skills in economic development. *Journal of Economic Literature* pp. 607–668.

Hoddinott, J. and Kinsey, B. 2001. Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics* 63:409–436.

Jones, S., Schipper, Y., Ruto, S., and Rajani, R. 2014. Can your child read and count? Measuring learning outcomes in East Africa. *Journal of African Economies* p. eju009.

Maccini, S. and Yang, D. 2009. Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review* 99:1006–26.

Marchetta, F. and Sahn, D. E. 2012. The role of education and family background in marriage, childbearing and labor market participation in senegal. CERDI Working Paper N° 201224, CERDI, Université d'Auvergne.

Schroeder, D., Martorell, R., Rivera, J., Ruel, M., and Habicht, J. 1995. Age differences in the impact of nutritional supplementation on growth. *The Journal of Nutrition* 125:1051S–1059S.

Chapter 1

The detrimental effect of malnutrition on school achievement: Evidence from two districts of Madagascar

1.1 Introduction

Despite their relatively high cost for developing countries, school-based health interventions like school feeding programs are widespread across the world. According to Galloway et al. (2009), the mean cost of a school feeding program (hereafter, SFP) is of US$40 per child per year in Africa. In Madagascar, for the school year 2011/2012, giving children a meal that corresponds to a food intake of 603 kilocalories for 175 days costs US$35 per child. This average cost is very sensitive to food and transport price rises. During the 2008 food price crisis, the average daily cost per child had risen from US$0.19 to US$0.30 in Madagascar in a few months, leading the World Food Program (WFP) to consider a cut in the treatment duration from ten to four months by focusing only on the lean season, the period of the year when the weakest pupils are most likely to leave school temporarily or definitively. Behind the generic term of school feeding lie different kinds of interventions. Rations are either distributed at school or could be brought at home. The ration could be a glass of milk, a snack or even a whole meal and could be fortified or not. The objectives are similar and are often very ambitious. According to Kazianga et al. (2009), whatever the kind of intervention, three main goals stand out: increasing enrollment rate and attendance, improving nutrition itself and, lastly, improving school achievement through a better nutritional status. Any of those three goals are expected to have a large long-term effect on the accumulation of human capital and, therefore, returns to SFP should largely encompass the initial cost for the society.

The first goal, increasing enrollment, is the educational part of a school-feeding program. Under this objective, school-feeding programs can be seen as a conditional cash transfer to parents whose children are enrolled. Giving a daily meal is expected to decrease the opportunity cost to send one more child to school (rather than keeping him at home for housework or to expand the household's labor force for agricultural activities) and increase time spent at school by students. In the developing world where school enrollment is very low, SFP are seen as an instrument to attract children to school and keep those who are already enrolled. Kazianga et al. (2009) have found a positive effect on enrollment of two kinds of school feeding programs (on-site feeding or take-home rations) but a negative effect on attendance. They argue that, in the presence of an imperfect labor market, parents can't replace children labor force. With the help of a randomized controlled evaluation design in Uganda, Alderman et al. (2012) found a positive effect of SFP on enrollment and attendance but highlighted that SFP could increase the time taken to complete school. Increasing enrollment is a straightforward goal that doesn't come without drawbacks: higher enrollment rates could lead to overcrowded classrooms if school inputs are not adjusted.

The second goal, improving children's nutritional status, is the pure nutrition part of the program where school is used as a hub to reach children easily.[1] In that respect, SFP should be seen as a food aid distribution toward pupils to prevent them from long-term (stunting) and short-term malnutrition (thinness) or anemia. It raises several concerns. Firstly, on a long-term point of view, is school-age a good timing for a nutrition program? According to Martorell (1999), the two first years of life are a crucial period for child growth, especially for the brain, as the growth rate is particularly high during this period. It is unlikely that any caloric supplementation at school-age could completely offset early childhood nutritional deficits as pupils may be too old to completely recover from growth deficit (Buttenheim et al. (2011)). Secondly, are SFP effective to improve children short-term nutritional status? According to Jacoby (2002) food supplementation stuck to children: there is no reallocation of food in the household in disfavor of the child who benefits from a SFP, with the exception of the poorer ones. On the other hand, Kazianga et al. (2009) have shown how take-home ration in Burkina-Faso could also have a positive effect on the nutritional status of siblings under five years old. While it confirms that food distribution can succeed in improving children nutritional status, this result also suggests that children under five years old are particularly sensitive to caloric supplementation and that take-home rations make it more difficult to target pupils only.

The third objective links nutrition to school achievement. Under the assumption that poor nutrition hampers pupils' learning process, a daily food intake is expected to reduce the negative effect of malnutrition. This pattern is more complex than the first two objectives.

[1]Because it reduces the cost to track children, reaching them at school is an efficient way to treat a large number of children. Health and nutrition programs such as micronutrient supplementation or deworming that sometimes go along with a SFP often use this solution. However, it could lead to a selection bias in countries where a large share of children is not enrolled.

Firstly, meal distribution at school could reduce short-term hunger, enhance pupils' attention span and then, increase their daily learning time. Given at the appropriate time during the school day, a food intake could help children to keep concentrating on the lesson. Except from the direct short-term impact through attention, the pattern is not obvious. A crucial point is to understand whether the negative effect of the short-term malnutrition matters more than the long-term one? If the effect of stunting is more detrimental to learning (in other words, if early childhood nutrition has long-lasting effect on school achievement), SFP may be less effective than a nutrition programs devoted to preschool children. In the case thinness at school-age has a negative effect on learning, independently from stunting, could a SFP shifts the nutritional status enough to have a positive effect on learning and could this effect be larger for stunted children? Impact evaluations only found little or no impact of SFP on school achievement. Kazianga et al. (2009) found no impact on learning, as measured by a mathematics test score. According to Grantham-McGregor et al. (1998) in an impact evaluation in Jamaica, giving a breakfast to pupils enhances their cognitive functions but the positive result depends heavily on schools structure.

The mixed impact of SFP on school achievement raises a lot of questions. Why aren't the impacts of school-feeding on tests scores conclusive? What is the perfect timing for a nutrition intervention? Is there a large enough causal effect of poor nutrition on school achievement that could justify a SFP? What is the more detrimental for school achievement between short and long-term nutritional status? The literature on the impact of health and nutrition on school achievement is very large but fails to give a clear answer to these questions. According to Behrman (1996), the main challenge to assess this relation is to deal with the endogeneity of the nutritional indexes as both health and education outcomes result in part from parents' choices. In the absence of a good identification strategy, the effect of nutrition on knowledge acquisition will be either overestimated or underestimated.

The goal of this paper is to take advantage of a large cross-sectional dataset designed as a baseline for the impact evaluation of a SFP, to assess the causal relation between poor nutrition and knowledge acquisition which is one of the main assumption behind any SFP. We explore the effect of short and long-term nutrition on mathematics and reading tests scores and find that mild long-term malnutrition (stunting), but also short-term malnutrition (thinness) have a detrimental effect on learning. We contribute to a large literature that use an instrumental variables strategy to control for the endogeneity of children long-term nutritional status (see Wisniewski (2010), Alderman et al. (2001) or Alderman et al. (2006)). Using rainfall shocks during the preschool period as exogenous instruments, we find that mild long-term malnutrition has a bigger negative impact than previously measured with a *naive* OLS estimation.

The next section presents the conceptual framework. Section three and four briefly discuss the context of the data collection and displays some descriptive statistics. Section five explains the identification strategy and section six presents the results.

1.2 Conceptual framework

Following the framework presented by Glewwe and Miguel (2007), we estimate the relationship between malnutrition and school achievement as a knowledge production function. We assume two periods. The first one corresponds to the preschool period when parents invest in the production of their child's health and nutrition. Children attend school during the second period.[2] Learning outcomes at time t (say pupil's knowledge expressed as a test score) rely on nutrition in both periods (t and $t+1$), parents' educational inputs in both periods, child's ability and observable characteristics and school characteristics.

We estimate the following education production function:

$$S_{ij} = \beta_0 + \beta_1 C_{ij} + \beta_2 LTN_{ij} + \beta_3 STN_{ij} + \beta_4 SC_j + \beta_5 H_{ij} + v_j + \varepsilon_{ij} \qquad (1.1)$$

where S_{ij}, the dependent variable, is the educational outcome (here a test score) for child i from school j ; C_{ij} is a vector of child's characteristics that includes age, gender and grade ; LTN_{ij} is the long-term nutritional status of the child ; STN_{ij} is the short-term nutritional status ; SC_j is a vector of school characteristics and H_{ij} is a vector of household's characteristics. LTN_{ij} is determined at the first period and STN_{ij} at the second. An estimation of this model should give us the estimated relationship between nutrition and learning once controlling for child, household and school observable characteristics. Heterogeneity across schools infrastructure can be controlled with the help of a fixed effects method (v_j).

The estimation of the nutrition indicators effect should be interpreted carefully. Behrman (1996) and Alderman et al. (2001) argue that both nutrition and education outcomes are determined by parents choices, on the basis of their assessment of the situation and preferences about their child's welfare, which is widely unobservable. As a consequence, the estimation of equation (1.1) may suffer from the endogeneity of the nutritional status, leading to biased estimates of parameters β_2 and β_3. According to the author, the direction of the bias (whether the effect will be underestimated or overestimated) is unknown at first sight. Let's assume two types of parents: parents who have a strong preference for their children's welfare in a general way (i.e. who invest in both health and education) and parents that don't. One can also imagine that parents may choose to encourage the education of a healthier child rather than the weaker ones. In this case, a simple estimation with OLS would overestimate the real effect of nutrition on school achievement since the same group of children will benefit from in-

[2]In a potential third period, individuals keep learning skills outside of school (Behrman et al. (2008)). As we only consider school-age children, this period is out of the scope of this paper.

vestments in nutrition and education. In a second time, let's assume parents differ in the way they invest in their child's welfare: parents with a strong preference for health and parents with a strong preference for education. In this case, the real effect will be underestimated with OLS because children will only benefit from an investment in one of the two outcomes. As Behrman and Lavy (1994) suggested it, unobserved heterogeneity at child (innate ability) and community level (allocation of school inputs or health infrastructures) is also likely to lead to an endogeneity issue.

1.3 Data

1.3.1 Context

Unlike the east coast, where the precipitations are heavy, the south of Madagascar is a dry climate region. As a consequence, agricultural production yield is lower than in the rest of the country. The production of rice, both the most widely grown and consumed cereal in the country, is limited by the lack of precipitation. The two main crops grown in the south are maize and cassava. Households face a severe lean (or hungry) season from November to March, which duration depends on the size of the previous harvest (from March to June) and time left to the next. The lean season takes place during the school year that goes from October to July and plays a crucial role in child's attendance.[3] There is suggestive evidence that, when food is scarce, parents keep children at home to expand the household's labor force for farming activities and gathering (Gubert and Robilliard (2008)). During the lean season, plucking scarcity food like the cactus fruit, *raketa* in Malagasy, is a very usual activity, especially for poorer households (Kaufmann (2004)). Sometimes children are just too weak to go to school, in particular if the school is far from the house. Therefore children's learning slows down and some of them leave school definitely. As Glick et al. (2005) points out, the south of the island (especially the former administrative region Toliary) faces the lowest primary enrollment rate of the country.

In order to tackle this issue, the Ministry of Education (MEN) decided to implement a SFP in primary schools from the Androy region (in 2009, 611 schools in five CISCO[4] were covered).[5] The goal was to increase the students' learning time by improving school atten-

[3] The lean season coincides with a substantial decline in pupils' attendance. At the time of the project, the Ministry of Education was thinking about moving the beginning of the school year from October to February to avoid a part of the hungry season.

[4] A CISCO (*Circonscription scolaire* in French) is an administrative zone, a district area that includes more than a hundred schools. There are 114 CISCO in the country.

[5] This SFP was led by the MEN, financed by external donors (including the World Bank) and implemented by the U.N. World Food Program. The program consists of a daily distribution of a non-fortified meal to every pupil who attends school the day of the distribution. The meal, which represents a calorie content of 603kcal, was composed of 100g of rice, 15g of leguminous plant and 15cl of cooking oil. The distribution is effective every

dance and by getting children more focused while they were in class. This program clearly rests on the assumption that short-term malnutrition is a main obstacle to learning.

In 2008, the MEN and its donors have decided to extend the program and to assess its impact with the help of a randomized controlled trial impact evaluation. Two additional CISCO (Amboasary and Betioky) were targeted because they share a border and a similar climate with the first five targeted CISCO. Data used in this paper were collected in March-April 2009 and should have been a baseline for the impact evaluation. Following the coup of March 2009, donors' funds have been frozen. As a consequence, the program extension and the evaluation project were cancelled. Unfortunately, data can't tell anything about the impact of a SFP but it can give information about the relationship between nutrition and school achievement.

1.3.2 Survey design

The survey was designed as a baseline for the impact evaluation. The main targets were primary school pupils, but also the teachers and director of the school as well as the pupils' parents. In order to get enough statistical power for the evaluation, 350 primary schools were randomly selected from an official list of schools. In the end, nearly every primary schools of the two CISCO were visited with the exception of schools with a single grade that were excluded beforehand[6] and the schools that were found closed. Some schools were closed from the beginning of the school year, some others were closed recently because of the path of a cyclone. The unrest that preceded the political coup had no direct impact on the survey implementation: troubles were limited to the main cities and rural schools were not closed. Neither the parents nor the schools' employees were aware of the extension of the school feeding program. The final sample, which is made up of 299 primary schools, is a heterogenous group. Schools differ by status and by size. Most of them are public but twenty-height are private and six are community schools (small newly opened schools waiting for being officially referenced as public school). Only half of schools had a complete cycle at the time of the survey.

In every school, students were randomly selected among those who were attending school the day of the survey. It is important to note that pupils were sampled during the first minutes of the school visit and that schools weren't aware of the visit. The aim of this strategy was to avoid sampling children that would not have attended school in an ordinary day. Five children

day of the school year (nearly 170 days) and is supposed to happen at the morning at the middle of the school day that usually goes from 7:00am to 12:30pm.

[6] Pupils from the five grades were interviewed but a special attention was paid to second grade children because they were offering a double advantage: with at least a year and a half of schooling (for those who hadn't repeated grades) they had enough reading and writing skills to be tested but were also young enough not to leave primary school during the two following years, when the follow-up surveys were foreseen. Primary schools with only first grade teaching were then excluded from the sample.

per grade (ten in second grade) were selected in each school. Then, depending on the school size, the sample is composed of fifteen to thirty students per school, for a total number of 6500 children. Pupils were measured (height, weight and mid upper arm circumference) and were given two individual oral tests to assess their knowledge in mathematics and reading.

The school survey was completed by a household survey for 1500 pupils over the 6500 sampled ones in order to gather information about the family inputs that can play a role in the accumulation of human capital. In each school, six pupils were randomly picked among the ten second grade children of the sample. In each household, school-age children, with the exception of the already sampled ones, were tested and measured even if they weren't enrolled at school. Unless it is specified, the siblings of the children selected at school are not included in the estimation sample.

The survey was implemented (between March 11 and April 5, 2009) during the second half of the school year, in the middle of a harsh hungry season. Given the survey design, the sample is not fully representative of the children of the CISCO. Since children have only been targeted at school during enumerators' visit, the sample misses children that were not enrolled at school and those that were enrolled but not present and there is no doubt that those children are different from the sampled ones. Under the assumption that pupils that have dropped out during the present school year are the weakest (those with the lowest nutritional status), their absence in the sample implies that any estimation of the effect of nutrition on education would be underestimated.

1.4 Descriptive statistics

1.4.1 Educational outcomes

Primary education in Madagascar features a relatively higher enrollment rate than the rest of Sub-Saharan Africa (79% of children were enrolled in 2003 on a national basis according to World Development Indicator), no gender bias in enrollment, a poor primary completion rate (51% in 2003 - World Bank) and a high repetition rate (30% in 2003). In theory, enrollment is free but fees are often collected by the association of parents of students in order to hire additional teachers (Glick et al. (2005)). In our household sample, nearly 83% of children aged five to fifteen were currently enrolled at school and the proportion is significantly higher for girls (86% over 79% for boys). We observe an enrollment peak at ten years old for girls (95% of girls enrolled) and nine years old for boys (88%). Even if they are probably overestimated because of the specific household sample selection,[7] those enrollment rates could cast doubt on the size of the effect of a SFP on total enrollment. However, being enrolled doesn't mean

[7]By definition, households in the sample had at least one enrolled child.

that students actually attend school during the whole academic year. Unfortunately, we are missing details about individual attendance throughout the year. Consistent with the literature, the secondary enrollment rate is worrying: less than 8% of children aged eleven to fifteen were currently enrolled in lower secondary school. Among the adult population, almost 16% of them have declared their highest grade reached is above primary school.

Even if the situation is very heterogeneous, pupils face very poor learning conditions. In our sample, only 45% of schools have a complete cycle (five grades) and one third of them do not have more than the first three grades. Large complete schools (sometimes with more than one classroom by grade) can be found in the main towns of the two CISCO. On remote areas, schools tend to be smaller and with fewer teachers: in 35% of the sampled schools, there is only one teacher. As a consequence, multigrade classroom is widely used (57% of pupils in the sample are in multigrade classroom) and classrooms are largely overcrowded. Consistent with what Michaelowa (2001) reported, the repetition rate is high: 37% of pupils have repeated at least a grade (self-declared) and 24% are currently repeating. The rate is decreasing with age. In this context, any increase in the total number of pupils following a SFP should be considered carefully.

In order to assess their cognitive skills, children were given two types of test: a mathematics test and a reading test. Both of them were oral and given individually to students regardless of their grade. Pupils were new to these kinds of tests. The mathematics test was composed of eighteen questions on basic knowledge from digit recognition to multiplication and division. The maximum score is twenty-four. According to the Cronbach's alpha value (0.88 on the overall sample), the test is largely consistent and reliable.[8] The test turns out to be more relevant for students of the first three grades. For upper grades, the test was not hard enough to distinguish between two pupils that already have acquired basic knowledge. In other words, the standard deviation is higher for the lower grades. Figure (1.1) shows the mathematics test score distribution by grade. The average score is 15.7 for the whole sample and goes from 11 for first grade pupils to 20 for fifth grade pupils. Only 1% of children had scored perfectly.

The reading test was in two parts. Pupils had sixty seconds to read as much letters as possible (first part) and sixty more to read as much words as possible in a short text in Malagasy (second part).[9] If the student stumbles over a letter or a word, enumerators had to suggest the following one. The final score is the total number of letters or words a pupil had correctly read. Contrary to the mathematics test, the reading test was more relevant for pupils in upper grades in the sense that the standard deviation increases with grade. In early

[8] Cronbach's alpha measures the internal consistency of a test score, i.e. the intercorrelations between answers to a test. The internal consistency is considered good when the coefficient is over 0.80 and excellent over 0.90. The coefficient can't be computed for the reading test.

[9] One important advantage of the letters reading test is that it is independent of a specific language (French or Malagasy).

grades, pupils had not acquired basic knowledge: first grade pupils only start to recognize and to write letters on slate during the second half of the first year. For upper grades, the test was sensitive enough to distinguish between two pupils with good reading skill. In theory, there is no upper bound for the score. Figures 1.2 and 1.3 show the distribution of the number of letters and words read by grade. The average number of letters (words) read is 31.4 (21.4) for the whole sample and reaches 65.0 (51.0) for fifth grade pupils. First grade children could only read five letters on average, and strikingly, a quarter of second grade children are not able to read more than five letters.[10]

Pupils with better mathematics skills tend to have better reading skills too: the correlation between mathematics and reading test scores is large and positive over the whole sample (0.60).This result mainly reflects the progression of knowledge through grades. The degree of correlation, while still positive and significant, is smaller once estimated by grade (from 0.30 for first grade to 0.15 for fifth grade). As expected, students' performance is linked to specific school characteristics. Similarly to Glick et al. (2005), average scores appear to be significantly lower in public schools (rather than private ones) and in multigrade classes. Moreover, pupils from remote schools, which tend to be smaller ones, performed worse than others to the tests.

1.4.2 Children nutritional status

In order to assess the long and short-term pupils nutritional status, several indicators have been computed with the help of weight and height measurement and expressed as standardized deviations from means of a population of reference.[11] The z-score of the height for age ratio (HAZ) is an indicator of chronic deficits (long-term nutritional status) that assesses stunting in children aged five to nineteen. A child is considered as stunted if its HAZ is two standard deviations below the mean of a healthy population. The z-score of the body mass index for age (BAZ) provides information about acute or recent deficits in nutrition (short-term nutritional status). A child is considered as wasted if its BAZ is more than two standard deviations below the mean of a healthy population. The z-score of the weight for age ratio

[10] The correlation is high between the two reading test scores (0.80). That's why we will mainly focus on the letters reading score on the results section.

[11] The measurement protocol respects the one established by UNICEF. The used tools, which are made of wood and 170cm high, follow the UNICEF model. Given their age, the children were measured in stand-up positions. For the weighing, commercial electronic bathroom scales were used. The children were weighed barefooted with their clothes. Z-scores were generated with the help of the World Health Organization's Anthroplus software. The calculation needs precise information about children's age. Computation was made on the basis of the date of birth. The day of birth was imputed, day 15 being assigned to each child to limit the size of the potential bias. The month of birth contains a lot of measurement errors: over 70% of children have been reported to be born in January. Such an error is supposed to lead to an underestimation of malnutrition. However, no difference in the prevalence of stunting or wasting has been noticed between children born in January and those born during the rest of the year. As a consequence, the month of birth was not modified.

(WAZ) reflects acute and chronic deficits in nutrition. A child is considered underweight if its WAZ is more than two standard deviations below the mean of the reference population. As WAZ is unreliable for children over ten and then inappropriate, HAZ and BAZ have been preferred for the econometric analysis.

According to those ratios, the situation in the south of Madagascar is worrying. The prevalence of malnutrition is very high in the sample. Nearly one third of the children (31.6%) are stunted and 10.3% suffer from severe stunting (HAZ below three standard deviations). The average HAZ in the school sample is -1.31, which suggests that children have experienced several episode of deficit in nutrition during their lifetime. Consistent with the medical literature on school-age children (Lwambo et al. (2000), Friedman et al. (2005)), stunting progresses with increasing age, suggesting a limited catch-up growth over the studied population: half children aged fourteen to fifteen are stunted (one quarter are severely stunted) whereas 17% of children aged five to seven are (5% are severely stunted). Girls show a better long-term nutritional status. This difference is significant over the whole sample (where girls are younger than boys) but also by age (especially for the younger and the older pupils). One important finding is that the prevalence of stunting is higher over the children measured at home (student's siblings from the household sample) than over children from the school sample (restricted to students aged five to ten): 32% of children measured at home (aged five to ten) are stunted whereas only 22% of children selected at school (aged five to ten) are. This difference may reflect the negative impact of poor nutrition on enrollment as it suggests that the more malnourished children are either not enrolled at school (and therefore out of our school sample) or at least more likely to miss school during the lean season (and being absent during the sample selection). This is also suggestive evidence that the analysis based on the school sample is likely to underestimate the overall effect of malnutrition.

The prevalence of short-term malnutrition is also high, meaning that children had experienced a recent deficit in nutrition. A quarter of pupils are thin and 7% suffer from severe thinness (BAZ below three standard deviations). The mean BAZ is -1.39. Similarly to the long-term nutritional status, the proportion of wasted pupils rises with age (going from 18.4% for children aged five to ten to 41.4% for children aged fourteen to fifteen) and is significantly higher for boys (28.2%) than for girls (22.4%). Surprisingly and contrary to the prevalence of stunting, children measured at home are less likely to be wasted than children measured at school.

Long-term and short-term nutritional statuses are correlated: the correlation between HAZ and BAZ is highly significant but the coefficient is low (0.17). Only 10.8% of pupils suffered from both stunting and thinness but nearly 43.0% of thin children were also stunted.

Children nutritional deficiencies are the consequence of food shortages. About 45% of households declared having eaten prickly pears (*raketa* in Malagasy), a cactus fruit which is considered as scarcity food (Kaufmann (2004)), during the last seven days.

1.4.3 Tests scores and nutritional status

Descriptive statistics can give a first insight of the relation between malnutrition and school achievement. Firstly, poor nutrition is negatively associated with enrollment, at least on the upper grades. The prevalence of stunting and thinness increases naturally with age and, as a consequence it increases with grade. In our sample however, the prevalence of malnutrition starts to decrease at fourth (for thinness) and fifth grade (for stunting), while age go on increasing. This finding suggests that students suffering for long or short-term malnutrition are more likely to drop out of school before completing the primary school cycle. Alternatively, this result could reflect the differences between children who attend complete and incomplete schools, the latter being in more remote areas.

Secondly, the link between nutrition and school achievement is not obvious at first sight: there is no clear difference between test scores by grade between stunted (or wasted) and non stunted pupils (figure 1.4). However, a large and significant difference appears once test scores are compared by age (figure 1.5). The average age for grade can explain this pattern: at a given grade, stunted pupils are a year older on average than the other pupils. The difference is the smallest at first grade (0.77 year) and the highest at second grade (1.37). The two main origins of this difference are a delay in enrollment or a higher probability to repeat grade for stunted children. It is hard to disentangle between these two assumptions with the database due to missing information on age at enrollment. Glick et al. (2005) found that in rural areas of Madagascar, nearly one third of children entered school after six. Nevertheless, the repetition rate is higher for stunted children (the difference is significant at second and third grade).

1.5 Empirical strategy

1.5.1 Endogeneity issues and instrumental variables

Knowledge and nutrition are stock outputs that benefit from inputs throughout a child's life. Therefore, the cross-sectional dimension of the database is an important limitation as it lacks direct measurement of the past health and nutrition status of children, a history of household wealth or information on parents' educational inputs. Moreover, the absence of a second round of information prevents us from controlling for children time-invariant unobservable characteristics, like innate ability.

An instrumental variables method is requested to estimate the causal relationship between a child nutritional status and knowledge acquisition, as expressed by equation (1.1).

Reliable instruments must explain a part of the nutritional status variation and should not be correlated with tests scores. Exogenous instruments of this kind are hard to find in a micro database at the individual level. An ideal database would provide information on the household wealth, consumption and agricultural production for every year of a child's life. Such information would help to explain the long-term nutritional status as every food shortage has an impact on child development that could not be compensated by a good harvest the following year. Agricultural production shocks and food shortage during a child preschool period have an impact on child long-term nutrition, if households can't smooth their consumption: Hoddinott and Kinsey (2001) demonstrated how droughts can hamper a child's growth and that catch-up growth is very limited. Agricultural shocks however are unlikely to have a direct effect on children learning if they occur before they are enrolled at school (pre-school period).[12]

Agricultural production records from birth to five years old are not available in this dataset. However, meteorological information such as rainfall or temperatures can be used as a proxy. They are among the main factors of agricultural production, especially in this part of the country where the climate is very dry and irrigation methods almost never used. Under the assumption that households can't entirely cope with them, large enough exogenous rainfall shocks would affect negatively households' agricultural production, consumption and therefore their children's growth.[13] Nutritional shocks are expected to have an impact on child growth until adulthood but this study focuses on the first five years for two main reasons. Firstly, children enter primary school at six ; at school-age, rainfall shocks could have a direct impact on children learning environment. For instance, parents might keep their children at home to use them as labor force for scarcity food gathering. Moreover, small schools used to close if the lean season is to hard or if a cyclone pass through the region. Secondly, the first years of life are decisive for child stature and brain development (Maccini and Yang (2009) ; Jukes et al. (2007)).[14]

This IV strategy is inappropriate to estimate the effect of short-term malnutrition. Thin-

[12]Several instruments were proposed by the literature. In the context of Sri Lanka, Wisniewski (2010) directly used agricultural production shocks at the community level, interacted with parental schooling, as exogenous instruments for the height for age z-score. In a convincing attempt to estimate the causal impact of nutrition on school performance, Alderman et al. (2001) used major commodities price shocks during early childhood to predict the height-for-age ratio at five years old. The identification strategy proposed by Alderman et al. (2006) relies on the specific context of Zimbabwe as it takes the exposure to the civil war and to a large national drought as exogenous instruments. On an adult sample, Behrman et al. (2008) used the exposition to a specific nutrition intervention during the first three years of life and being a twin as exogenous instruments.

[13]A fall in agricultural production following a covariant rainfall shock could lead to higher local cereal prices. A price rise is bad for consumers but would benefit to cereal producers. However, for the main parts of the households in the sample, cereal stocks are never large enough to cover a full year.

[14]It is worth noting that Thai and Falaris (2014) raised concerns about the use of rainfall shocks experienced in early childhood as instruments. They argue that height for age is only one component of a children's health and that unobserved health outcomes like cognitive ability or immune system development are likely to be correlated with the rainfall shocks.

ness is determined by recent food shortage. Rainfall shocks over the previous two years before the survey could be a proxy for agricultural production and predict a part of a child's BAZ. However, shocks are likely to have an impact on education outcomes too, through absenteeism or drop-out.

1.5.2 Meteorological data

The Center for Climatic Research (University of Delaware) provides a very large meteorological dataset (see Legates and Willmott (1990) for a description of the dataset and methodology). It contains worldwide monthly precipitation and air temperature records from 1900 to 2009. Over this period, they use 1600 to 12000 weather stations to make an interpolation in order to get a record every 0.5 degree of latitude and longitude all over the globe. The advantage of this database is the length of the time series that allows calculation of a long-term tendency. The drawback is that series are highly correlated with their nearest neighbor point (by calculation).[15]

In this paper, we call a season a period of twelve months from September to August (the two months with the lowest precipitation). It includes the rainfall season that goes from November to March (precipitation are particularly high in December, January and February), a period that corresponds to the lean season, followed by the main harvest from April to May. There is no off-season harvest in this part of the country.

The climate of the south of the island is very harsh. First of all, precipitations are very low. Over the last fifty years, the average annual total rainfall is 766mm in the two districts whereas the national mean is 1418mm. Secondly, the total amount of precipitation is particularly erratic from one year to another. The average deviation from the long-term rainfall mean is 9% at the national level but reaches 18% in Amboasary and Betioky. During the last fifty years, the difference in total precipitation from a season to another is larger than 20% half the years in the south whereas it occurs only once every five years on average on a national basis. Unsurprisingly, the average annual temperature is slightly higher in Amboasary and Betioky (23.9°C in average during the last ten years) than the rest of the country (23.3°C). The difference in average temperature reaches one degree during the hottest months of the year (December to March).

[15] Since the south of island is the most vulnerable region of Madagascar, it is under close surveillance. Between 1997 and 2007, The Early Warning System (Système d'Alerte Précoce) has collected rainfall and price data in 7 districts (including Amboasary and Betioky). Prices were collected on a monthly basis and precipitation recorded three times a month. Data are available for nearly forty weather stations and seventy-five markets (sixteen stations and twenty-three markets in Amboasary and Betioky). Despite its richness, this dataset wasn't appropriate for this analysis. Firstly, the time period length (ten years) was not sufficient to compute a reliable long-term trend. Secondly, the period doesn't cover early childhood for the oldest children of our sample. Finally, rainfall and prices series contain too much missing data to be reliable. However this rainfall dataset has been used to check the reliability of the CCR data: the correlation between the two datasets is high.

1.5.3 Instrumental variables computation

In the literature, rainfall data are used in different ways to expressed income shocks. In a Thai context, Paxson (1992) divides the year into four seasons, expresses rainfall variation as deviation from the long-term mean and includes the four variables and their squares in income and saving equations. Kazianga (2012), Thai and Falaris (2014) and Maccini and Yang (2009) use yearly rainfall deviation from the long-term mean, along with data from Burkina-Faso, Vietnam and Indonesia. Akresh et al. (2011) also use deviation from the mean, expressed it in millimeters, and rank shocks into quartile dummies.

In the present paper, rainfall shocks are expressed as the variation in the total amount of precipitation from a season to another. A decrease in the total amount of precipitation will lead to a smaller agricultural production from a year to another (Levine and Yang (2014)). If the household can't cope with this cut in the agricultural production (i.e. an unexpected reduction of their permanent income), children's growth could be hampered. Conversely, an increase in the total amount of precipitation will lead to a higher agricultural production and higher consumption, which can potentially lead to a better long-term nutritional status. Over the thirteen seasons between 1991 and 2004, the average variation is 5%. The maximum negative variation is -66%. The maximum positive variation is 133%.

In Amboasary and Betioky, the distance between two weather points is close to 50km. Because neither schools, nor households have been geo-positioned during the survey, children were matched to the nearest weather station on the basis of the centroid of the commune/ZAP[16] they are living in. We assume that every school matched to the same weather point has experienced the same weather conditions and shocks. However, children of various ages experience these shocks at different periods of their lives. In other words, there is variation between and within weather points but no variation for children of the same age under the same weather point. This strategy relies on the assumption that households were living near the same weather point since the child birth at least. Therefore, household migration or child fostering out of the weather point area may bias the analysis. These two phenomenons are likely to be low in rural area such as Amboasary and Betioky. In our household sample, 80% of heads states to be living in their current village for more than fifteen years.[17]

[16] A ZAP (*Zone d'Administration Pédagogique*) is an administrative area that counts ten schools on average. There are twelve ZAP in Amboasary and twenty ZAP in Betioky

[17] The rest of them have moved to their current village recently but their former villages are unknown and may be under the same weather point.

1.6 Results

1.6.1 OLS estimations

Firstly, equation (1.1) is estimated with ordinary least square. Mathematics test score, number of letters read and the number of words read are used as dependent variables to assess pupils' school performance. Results are presented in Tables 1.3, 1.4 and 1.5. Community (ZAP) or school fixed effects are alternatively applied in order to control for unobservable community and school characteristics.[18] The first three columns present results with ZAP fixed effects and school characteristics. The effects of short and long-term indexes are estimated separately and together. The fourth column presents results with school fixed effects. Contrary to the first four columns which are based on the school sample (including pupils from the five grades), the fourth column focus on second grade students for which household's characteristics data are available.

As expected, age and grade are strong predictors of the scores. Consistent with the literature (see Glick and Sahn (2009)), the higher the grade, the higher the score. While this result should be interpreted cautiously as its potential endogeneity has not been controlled for, the substantial size of the effect of grade achievement is worth noting: the mathematics score is 20% higher for students enrolled in second grade rather than in the first one (48% for both reading test scores). For a given grade, older students perform better to the tests. This effect is not linear, each year older being less and less important but the turning point is predicted long after children should have left school. Community characteristics and school infrastructures also impact the acquisition of knowledge. Being in a multigrade classroom has a negative impact on both reading tests scores. Being in an overcrowded classroom is also associated with lower performance but this finding is only significant while we restrict our sample to the students whose parents were interviewed (second grade students). The sensibility of children's performance to school's infrastructures raise concerns about the potential negative side-effects a school feeding program if the surge of new pupils attracted is not offset by an investment in infrastructures.

The long-term nutritional status (HAZ) is associated with better results. The parameter is significant at 1% level for the three scores. All other things held constant, a one point increase in HAZ is associated with a 0.07 standard deviation (SD) higher score on mathematics test (which corresponds to a 1.71% increase according to the average score), 0.04 SD higher number of letters read (3.5%) read and 0.02 SD higher number of words read (2.2%). The size of this relationship is modest in comparison to the effect of grade progression but it should be

[18]The community or school fixed effects will capture the effect of community characteristics that are permanent and common on students' learning but will fail to control for inputs targeted to a specific group within a community (say the allocation of additional teacher time to low-abilities students.

kept in mind that this effect is net of the impact of nutrition on grade progression. It is robust to the inclusion of the short-term nutrition indicator: once the BAZ variable is introduced the HAZ parameter is not altered, suggesting that short and long-term malnutrition (i.e. recent and past episodes of food shortage) have an independent effect on school achievement. A one SD increase in BAZ is associated with an increase in the mathematics score of less than 0.6% of the average score, 2.4% more letters read on average and 3.1% more words read.

Nutrition indicators are also robust to the inclusion of household characteristics. Even after controlling for cattle ownership, education of head, household size and wealth, both short and long-term nutrition remains significant at 1% level. Household's wealth, a proxy for the family pedagogical inputs,[19] is only significant for the two reading tests: living in a wealthier household is associated with better school performance. The lack of effect of the household head's education is surprising but not inconsistent with the literature: a large majority of the household head in our sample are male and, mother's education is regularly reported to be of greater importance than father's education.

As previously noted, OLS may not give consistent estimates of the relationship between the nutritional status. Alternative estimation strategies are considered in the next sections.

1.6.2 Household fixed effects

In a first attempt to control for the unobserved heterogeneity, we take advantage of the household sample to estimate the effect of malnutrition on achievement by controlling for the unobserved household characteristics. In every visited household, the siblings of the students selected at school were also measured and tested for their mathematics knowledge (they were not given the reading test). Over the 1500 second grade pupils of the household sample, 845 had siblings aged five to ten for which tests scores and anthropometrics data are available. This strategy relies on the assumption that the student's innate ability is family specific (Todd and Wolpin (2003)). Estimating the variation within a household would therefore eliminate the effect of the common heterogeneity between siblings.

The mean mathematics test score is 11.24 (SD=6.11). Table 1.6 presents the results of this investigation. Equation (1.1) is estimated using ZAP (first column), school (second column) and household fixed effects (columns third to five). The HAZ parameter is remarkably stable over the various specifications: controlling alternatively for the unobserved heterogeneity at the community, school or household level doesn't affect the size of the parameter, which remains highly significant. The BAZ coefficient decreases a little bit with the household fixed effects and loses significance but remains significant at the 10% level.

[19]Household's wealth is measured by an asset index that was computed with the principal component analysis method on the basis of assets ownership (twenty goods) and housing characteristics (number of rooms, type of floor, walls and roof and access to water).

As expected, being enrolled (especially in a higher grade) is associated with higher performance. The performance of the students selected at school is widely better, conditional with the other covariates. This finding, that can be the consequence of the different way the test was applied at school and in the household, could also be interpreted as a sign of selection into the school sample.

1.6.3 IV estimations

As OLS can't provide consistent estimates of the causal relationship between malnutrition the school achievement, we use an instrumental variable strategy to estimate equation (1.1). Results for the mathematics score and the letters reading score are presented in Tables 1.7 and 1.8.[20] As in the previous sections, results are presented with community and school fixed-effects alternatively. Columns one and three show the first stage equations. Columns two and four show the IV results.

The variation in the total amount of precipitation from a season to another during early childhood has a positive and significant effect on HAZ. The rainfall shock experienced during the fifth year of life was not included in the regression as it appears to be a redundant instrument. On the mathematics test sample, the rainfall parameters are significant for the four first years of life (three years out of four when fixed effects are at the community level).[21] The four instruments are jointly significant in every specification. Depending on the year, considered, an increase by 100% in the total amount of rainfall leads to a 10 to 16% increase in HAZ. Conversely, a decrease in rainfall from a year to another leads to poorer long-term nutritional status. In other words, child growth benefits from the unexpected increase in the household resources that follows a positive variation in the rainfall precipitation and is hampered by unexpected cuts in the resources. The proportion of the explained variation of HAZ is low (the R^2 ranges between 0.13 to 0.16) but the conclusion of this first stage regression is important: rainfall shocks at early childhood have a permanent impact on child growth and this effect is large enough to be detected five to fourteen years later.

A noteworthy feature of our findings from the first stage estimates is the parameters of the grade dummies, which are significantly associated with better HAZ while age shows an expected negative impact. This finding strengthens the assumption that students with better nutritional status progresses better through grades and/or that the weakest children leave school before completing primary school.

The endogeneity of the long-term nutritional status is confirmed by the endogenity test on every specification. Once instrumented, the effect of HAZ on the mathematics score is larger than the OLS estimates, suggesting that the parameter was underestimated. A one

[20]Results for the words reading test, which are similar to the letters reading test, are not presented here.
[21]Two to three instruments out of four are significant in first stage regression on the letters reading test.

point increase in HAZ is associated with an 8.8% increase in the mathematics score while the OLS estimates exhibit a 1.7% increase only. This finding is consistent with Alderman et al. (2001) and Alderman et al. (2006) who reported substantially higher impact of preschool nutrition on children enrollment and grade attainment when their model is estimated with an IV method rather than with OLS. The model is not underidentified (Kleibergen-Paap LM-test) and the instruments are valid according to the Hansen J-test statistics. However, estimates may suffer from a weak identification as the F-statistics for excluded instruments is under the threshold of ten in all specifications.

Alternative specifications are presented in Table 1.9. Results from a limited-information maximum likelihood estimation, a recommended estimator robust in the presence of weak instruments, are presented in the first column: neither the size of the parameter nor its significance is affected when the method is applied. The second and third columns present the results when the sample is restricted to the first two grades (second column) and last three grades (third column). Malnutrition is more detrimental during the first schooling years: the effect of malnutrition on the acquisition on the basic mathematics knowledge is larger among students enrolled in lower grades (+13%) but insignificant for those who were enrolled in higher grades. In order to rule out the sample selection issue between the students who did not take the reading tests, we restrict the sample to the students who were given the reading test and find similar results (fourth column). Finally, in the fifth column, we switch the height for age z-score to a simple dummy variable for stunted children (HAZ<-2). We find that being stunted reduce the average mathematics score by nearly one standard deviation, which corresponds to a -25% decrease.

The IV method is less convincing with the reading tests (Table 1.8). The results also show a larger estimated effect of the long-term malnutrition but the coefficient is no longer significant. This result may be the consequence of the attrition bias within the upper grades. Learning to read is a long process and only students above the third grade are able to read correctly, which means in our data that the variation in the reading tests scores is limited in the lower grades. In the same time, following results on the mathematics score, the malnutrition is larger in lower grades. As the more malnourished children have dropped out of school or have never reached the higher grades, the variation in the nutritional status is more limited within the grades where the variation in literacy is higher. If this attrition bias is the consequence of early childhood rainfall shocks, the IV strategy may not be effective.

1.7 Discussion and conclusion

In this paper, we investigated the relationship between malnutrition and school achievement among a school based sample of primary school pupils. In order to deal with the unobserved heterogeneity (like children's innate ability or parent's preferences) that could bias our estimates of this relationship, we applied a fixed-effect method at the household level to purge the results from the family specific component of a child's innate ability. Subsequently, we used rainfall shocks experienced during early childhood as exogenous instruments in an instrumental variables strategy. This later (and preferred) specification allows us to estimate the causal effect of long-term malnutrition on school performance.

Our analysis present suggestive evidence that, besides its indirect effect on knowledge acquisition through grade progression, long-term malnutrition has a direct negative effect on learning. For a given grade and age, stunted children had worse performance to mathematics and reading tests. While the estimated contribution of nutrition to the production of cognitive skills is modest according to the OLS estimates, our results suggest that this effect was underestimated. Food shortage, resulting from an unexpected reduction in the agricultural production, experienced during the first years of a child's life, could have a long-lasting effect on its learning capacity at school age. Short-term malnutrition, an indicator of recent episodes of food shortage is also associated with lower cognitive skills but its causal effect couldn't be estimated with our IV strategy.

While we couldn't rule out the importance of short-term nutrition (especially on attendance), the large effect of stunting suggests that school based health interventions like school feeding programs might come to late in the life of a child. Considering the long-lasting effect of income shocks experienced during the first years of life on human development, helping households to cope with those shocks may have substantial positive impacts on both nutrition and school achievement. As showed by Schroeder et al. (1995), supplementary feeding is particularly efficient to improve child's growth during the first three years of life and, according to Maluccio et al. (2009), it can have long-lasting effects on school achievement.

Bibliography

AKRESH, R., VERWIMP, P., AND BUNDERVOET, T. 2011. Civil war, crop failure, and child stunting in Rwanda. *Economic Development and Cultural Change* 59:777–810.

ALDERMAN, H., BEHRMAN, J., LAVY, V., AND MENON, R. 2001. Child health and school enrollment: A longitudinal analysis. *Journal of Human Resources* 26:185–205.

ALDERMAN, H., GILLIGAN, D. O., AND LEHRER, K. 2012. The impact of food for education programs on school participation in northern Uganda. *Economic Development and Cultural Change* 61:187–218.

ALDERMAN, H., HODDINOTT, J., AND KINSEY, B. 2006. Long term consequences of early childhood malnutrition. *Oxford Economic Papers* 58:450–474.

BEHRMAN, J. 1996. The impact of health and nutrition on education. *The World Bank Research Observer* 11:23–37.

BEHRMAN, J. R., HODDINOTT, J. F., MALUCCIO, J. A., SOLER-HAMPEJSEK, E., BEHRMAN, E. L., MARTORELL, R., RAMIREZ-ZEA, M., AND STEIN, A. D. 2008. What determines adult cognitive skills? Impacts of pre-schooling, schooling and post-schooling experiences in Guatemala. IFPRI Discussion Paper 826, International Food Policy Research Institute.

BEHRMAN, J. R. AND LAVY, V. 1994. Children's health and achievement in school. Living Standards Measurement Study Working Paper No. 104.

BUTTENHEIM, A., ALDERMAN, H., AND FRIEDMAN, J. 2011. Impact evaluation of school feeding programs in Lao PDR. Policy Research Working Paper Series 5518, The World Bank.

FRIEDMAN, J., PHILLIPS-HOWARD, P., MIREL, L., TERLOUW, D., OKELLO, N., VULULE, J., HAWLEY, W., NAHLEN, B., AND TER KUILE, F. 2005. Progression of stunting and its predictors among school-aged children in western Kenya. *European Journal of Clinical Nutrition* 59:914–922.

Galloway, R., Kristjansson, E., Gelli, A., Meir, U., Espejo, F., and Bundy, D. 2009. School feeding: Outcomes and costs. *Food and Nutrition Bulletin* 30:171–182.

Glewwe, P. and Miguel, E. 2007. The impact of child health and nutrition on education in less developed countries. *Handbook of Development Economics* 4:3561–3606.

Glick, P., Rajemison, H., Ravelo, A., Raveloarison, Y., Razakamanantsoa, M., and Sahn, D. 2005. The progression through school and academic performance in Madagascar study: Preliminary descriptive results. *Cornell Food and Nutrition Policy Program Working Paper 166, Cornell University, Ithaca, NY* .

Glick, P. and Sahn, D. E. 2009. Cognitive skills among children in senegal: Disentangling the roles of schooling and family background. *Economics of Education Review* 28:178–188.

Grantham-McGregor, S., Chang, S., and Walker, S. 1998. Evaluation of school feeding programs: Some Jamaican examples. *The American Journal of Clinical Nutrition* 67:785S–789S.

Gubert, F. and Robilliard, A.-S. 2008. Risk and schooling decisions in rural Madagascar: A panel data-analysis. *Journal of African Economies* 17:207–238.

Hoddinott, J. and Kinsey, B. 2001. Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics* 63:409–436.

Jacoby, H. 2002. Is there an intrahousehold 'flypaper effect'? Evidence from a school feeding program. *The Economic Journal* 112:196–221.

Jukes, M., Drake, L., and Bundy, D. 2007. School health, nutrition and education for all: Levelling the playing field. CABI Publishing, Wallingford.

Kaufmann, J. 2004. Prickly pear cactus and pastoralism in southwest Madagascar. *Ethnology* 43:345–361.

Kazianga, H. 2012. Income risk and household schooling decisions in Burkina Faso. *World Development* 40:1647–1662.

Kazianga, H., de Walque, D., and Alderman, H. 2009. Educational and health impacts of two school feeding schemes: evidence from a randomized trial in rural Burkina Faso. Policy Research Working Paper Series 4976, The World Bank.

Legates, D. and Willmott, C. 1990. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology* 10:111–127.

Levine, D. I. and Yang, D. 2014. The impact of rainfall on rice output in indonesia. Technical report, National Bureau of Economic Research.

Lwambo, N., Brooker, S., Siza, J., Bundy, D., and Guyatt, H. 2000. Age patterns in stunting and anaemia in African schoolchildren: a cross-sectional study in Tanzania. *European Journal of Clinical Nutrition* 54:36–40.

Maccini, S. and Yang, D. 2009. Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review* 99:1006–26.

Maluccio, J., Hoddinott, J., Behrman, J., Martorell, R., Quisumbing, A., and Stein, A. 2009. The impact of improving nutrition during early childhood on education among Guatemalan adults. *The Economic Journal* 119:734–763.

Martorell, R. 1999. The nature of child malnutrition and its long-term implications. *Food and Nutrition Bulletin-United Nations University Press* 20:288–292.

Michaelowa, K. 2001. Primary education quality in francophone Sub-Saharan Africa: Determinants of learning achievement and efficiency considerations. *World Development* 29:1699–1716.

Paxson, C. 1992. Using weather variability to estimate the response of savings to transitory income in Thailand. *American Economic Review* 82:15–33.

Schroeder, D., Martorell, R., Rivera, J., Ruel, M., and Habicht, J. 1995. Age differences in the impact of nutritional supplementation on growth. *The Journal of Nutrition* 125:1051S–1059S.

Thai, T. Q. and Falaris, E. M. 2014. Child schooling, child health, and rainfall shocks: Evidence from rural Vietnam. *Journal of Development Studies* pp. 1–13.

Todd, P. E. and Wolpin, K. I. 2003. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal* 113:F3–F33.

Wisniewski, S. 2010. Child nutrition, health problems, and school achievement in Sri Lanka. *World Development* 38:315–332.

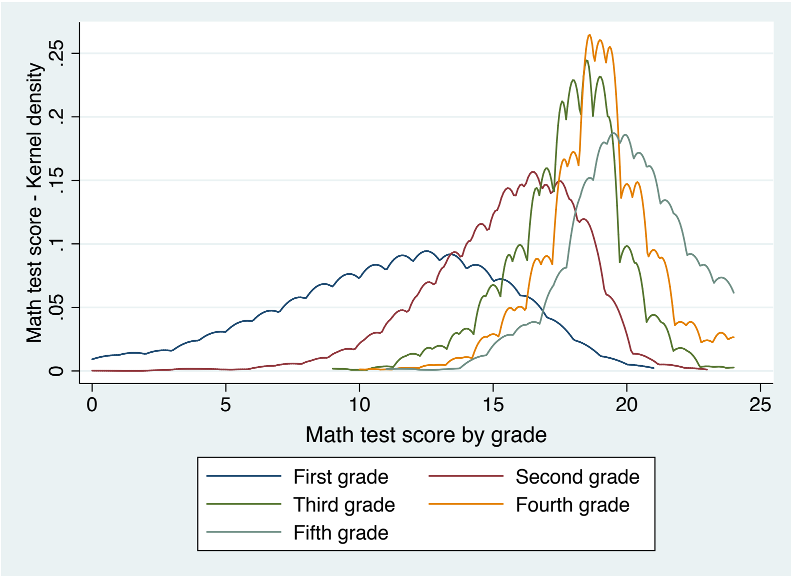

Figure 1.1: Distribution of mathematics test score by grade

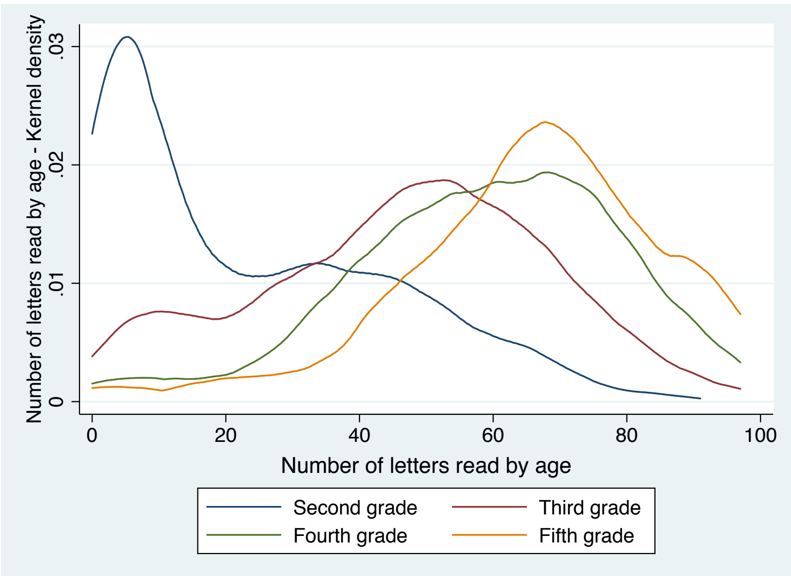

Figure 1.2: Distribution of the number of letters read by grade

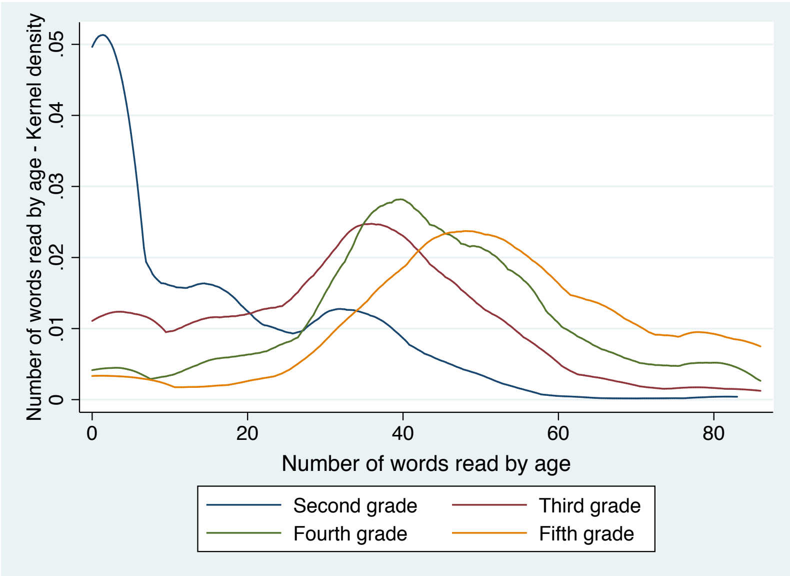

Figure 1.3: Distribution of the number of words read by grade

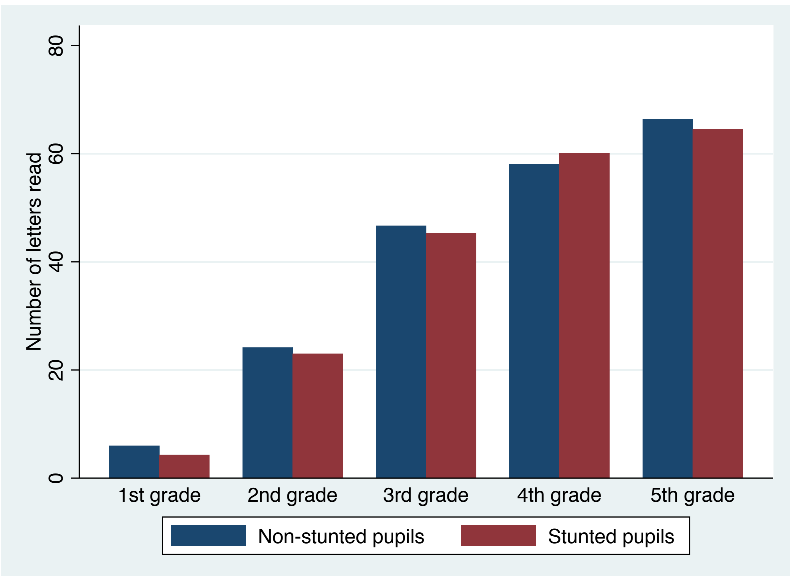

Figure 1.4: Number of letters read by nutritional status and grade

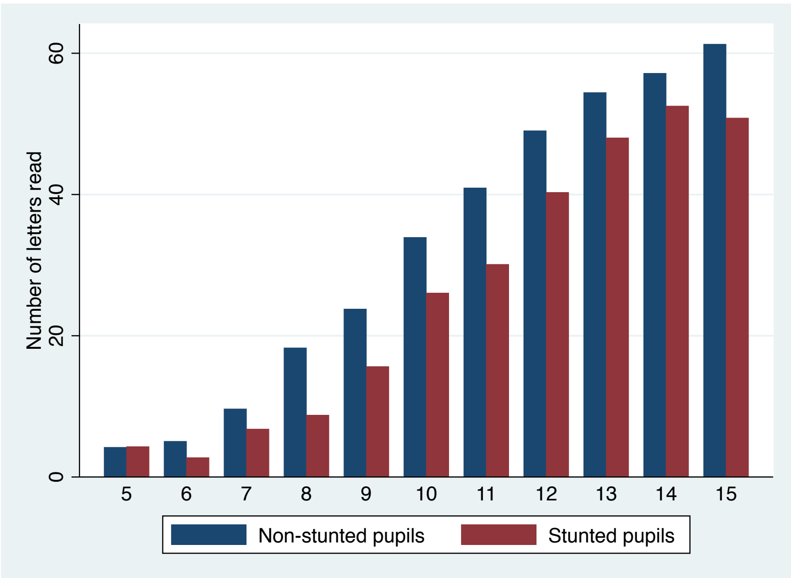

Figure 1.5: Number of letters read by nutritional status and age

Table 1.1: List of variables

Variables	Description
Mathematics test score	Child's performance to an oral mathematics test. The minimum possible score is 0 and the maximum 24.
Number of letters read	Child's performance to an oral reading test. The score is the total number of letters a pupil has correctly read within a minute.
Number of words read	Child's performance to an oral reading test. The score is the total number of words a pupil has correctly read within a minute. The text was in Malagasy.
Height for age z-score	Height for age z-score. Extreme values beyond -6 and over 6 have been dropped.
BMI for age z-score	Body mass index for age z-score. Extreme values beyond -5 and over 5 have been dropped.
Age in years	
child's age in years Child's gender (girl=1)	Dummy variable for gender. Equals 1 if the child is a girl and 0 otherwise.
CP1 (first grade=1)	Dummy variable for first grade enrollment. Equals 1 if the child is enrolled in first grade, 0 otherwise.
CP2 (second grade=1)	Dummy variable for second grade enrollment. Equals 1 if the child is enrolled in second grade, 0 otherwise.
CE (third grade=1)	Dummy variable for third grade enrollment. Equals 1 if the child is enrolled in third grade, 0 otherwise.
CM1 (fourth grade=1)	Dummy variable for fourth grade enrollment. Equals 1 if the child is enrolled in fourth grade, 0 otherwise.
CM2 (fifth grade=1)	Dummy variable for fifth grade enrollment. Equals 1 if the child is enrolled in fifth grade, 0 otherwise.
Multigrade classroom	Dummy variable for multigrade classroom. Equals 1 if the pupil is enrolled in a multigrade classroom, and 0 otherwise.
Teachers speak french	Proportion of french speaking teachers in school.
Teachers' age	Average age of teachers in school.
Private school	Dummy variable for school type. Equals 1 if school is a private one.
Pupils per teacher ratio	Average number of pupils per teacher in school.
Main town in the CISCO	Dummy variable that equals 1 if the village is the main town in the CISCO.
Main village in the municipality	Dummy variable that equals 1 if the village is the main town in the municipality.
Taxibrousse in the village	Dummy variable that equals 1 if the village is served by a taxibrousse.
Number of adults	Total number of adults (>15 years old) in household.
Number of children	Total number of children (≤15 years old) in household.
Cattle size	Total number of cows in the household cattle.
Household's Wealth Index	Wealth index computed with the help of a principal component analysis procedure on assets
Education of household head in years	Household's head highest level of education (no matter if he is the father of the pupil).

Table 1.2: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Mathematics test score	15.73	4.18	0	24	6439
Letters reading test score	31.38	27.68	0	97	5995
Words reading test score	21.42	22.65	0	86	6095
Age in years	10.21	2.48	5	17	6526
Child's gender (girl=1)	0.55	0.5	0	1	6475
Height for age z-score	-1.34	1.41	-5.92	5.46	6314
Stunting (=1 if HAZ< -2)	0.32	0.47	0	1	6314
Severe stunting (=1 if HAZ< -3)	0.1	0.3	0	1	6314
BMI for age z-score	-1.36	1.07	-4.99	4.98	6252
Thinness (=1 if BAZ< -2)	0.25	0.43	0	1	6252
Severe thinness (=1 if BAZ< -3)	0.07	0.26	0	1	6252
Weight for age z-score	-1.34	1.19	-5.57	4.78	2801
Wasting (=1 if WAZ< -2)	0.28	0.45	0	1	2801
Severe underweight (=1 if WAZ< -3)	0.08	0.27	0	1	2801
CP1 (first grade=1)	0.22	0.42	0	1	6526
CP2 (second grade=1)	0.39	0.49	0	1	6526
CE (third grade=1)	0.18	0.38	0	1	6526
CM1 (fourth grade=1)	0.13	0.33	0	1	6526
CM2 (fifth grade=1)	0.09	0.29	0	1	6526
Student is in a multigrade classroom	0.57	0.5	0	1	6473
Private school	0.11	0.32	0	1	6526
Taxibrousse in the village	0.27	0.44	0	1	6513
District dummy variable (Amboasary=1)	0.45	0.5	0	1	6526
Household size	7.06	2.54	2	14	1553
Education of household head in years	2.74	3.6	0	13	1543
Household head gender (female=1)	0.15	0.36	0	1	1553

Table 1.3: OLS regressions - Mathematics test score on nutrition

	(1) Math score	(2) Math score	(3) Math score	(4) Math score	(5) Math score
Height for age z-score	0.066***		0.064***	0.053***	0.058***
	(0.010)		(0.010)	(0.008)	(0.014)
BMI for age z-score		0.033***	0.024***	0.014	0.057***
		(0.009)	(0.008)	(0.009)	(0.018)
Age in years	0.446***	0.425***	0.446***	0.424***	0.240***
	(0.056)	(0.060)	(0.056)	(0.046)	(0.084)
Age in years^2	−0.017***	−0.016***	−0.016***	−0.016***	−0.007*
	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)
Child's gender (girl=1)	−0.009	−0.009	−0.012	−0.014	−0.019
	(0.015)	(0.015)	(0.015)	(0.015)	(0.026)
CP2 (second grade=1)	0.780***	0.806***	0.779***	0.799***	
	(0.057)	(0.059)	(0.057)	(0.043)	
CE (third grade=1)	1.164***	1.210***	1.163***	1.199***	
	(0.055)	(0.057)	(0.055)	(0.052)	
CM1 (fourth grade=1)	1.378***	1.433***	1.374***	1.421***	
	(0.056)	(0.059)	(0.056)	(0.061)	
CM2 (fifth grade=1)	1.600***	1.674***	1.595***	1.648***	
	(0.061)	(0.064)	(0.062)	(0.067)	
Multigrade classroom	−0.043	−0.042	−0.042	−0.063	−0.036
	(0.046)	(0.045)	(0.046)	(0.049)	(0.080)
Teachers speak French	0.098*	0.101*	0.098*		0.174**
	(0.051)	(0.053)	(0.051)		(0.080)
Teachers' age	−0.001	−0.002	−0.001		−0.006
	(0.004)	(0.004)	(0.004)		(0.004)
Private school	0.077	0.073	0.074		−0.048
	(0.050)	(0.051)	(0.051)		(0.072)
Pupils-per-teacher ratio	−0.002	−0.001	−0.002		−0.001
	(0.001)	(0.001)	(0.001)		(0.001)
Main town in the CISCO	0.229***	0.280***	0.227***		0.288***
	(0.032)	(0.038)	(0.034)		(0.056)
Main village in the commune	0.041	0.043	0.042		0.048
	(0.044)	(0.044)	(0.044)		(0.068)
Taxibrousse in the village	−0.033	−0.034	−0.033		−0.035
	(0.039)	(0.039)	(0.039)		(0.064)
Number of adults					−0.010
					(0.013)
Number of children					0.004
					(0.008)
Cattle size					−0.001
					(0.001)
Household's wealth index					−0.006
					(0.026)
Education of household head					0.005
					(0.005)
Constant	−3.354***	−3.245***	−3.331***	−3.295***	−1.280**
	(0.304)	(0.320)	(0.304)	(0.234)	(0.497)
N	5767	5767	5767	5767	1383
R^2	0.569	0.563	0.570	0.613	0.092
F-stat	317.158	297.240	338.590	163.856	8.875
Fixed effects	ZAP	ZAP	ZAP	Schools	ZAP
Number of groups	34	34	34	288	34

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 1.4: OLS regressions - Letters reading test score on nutrition

	(1) Letters read	(2) Letters read	(3) Letters read	(4) Letters read	(5) Letters read
Height for age z-score	0.040***		0.037***	0.024***	0.063***
	(0.011)		(0.011)	(0.007)	(0.016)
BMI for age z-score		0.032***	0.027***	0.018**	0.052**
		(0.009)	(0.009)	(0.008)	(0.020)
Age in years	0.113***	0.100***	0.112***	0.062*	0.243**
	(0.036)	(0.033)	(0.036)	(0.036)	(0.101)
Age in years^2	−0.003*	−0.003*	−0.003*	−0.001	−0.007
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Child's gender (girl=1)	0.031**	0.030*	0.028*	0.016	0.066*
	(0.015)	(0.016)	(0.016)	(0.018)	(0.038)
CP2 (second grade=1)	0.548***	0.564***	0.547***	0.566***	
	(0.047)	(0.047)	(0.047)	(0.034)	
CE (third grade=1)	1.298***	1.326***	1.297***	1.355***	
	(0.057)	(0.057)	(0.057)	(0.045)	
CM1 (fourth grade=1)	1.709***	1.741***	1.704***	1.772***	
	(0.065)	(0.065)	(0.064)	(0.052)	
CM2 (fifth grade=1)	1.846***	1.886***	1.838***	1.932***	
	(0.059)	(0.056)	(0.059)	(0.059)	
Multigrade classroom	−0.081*	−0.080*	−0.081*	−0.007	−0.045
	(0.045)	(0.045)	(0.045)	(0.062)	(0.074)
Teachers speak French	−0.036	−0.034	−0.036		0.023
	(0.038)	(0.038)	(0.038)		(0.072)
Teacher's age	−0.004	−0.004	−0.004		−0.007
	(0.003)	(0.003)	(0.003)		(0.005)
Private school	0.134*	0.130	0.130		0.161
	(0.079)	(0.078)	(0.078)		(0.105)
Pupils-per-teacher ratio	−0.001	−0.001	−0.001		−0.004***
	(0.001)	(0.001)	(0.001)		(0.001)
Main town in the CISCO	0.089	0.118	0.087		0.121
	(0.087)	(0.094)	(0.090)		(0.081)
Main village in the commune	0.083	0.087	0.085		0.074
	(0.075)	(0.075)	(0.074)		(0.118)
Taxibrousse in the village	0.052	0.051	0.051		−0.005
	(0.043)	(0.042)	(0.043)		(0.055)
Number of adults					−0.011
					(0.013)
Number of children					0.008
					(0.010)
Cattle size					−0.001
					(0.001)
Household's wealth index					0.090***
					(0.027)
Education of household head					0.002
					(0.006)
Constant	−1.439***	−1.360***	−1.413***	−1.369***	−1.492**
	(0.210)	(0.199)	(0.208)	(0.175)	(0.552)
N	5296	5296	5296	5296	1300
R^2	0.603	0.601	0.603	0.629	0.104
F-stat	327.701	335.786	334.542	307.208	13.743
Fixed effects	ZAP	ZAP	ZAP	Schools	ZAP
Number of groups	34	34	34	277	34

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 1.5: OLS regressions - Words reading test score on nutrition

	(1) Words read	(2) Words read	(3) Words read	(4) Words read	(5) Words read
Height for age z-score	0.021**		0.018**	0.013	0.041***
	(0.009)		(0.009)	(0.008)	(0.013)
BMI for age z-score		0.032***	0.029***	0.014	0.057***
		(0.010)	(0.009)	(0.009)	(0.016)
Age in years	0.031	0.024	0.030	−0.027	0.173*
	(0.036)	(0.033)	(0.036)	(0.039)	(0.099)
Age in years^2	0.000	0.000	0.000	0.002	−0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Child's gender (girl=1)	0.077***	0.073***	0.073***	0.057***	0.101***
	(0.020)	(0.021)	(0.020)	(0.018)	(0.032)
CP2 (second grade=1)	0.453***	0.461***	0.453***	0.484***	
	(0.053)	(0.051)	(0.053)	(0.037)	
CE (third grade=1)	1.215***	1.227***	1.213***	1.266***	
	(0.067)	(0.064)	(0.067)	(0.051)	
CM1 (fourth grade=1)	1.658***	1.671***	1.653***	1.702***	
	(0.062)	(0.059)	(0.062)	(0.056)	
CM2 (fifth grade=1)	1.936***	1.950***	1.927***	1.975***	
	(0.077)	(0.073)	(0.077)	(0.071)	
Multigrade classroom	−0.085**	−0.084*	−0.084**	−0.019	−0.062
	(0.041)	(0.041)	(0.041)	(0.063)	(0.060)
Teachers speak French	−0.048	−0.047	−0.048		0.006
	(0.049)	(0.049)	(0.049)		(0.075)
Teacher's age	−0.007*	−0.007*	−0.007*		−0.011*
	(0.004)	(0.004)	(0.004)		(0.005)
Private school	0.128	0.124	0.124		0.149
	(0.086)	(0.085)	(0.085)		(0.095)
Pupils-per-teacher ratio	0.001	0.001	0.001		−0.002*
	(0.001)	(0.001)	(0.001)		(0.001)
Main town in the CISCO	0.211**	0.225**	0.210*		0.381**
	(0.101)	(0.109)	(0.107)		(0.171)
Main village in the commune	0.141*	0.144*	0.143*		0.171
	(0.081)	(0.081)	(0.080)		(0.102)
Taxibrousse in the village	0.023	0.022	0.023		−0.106
	(0.057)	(0.057)	(0.057)		(0.063)
Number of adults					0.004
					(0.012)
Number of children					0.004
					(0.011)
Cattle size					−0.001
					(0.001)
Household's wealth index					0.089**
					(0.034)
Education of household head					0.009
					(0.007)
Constant	−0.879***	−0.825***	−0.851***	−0.796***	−1.102*
	(0.226)	(0.215)	(0.223)	(0.191)	(0.594)
N	5308	5308	5308	5308	1300
R^2	0.544	0.545	0.545	0.555	0.117
F-stat	303.455	263.358	296.056	224.966	30.544
Fixed effects	ZAP	ZAP	ZAP	Schools	ZAP
Number of groups	34	34	34	277	34

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 1.6: OLS regressions - Mathematics score on nutrition (household sample)

	(1) Math score	(2) Math score	(3) Math score	(4) Math score	(5) Math score
Height for age z-score	0.044***	0.041***	0.045***	0.044***	
	(0.014)	(0.014)	(0.016)	(0.016)	
BMI for age z-score	0.038***	0.024*	0.027*		0.025
	(0.014)	(0.014)	(0.016)		(0.017)
Age in years	0.535***	0.558***	0.553***	0.553***	0.544***
	(0.055)	(0.066)	(0.078)	(0.079)	(0.079)
Age in years^2	−0.023***	−0.024***	−0.023***	−0.023***	−0.023***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Child's gender (girl=1)	0.044	0.055*	0.031	0.036	0.032
	(0.030)	(0.029)	(0.037)	(0.036)	(0.037)
Currently enrolled at school	0.425***	0.423***	0.466***	0.460***	0.466***
	(0.087)	(0.096)	(0.084)	(0.085)	(0.086)
Highest grade attained	0.301***	0.296***	0.272***	0.275***	0.277***
	(0.053)	(0.065)	(0.057)	(0.058)	(0.058)
School sampled child	0.577***	0.573***	0.560***	0.554***	0.568***
	(0.064)	(0.060)	(0.052)	(0.053)	(0.054)
Number of adults	−0.003	0.001			
	(0.011)	(0.012)			
Number of children	0.008	0.005			
	(0.009)	(0.010)			
Cattle size	−0.004***	−0.002***			
	(0.001)	(0.001)			
Household's wealth index	0.048*	0.025			
	(0.027)	(0.035)			
Education of household head in years	0.002	0.003			
	(0.005)	(0.006)			
Teachers speak French	0.125				
	(0.079)				
Teachers' age	0.002				
	(0.004)				
Private school	0.040				
	(0.055)				
Pupils-per-teacher ratio	−0.002				
	(0.001)				
Main town in the CISCO	0.078				
	(0.199)				
Main village in the commune	0.068				
	(0.068)				
Taxibrousse in the village	−0.034				
	(0.056)				
Constant	−3.857***	−3.964***	−3.965***	−3.988***	−3.977***
	(0.236)	(0.309)	(0.338)	(0.340)	(0.342)
N	1829	1829	1829	1829	1829
R^2	0.605	0.651	0.711	0.710	0.709
F	211.737	156.652	306.879	348.261	346.003
Fixed effects	ZAP	Schools	Households	Households	Households

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 1.7: IV regressions - Mathematics score on nutrition

	(1) First stage	(2) IV	(3) First stage	(4) IV
Height for age z-score		0.333**		0.303*
		(0.162)		(0.156)
Rainfall variation (year 1)	0.133*		0.136**	
	(0.070)		(0.066)	
Rainfall variation (year 2)	0.201***		0.167***	
	(0.074)		(0.055)	
Rainfall variation (year 3)	0.113*		0.061	
	(0.063)		(0.066)	
Rainfall variation (year 4)	0.165***		0.118**	
	(0.047)		(0.047)	
Age in years	−0.315***	0.525***	−0.264**	0.517***
	(0.090)	(0.085)	(0.128)	(0.085)
Age in years^2	0.002	−0.017***	−0.001	−0.017***
	(0.004)	(0.003)	(0.006)	(0.002)
Child's gender (girl=1)	0.084**	−0.039*	0.080**	−0.030
	(0.033)	(0.022)	(0.030)	(0.020)
CP2 (second grade=1)	0.398***	0.696***	0.415***	0.690***
	(0.060)	(0.084)	(0.082)	(0.094)
CE (third grade=1)	0.765***	0.996***	0.746***	1.000***
	(0.076)	(0.140)	(0.098)	(0.143)
CM1 (fourth grade=1)	0.940***	1.174***	0.951***	1.167***
	(0.084)	(0.170)	(0.099)	(0.176)
CM2 (fifth grade=1)	1.265***	1.296***	1.311***	1.291***
	(0.101)	(0.219)	(0.101)	(0.226)
Multigrade classroom	0.018	−0.077	0.007	−0.050
	(0.074)	(0.055)	(0.058)	(0.049)
Teachers speak French			0.046	0.091*
			(0.119)	(0.053)
Teacher's age			−0.006	−0.000
			(0.005)	(0.003)
Private school			−0.007	0.079
			(0.076)	(0.054)
Pupils-per-teacher ratio			0.005***	−0.003
			(0.002)	(0.002)
Main town in the CISCO			0.817***	0.039
			(0.171)	(0.151)
Main village in the commune			0.001	0.034
			(0.111)	(0.052)
Taxibrousse in the village			−0.010	−0.031
			(0.063)	(0.044)
N	5745	5745	5745	5745
R^2	0.131	0.490	0.159	0.479
F-stat	38.492	171.615	101.338	302.663
F-stat (excluded instruments)		3.838		5.101
Underidentification test (p-value)		0.005		0.021
Hansen J statistic (p-value)		0.945		0.880
Fixed effects	Schools	Schools	ZAP	ZAP
Number of groups	284	284	33	33

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 1.8: IV regressions - Number of letters read on nutrition

	(1) First stage	(2) IV	(3) First stage	(4) IV
Height for age z-score		0.075		0.090
		(0.150)		(0.136)
Rainfall variation (year 1)	0.079		0.131*	
	(0.077)		(0.071)	
Rainfall variation (year 2)	0.164**		0.173***	
	(0.083)		(0.062)	
Rainfall variation (year 3)	0.064		0.044	
	(0.064)		(0.076)	
Rainfall variation (year 4)	0.124**		0.106*	
	(0.052)		(0.053)	
Age in years	−0.341***	0.082	−0.278**	0.132**
	(0.095)	(0.066)	(0.123)	(0.059)
Age in years^2	0.003	−0.001	−0.001	−0.003*
	(0.004)	(0.002)	(0.006)	(0.002)
Child's gender (girl=1)	0.085**	0.013	0.073**	0.026
	(0.035)	(0.022)	(0.031)	(0.017)
CP2 (second grade=1)	0.393***	0.543***	0.426***	0.521***
	(0.067)	(0.068)	(0.076)	(0.083)
CE (third grade=1)	0.791***	1.306***	0.768***	1.247***
	(0.082)	(0.125)	(0.088)	(0.122)
CM1 (fourth grade=1)	0.969***	1.714***	0.980***	1.647***
	(0.091)	(0.157)	(0.089)	(0.143)
CM2 (fifth grade=1)	1.316***	1.866***	1.334***	1.775***
	(0.112)	(0.213)	(0.100)	(0.190)
Multigrade classroom	0.073	−0.009	0.044	−0.077
	(0.078)	(0.063)	(0.063)	(0.047)
Teachers speak French			0.064	−0.052
			(0.126)	(0.039)
Teacher's age			−0.006	−0.004
			(0.005)	(0.003)
Private school			0.001	0.139*
			(0.074)	(0.076)
Pupils-per-teacher ratio			0.004**	−0.001
			(0.002)	(0.001)
Main town in the CISCO			0.824***	0.050
			(0.178)	(0.153)
Main village in the commune			0.042	0.086
			(0.095)	(0.075)
Taxibrousse in the village			0.000	0.050
			(0.061)	(0.044)
N	5261	5261	5261	5261
R^2	0.133	0.620	0.161	0.595
F-stat	37.026	340.306	118.816	327.425
F-stat (excluded instruments)		2.251		4.209
Underidentification test (p-value)		0.065		0.018
Hansen J statistic (p-value)		0.691		0.739
Fixed effects	Schools	Schools	ZAP	ZAP
Number of groups	273	273	33	33

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 1.9: IV regressions - Mathematics test score on nutrition (alternative specifications)

	(1) Math score	(2) Math score	(3) Math score	(4) Math score	(5) Math score
Height for age z-score	0.342**	0.485**	−0.962	0.384**	
	(0.168)	(0.224)	(2.282)	(0.176)	
Stunting dummy					−0.948*
					(0.549)
Age in years	0.528***	0.640***	−0.879	0.550***	0.454***
	(0.087)	(0.104)	(2.069)	(0.096)	(0.063)
Age in years^2	−0.017***	−0.020***	0.027	−0.017***	−0.014***
	(0.003)	(0.004)	(0.062)	(0.003)	(0.002)
Child's gender (girl=1)	−0.040*	−0.040	0.031	−0.044*	−0.040*
	(0.022)	(0.034)	(0.193)	(0.026)	(0.024)
CP2 (second grade=1)	0.692***	0.620***		0.707***	0.740***
	(0.086)	(0.098)		(0.095)	(0.071)
CE (third grade=1)	0.989***			0.950***	1.074***
	(0.145)			(0.164)	(0.121)
CM1 (fourth grade=1)	1.166***		0.515	1.136***	1.248***
	(0.176)		(0.587)	(0.192)	(0.159)
CM2 (fifth grade=1)	1.285***		1.063	1.258***	1.388***
	(0.226)		(1.347)	(0.248)	(0.207)
Multigrade classroom	−0.077	−0.067	−0.239	−0.095	−0.069
	(0.055)	(0.159)	(0.385)	(0.063)	(0.055)
N	5745	3544	2196	5087	5745
R^2	0.482	0.139	−5.616	0.441	0.454
F-stat	170.774	115.378	7.872	153.316	164.630
F-stat (excluded instruments)	3.838	2.780	1.476	3.456	3.475
Underidentification test (p-value)	0.005	0.028	0.227	0.009	0.014
Hansen J statistic (p-value)	0.945	0.783	0.734	0.971	0.676
Fixed effects	Schools	Schools	Schools	Schools	Schools
Number of groups	284	283	230	273	284

Robust standard errors in parentheses are clustered by communities (ZAP) or by schools according to the fixed effects. * significant at 10%, ** significant at 5%, *** significant at 1%.

Chapter 2

Food distribution cycle, consumption and malnutrition in refugee camps

This chapter was written in collaboration with Marie-Charlotte Buisson.[1]

2.1 Introduction

In 2012, Sub-Saharan Africa hosted 2.667 millions of refugees (UNHCR (2013)). While humanitarian crises leading to migration and resettlement are usually given large media coverage, the long process of being a refugee is most of the time hidden. Indeed, being a refugee is a long-term situation. From an economic perspective, the initial migration shock implies for a refugee to lost its home, lands and assets as well as its livelihoods and sources of income. From a social perspective, the migration may also induce a new structure for the household and is always followed by the loss of the formal and informal social landmarks. Fleeing from violence and seeking shelter away initiate the migration and the resettlement that will be the starting point of an economic and social vulnerability for the household, the duration of this period being completely uncertain.

The official recognition of the refugee status gives access to the United Nations (UN) agencies, as the United Nations High Commissioner for Refugees (UNHCR) and the World Food Program (WFP), and relief organizations that grant protection and assistance to mitigate vulnerability. Settlement in a camp, intended for ensuring safety, facilitating the aid distribution and avoiding uncontrolled migration to the host country, is usually the first step of this assistance. The camp location is decided in an emergency context in coordination between the host country authorities and the UN agencies. This choice, guided by political, security and logistical reasons, has important consequences for the social and economic life

[1] Researcher at IWMI - M.Buisson@cgiar.org

of the refugees. Half the African refugees are currently living in camps based in rural areas.[2] Unlike households seeking refuge in an urban area (whether settled in camps or not), rural refugees have access to rural livelihoods like agriculture or livestock rearing. During the emergency phase when refugees are entirely dependent from aid, assistance covers their basic needs: housing, health and food security. When the situation persists, as the return cannot be considered, camps steadily organize themselves as villages and endogenous social and economic life rise again. From the initial humanitarian assistance, relief agencies shift to a development phase to support the progressive self-sufficiency of refugee households. Depending on the local context, refugees are given a piece of land, seeds, agricultural tools, livestock, training and assets for starting income generating activities (IGA).

The transition from humanitarian support to development means that aid has to be reduced as households become more and more self-sufficient. Nevertheless, this transition remains a political and operational challenge: a fine-tuning of the aid options and a targeting of the recipients are required. Reducing or removing food aid is a very critical issue that every camp settled for the long run have to deal with. The present analysis lies in this transition challenge and focus on a potentially perverse effect of food aid for alleviating long-term vulnerability. Regular food distributions are expected to prevent food shortage and promote food security among the refugees. However, the expected safety net role of food aid relies on the households' capacity to smooth their consumption. With the help of data collected in Southern Chad from Central African Republic refugees, we ascertain the role of a regular and universal food distribution cycle on consumption and children nutritional status. The survey, designed to measure an exogenous variation in the time interval between the previous food distribution and households interviews, allows us to assess whether the rhythm of food aid distribution can affect household's consumption and children's health. Our results suggest that cereal consumption decreases significantly as time goes away from the distribution and that the effect is sufficiently large to impact children short-term nutritional status.

This chapter first discusses the current literature related to aid distribution, consumption smoothing and refugees population, and underlines the background as well as the knowledge gaps this analysis is filling. Section three focuses on the context of southern Chad in which refugees evolve. The conceptual framework and data are discussed respectively in sections four and five. The empirical models and subsequent results are presented in section six and discussed in section seven.

[2]It is worth noting that the location decision as well as refugees rights results from a discussion between UNHCR and the host country. As a consequence, the right to move or to engage in economic activities may differ from a camp to another.

2.2 Literature

This paper falls within the scope of the literatures on refugees, on food aid effectiveness and on consumption smoothing. Surprisingly, the literature on the specific economic life of refugees is quite limited and focuses mainly on the impact of refugee's influx on the host country and natives' welfare (see Alix-Garcia and Saah (2010) or Baez (2011)) and the consequences of the initial displacement shock.[3] For instance, Adelman et al. (2013) show how being displaced has a long-term negative impact on child growth, with the help of regression discontinuity estimations. The consequences of being a refugee in the long run is, however, rarely discussed, with the exception of the medical literature where the nutritional status of refugee populations is regularly assessed. Consistently with our own analysis, several studies report a large prevalence of malnutrition among refugees: Toole et al. (1988) link malnutrition to the delivery of an inadequate food ration while Seal et al. (2005) emphasize the deficiencies on nutrients such as iron and vitamin A.

The allocation of aid, one of the main tools against malnutrition in camps and a key feature of the economic life of refugees, has been largely debated. Several studies show that food aid in an emergency context has an immediate positive effect on children survival and growth (Nieburg et al. (1992); Quisumbing (2003)), as well as on food consumption (Abebaw et al. (2010)). According to Gilligan and Hoddinott (2007) this positive effect is persistent over time if food assistance program prevent households from selling their productive assets. While food aid appears essential during a crisis, its effectiveness is often questioned. Dercon and Krishnan (2005) claim that food aid targeting is generally very poor at the micro level and, when delivered, could jeopardize informal aid networks. In the same way, Yamano et al. (2005) also argue that poor targeting is a major constraint to the positive effect of food aid on child stunting. Finally, in the long run, assistance is often suspected to have pernicious effects on an individual's self-sufficiency. Abdulai et al. (2005) argue that food aid has no deterrent effect on food production.

A large part of the literature on consumption smoothing in developing countries focus on the effect of income variability over the year to test the permanent income theory and highlight the role of the credit market on consumption (see Townsend (1995)). Paxson (1993) shows how seasonal variations in consumption is not linked to the variability of ressources but to seasonal variations in price and preferences. Deaton (1992) demonstrates with data from Cote d'Ivoire how farmers who anticipated a reduction of their future agricultural production used to save saved more than the others. We differ from this literature by focusing on the day-to-day variation of consumption over a short period of time between two flow of resources.

[3]See Jacobsen (2005) for an overview of the literature and Werker (2007) for a specific look at the economy of a refugee camp.

2.3 Refugees in the context of Southern Chad

2.3.1 Political and historical framework

From fall 2002 until the overthrow of President Patassé in March 2003, the occupation of northern Central African Republic by rebel troops of General Bozizé and induced violence have led the population to flee their villages and to find refuge, across the border, on Chadian soil where the population share languages and ethnicities. After having been welcomed by frontier zone villages at first, refugees were settled further inland in the first camp of Amboko, located at six kilometers from the small town of Goré, to secure the frontier.[4] This first influx of refugees was followed by two other waves after the presidential election, in June 2005 and in February 2006. As a consequence, two more camps, Gondjé and Dosseye, opened respectively in 2005 and 2006. These groups were made of households from the same area struck by violence and destruction of their villages. As a consequence, the refugee status was granted on the *prima facie* principle without individual interviews.

At the time of the survey in 2010, the situation in Central African Republic was not considered sufficiently safe by UNHCR to plan a come back through a reintegration program. Similarly, most of the refugee households reported us not considering a return as a potential option in the short run. The recent rise in violence in CAR reduced again the opportunities of coming back for the refugees settled for more than 10 years in the camps while new influx increased the population of the camps and even led to the creation of new camps. Indeed, in June 2012, the same camps have seen an influx of 1550 new refugees fleeing from violent clashes in Northern Central African Republic. The situation has remained critical since then, and UNHCR estimates that more than 60, 800 persons have been forced into exile in neighboring countries (Democratic Republic of the Congo, Cameroon and Chad) since December 2012.

2.3.2 Relief and aid: from humanitarian to development policy

From the original shock to the present situation, the dynamics of assistance follows the evolution of refugees' needs under strong logistical constraints. As they fled away, refugees left everything behind them and had nothing at their arrival; in our sample, only 13.7% of them arrived with a luggage. While some tried to bring with them what they owned, many report to have been stolen by rebels or soldiers as they were fleeing away. Once in camps, the first steps of assistance were to provide households with security, housing and food. Housing conditions reflect the evolution of refugees' living standards since their arrival: permanent

[4]Goré is the main town of the department of Nya Pende, which explains the presence of administrative organization. It is also the main market of the area.

houses, built by refugees and similar to natives' ones, have steadily replaced the initial tents. At the time of the survey, no household was still living in a tent. Over the first months, refugees get access to water through tube wells bored in camps, sanitation, free health care and, eventually, to education as primary schools were built in the camps. As per our computation based on figures provided by CARE International, UNHCR, WFP and UNICEF, the average cost to get access to water, sanitation, food distribution, health care and education was equivalent to a subsidy of 91 USD per person per year.[5]

In addition to this assistance, restoring households' productive capacity was a concern from the very beginning, with the aim of developing households' self-sufficiency. Every household was given one hectare of cultivable land as well as seeds and basic tools to start cultivation, which was the main activity for a large part of the refugees before the displacement. Thanks to a favorable climate and fertile soils, agricultural production is profitable in southern Chad. However, the small plots size and the impossibility to let the land lie fallow have reduced soil fertility, which may have a deterrent effect on refugees' will to engage in agricultural work.[6] Apart from agriculture, refugees are free to engage in income-generating activities (IGA). Trade, livestock rearing, working as a day laborer, tailoring and other craft activities are among the most common activities in the camps. IGA are supported by relief agencies that granted training and/or initial equipment in some cases. It is worth noting that IGA are not limited to the camps. Refugees and natives are free to move in and out of the camps[7], which facilitate economic activities, especially trade with the surrounding villages and the city of Goré.

Over the years, the aid policy shifted from the resolution of a humanitarian crisis to a development policy in all three camps. As a consequence, camps steadily became villages with close economic relationships with their native neighbors, both populations sharing markets, schools and health infrastructures. This progressive shift of policy has important consequences for refugees as the payment of school fees, health consultations, medicines and access to water was gradually introduced in 2010.

2.3.3 Food aid distribution

Access to food aid follows the same trends. Important rates of acute malnutrition have been registered among newly arrived refugees. In Amboko, 30% of children under five have been suffering from acute malnutrition when the camp was established (UNHCR & WFP (2004)).

[5]This estimation does not include the investment costs for tube wells, schools and health center, or the transport for food distribution.

[6]Refugees can expand the size of their cultivable land by renting a plot from natives. Unlike working as a day laborer, which is quite common among refugees, renting a plot is a very limited phenomenon.

[7]Refugees require a formal authorization from the local authority to move outside of the camp, but these authorizations are not needed for daily moves.

The initial goal of food distribution was to eradicate short-term malnutrition by removing food shortage resulting from the displacement shock. At first, refugees were given a ration that equals of 2100 kcal per person per day. This "full ration", that was aimed to cover adults daily needs, was composed of 400 grams of cereals, 70 grams of pulses, 25 grams of cooking oil, 5 grams of salt and 50 grams of "CSB" (Corn Soya Blend). Rations were not fortified with extra micronutrients. At the end of the emergency phase the daily ration felt to 1200 calories per person in December 2006 in Amboko and Gondjé, and in April 2010 in Dosseye (see Table 2.1 for a timeline). The impact of this change on malnutrition in the camps remains uncertain as the decision to reduce the ration size is (at least partly) related to the average refugees' self-sufficiency observed by relief agencies.

Selling the food ration is allowed and common. As per our data, 55% of the households sold part or their entire ration from the previous distribution. According to Reed and Habicht (1998), the two main reasons why refugees would sell their ration are the desire to balance their diet (with vegetables or meat) and the inadequacy of the distributed cereals with refugees' cultural preferences. A similar pattern was observed in our context: the rice ration was usually sold (or exchanged) in order to buy sorghum or other food items than cereals. Therefore, the ration cannot be seen as a pure "in kind" transfer. Interestingly, 87% of the households selling their ration do so in the week following the distribution, suggesting a strong need for cash.

The food distribution is the main event in the camps. In theory, food was distributed on a monthly basis. However, the food supply chain faced logistical and financial constraints that often resulted in late distributions as well as inadequate quantity or composition of the ration. These logistical issues may have direct and indirect effects on households' welfare. Firstly, as their selling prices largely differ, the monetary value of a ration composed of rice was significantly higher than a ration composed of sorghum. Secondly, the exact distribution dates are only announced few days in advance. While refugees do not question food delivery in itself[8], it is impossible for a household to anticipate the exact distribution date. Moreover, it is worth noting that the rate of the distribution was expanded from one to two months in October 2010. The daily quantity per person remained the same but the storage and therefore smoothing of the food was transferred to the households.

[8]During our visits in camps, refugees expressed doubts about the availability of food aid in the long run but not in the short run.

2.4 Conceptual framework

The basic framework of this analysis consists of an indefinitely repeated period whose starting point is the displacement shock ($t \in [S - \infty[$). At each period t, refugees received an exogenous amount of food aid (F_t) that is strictly the same for every individual but does not cover the total individual needs (N_{it}) for the period ($\forall t : F_{it} < N_{it}$). We make the assumption that the amount of food aid received is exogenous to the refugees. The ration is individual. Each household receives as many rations as the number of members on their distribution card. In other words, the ration is distributed equally according to the household size. Households can decide to leave their refugee status, and therefore assistance, but cannot directly influence the amount of food aid they receive.[9]

In order to maximize food consumption and cover their needs, households can adjust labor supply and engage its members into income-generating activities that will generate either a daily income (dI), like daily laborer for agriculture or taxi driver, or a distant and uncertain income (uI) from a labor-intensive activity (farming or livestock rearing). In addition, at each period, households can sell assets and borrow from the credit market to fill in their resource gap. We consider it as an adjustment income (aI). Therefore, the household's resources, I_t are define as:

$$I_t = \sum_t dI + uI + aI \tag{2.1}$$

At each period, household may adopt survival strategies in order to meet their basics needs. These survival strategies will determine dI, uI and aI. At the household level, they consequently have to respect this equivalence to meet their basic needs:

$$\sum N_{it} \approx \sum F_{it} + \sum_t dI + uI + aI \tag{2.2}$$

Survival strategies adopted at time t depend on welfare at time $t-1$, strategies availability and previous experiences of the household. Relief agencies can help households engage into activities by promoting them and providing needed inputs.

This sequence is preceded by an exogenous shock that leads households to flee their homes, lands and activities. Even if one purpose of relief assistance is to help households recovering from the shock, the direct effect of the shock is out of the scope of this paper and could not be assessed properly with our data set.

The rate at which aid is distributed is beyond household influence and depends mostly on logistical issues. In addition, due to the lack on information households feel uncertain about

[9]In practice, the amount of food aid received by refugees is not completely exogenous to them. They can expand it by artificially raising the size of the household; for example, they can hide the death of a member or declare as refugees, relatives that are not living in the camp. We expect this phenomenon to be marginal and we are therefore confident that this issue does not jeopardize our identification strategy.

the date of the next distribution. It means that the length of each period t is unknown for the household.

Within each period, a household has to decide the allocation of its resources all over the period in order to maximize its welfare. In a limited resources context, smoothing consumption is a way to avoid facing a food shortage at the end of the period. The certainty that a significant part of their resources will be provided the following period may help households smooth their consumption. On the other hand, if information is limited, uncertainty may lead households to have a strong preference for present. In a dynamic perspective, households with limited resources and a strong preference for present may engage in a vicious circle where resources received at time t are used to recover from deficiencies or to pay back a loan from $t-1$ in order to face food shortage.

2.5 Data and descriptive statistics

2.5.1 Survey design

Data used in this paper were collected in the context of the assessment of refugee household's self-sufficiency[10]. In the perspective of an integrated development policy and reduced funding, relief organizations wanted to understand how aid can be better targeted to most vulnerable households and how refugees could handle a cut in the amount of food aid. The survey was implemented from September to October 2010 among three refugee camps. The sample includes 977 refugee households, all randomly selected.[11] The availability of an exhaustive list of refugee households registered in the camps allowed us to select them randomly. Using their district of location and the ethnicity of household head, we draw a representative stratified sample of refugees' households. However, it is worth noting that being registered does not imply that a household and its members were currently living in the camp, as they were free to move out. When a selected household was not living in the camp at the time of the survey, another household of the same district and ethnic group from a second replacement list substituted him.[12] For the interest of this analysis, the questionnaire included detailed information on food consumption, agricultural production, household's activities and nutritional status of children under five years old.

The close and ambiguous relationship between refugees and donors and the specificity of refugee camps made the data collection process challenging. Donors' staff used to visit camps on a daily basis to supervise programs implementation and outcomes, discuss and as-

[10]The study was implemented by CARE International and funded by UNHCR.

[11]In order to allow comparison, 502 native households were also interviewed. As they were not concerned by food assistance, they are excluded from our analysis, with the exception of descriptive statistics.

[12]Nearly 54% households in our sample were substitutes from a second list of randomly selected households. There are no statistical differences between households selected first and substitutes.

sess refugees' needs. In this context, refugees may easily consider that their answers would have a direct impact on the allocation of any kind of future aid. In other words, incentives to give wrong information were strong. To avoid these potential biases, our strategy was *(i)* to hide as much as possible our links with UNHCR and other donors by presenting ourselves as an independent research team, *(ii)* to respect the randomly selected sample by rejecting not-selected households appeal to be interviewed, *(iii)* to ask for objectives and verifiable information like assets ownership, housing characteristics or children anthropometric measurement and finally *(iv)* to establish a trusting relationship between enumerators and refugees, households being interviewed in their maternal language, without interpreter and enumerators being settled in the camps for the time of the survey.[13]

The survey took place between two food distributions two months away from one another. At the time of the survey, WFP and UNHCR had decided to extend the duration between distributions from one to two months. Except for few exceptions due to logistic issues, it was the first two months ration distributed to the refugee population. It is also worth noting that the survey took place at the beginning of the harvest period. Agricultural production records refer to the previous harvest.

2.5.2 Food consumption

Food consumption was estimated in two different ways. Firstly, households were asked to recall the quantity of cereals consumed and how often they have consumed specific items during the previous seven days. This usual way to assess food consumption suffers from measurement errors due to the difficulty for respondents to recall precisely what they have eaten, how much and who shared the meal with them. An alternative option is to use a dietary recall survey where respondents describe the household food consumption over the previous 24 hours in details (see Jacoby (2002) for an example). For each of the three main meals of the day, respondents gave a precise description of the type and quantity of food eaten as well as the number of persons present.[14] The main advantage of this second method is to provide a precise measurement of the daily caloric intake. It has to be kept in mind that this measure may not be representative of the household average consumption in a longer run. For instance, daily consumption may respond to the daily labor supply, to an unexpected cash inflow or to a special occasion. Nevertheless, the 24-hours recall method collects recent information in the memory and measurement errors are less likely, and consequently this method is consequently preferred to the usual seven days recall method.

Nearly 60% of sampled households declared having eaten two meals the day before the

[13]Enumerators were not hired locally but in N'Djamena to avoid any collusion.

[14]Unlike Jacoby (2002), we assessed food dietary at the household level and focused only on the three main meals. Individual calorie intakes out of the main meals like snacks cannot be estimated with our strategy.

interview, while 30% have eaten three. Half of the households skipped lunch, which is quite common during the harvest time: workers eat before and after farming activities, which usually take place far away from house. Cereals are the basis of household's diet. Every household has eaten cereals during the seven previous days and 92% ate cereals the day before the interview. The two main cereals consumed are sorghum, which is locally produced, and rice, which is part of the food aid or can be purchased (at a higher price than sorghum). On average, individual intake from cereal consumption the day before the interview was 1379 kcal. Native household's cereal consumption was significantly higher (1766 kcal).

2.5.3 Anthropometric measurements

In the specific context of refugee camps, anthropometrics measurements are a precious unbiased measure of poverty. A total of 704 refugee children under five years old have been measured for their height, weight and mid upper am circumference (MUAC).[15] These measures allow us to assess children's long and short-term nutritional status. Height for age z-score (HAZ) is an index for long-term or chronic malnutrition that is influenced by the whole child history of food shortage. Weight for height z-score (WHZ) reflects the current nutritional status, i.e. the effect of recent episodes of food shortage. Weight for age z-score (WAZ) is a combination of these two indicators. Alongside those traditional indicators of malnutrition, MUAC is often used in emergency contexts for practical reasons, measurement being cheap and easy, but also because it is very sensitive to abrupt changes in nutritional status (Bern and Nathanail (1995)). Moreover, MUAC appears to be a complementary tool to WHZ in monitoring nutritional status and predicting death of malnourished children (Berkley et al. (2005)). The average growth of the arm circumference between one and five years old is very low (Frisancho (1974)) but raw MUAC should nevertheless be expressed into z-score indexes by age or height, as De Onis and Habicht (1996) recommend. Table 2.6 shows the expected correlations between the various indexes. As expected, the correlation between short-term indexes (MUAC, WAZ and WHZ) is highly positive (0.36 to 0.52) while the correlation between short and long-term ones is small, though only marginally significant.

High prevalence of malnutrition was one of the motivations for the paper. Despite most of the children being born in camps, and having therefore always lived under the food distribution coverage, 37% of the children under five sampled in camps are considered as stunted (HAZ<-2), while mean HAZ is -1.38 (Table 2.4).[16] Considering the whole sample of children

[15]The measurement protocol respects the one established by UNICEF. Children were measured with wooden height gauge (in lying or standing-position depending on their age), suspended scales and measuring tape. Z-scores have been generated with the help of the World Health Organization's Anthroplus software. A total of 480 native children have also been measured but are not included in the econometric analysis.

[16]Nearly 15% of the sampled children were born before their parents became refugees. Most of them came with the last influx of refugees and are settled in the camp of Dosseye. There is no statistical difference in HAZ between refugee children born inside and outside of the camp for a given age.

under five years old, there is no difference in stunting between refugee children and natives one: mean HAZ is -1.38 for the former and -1.43 for the latter. However, both populations differ in the way mean HAZ decreases with age. Native children have a better HAZ than refugee children when they are less than two years old but a poorer HAZ when they are two and older. In both cases, difference in mean is significant and robust to the exclusion of refugees born outside (Table 2.5). This result may reflect the long-term consequences of the food distribution but, as rations are not the only difference between natives and refugees, this question could not be assessed directly.

Prevalence of short-term malnutrition was low at the time of the survey. Twelve percent of refugee children under five could be considered as wasted (WHZ<-2) and four percent severely wasted (WHZ<-3), with an average mean of WHZ of -0.34, which means that their households have faced recent episodes of food shortage. In addition, short-term malnutrition appears significantly more prevalent in camps than in neighboring villages, whatever age category is considered. This finding is coherent with the differences found in cereal consumption between refugees and natives but may question aid efficiency. It is worth noting than prevalence of wasting is significantly higher in Dosseye (20.6%) where last arriving households have settled. Finally, MUAC is a more responsive anthropometric index to short-term food shortage and gives complementary information on short-term malnutrition. Nearly 8% of the refugee children in the sample are under the cutoff used by nutrition specialists of the field to diagnose wasting (12.5 cm). Mean MUAC for height and MUAC for age are respectively -1.10 and -1.18.

2.5.4 Time interval variable

The time interval variable used in the following regressions measures the number of days between the previous distribution and the date of interview (and anthropological measurements) for each household. Computing time interval from the food distribution to the day of the interview is straightforward but requires to address challenges on the field in order to make sure the time variable is not correlated with camps or households unobservable characteristics. All three camps have been surveyed at the same time, with three teams of enumerators working simultaneously. Our strategy was to split the sample into two identical randomly selected sub-samples and to interview those two groups at two different periods. Specifically, the first five hundred refugee households were interviewed during the first week of survey whereas the last five hundred were interviewed during the third (and last) week of survey[17] As a consequence, if we consider households of a specific ethnic group living in a given district, half of them were interviewed during the first third of the survey and the other half at the end.

[17]The intermediate week was devoted to survey native households in surrounding villages.

The average time interval is 31.8 days and the standard deviation is 7.3. Nine days (three per camp) are usually needed to distribute food to refugees. Since camps were visited over the same period but the distribution dates differ slightly, the average time since distribution is different for the three camps: 26 days on average in Amboko, 31 days for Gondjé and 37 for Dosseye (Tables 2.1 and 3.1). The survey design allows us to have a large variation in the time interval variable: the closest households to the distribution were interviewed 17 days after, while the maximum time interval is 46 days.

Using our design, the refugees' sample, as well as each camp sub-sample, can be divided into two groups of equal size whether households were interviewed during the first or the last part of the survey. The average difference between the two groups mean time since distribution is 11 days (26 days on average for the first half sample; 37 days for the last half). Table 2.3 shows mean-comparison tests between the two samples. Except for the household size, there is no significant differences for household characteristics that are relatively stable over time like assets and housing index, gender and age of household head and measured children, previous season cereal production or presence of a French speaking member (a proxy for the education level) within the household. Household size differs widely between the two groups. Though the magnitude of the difference is unexpected, this result is not surprising if we consider our definition of a household and the fact that some members used to move in and out the camp for long periods.[18] However, there is no difference between the two groups if we consider the number of persons present during each meal.

If these households characteristics do not differ based on the time since the distribution, this is not the case for the indicators of food consumption and malnutrition. As expected, cereal consumption is significantly lower on average among households interviewed during the last part of the survey. Moreover, both MUAC z-scores are significantly lower for children measured during the last part of survey whereas there is no difference for HAZ, WAZ and even WHZ.

2.6 Results

2.6.1 Results on consumption

First, we examine how the elapsed time since the previous distribution affects daily household consumption by estimating the following empirical model:

$$C_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 H_{ij} + \beta_3 S_{ij} + \upsilon_j + \varepsilon_{ij} \qquad (2.3)$$

[18]In our questionnaire, a household is composed of permanent members who leave together and share the same meals. A member is considered permanent if he or she is present during the interview or if he was not absent for more than a month.

where C_{ij} is the daily cereal consumption of the household i from block j of one of the three camps; T_{ij} is the time since the distribution; H_{ij} is a vector of fixed households characteristics, S_{ij} is a vector of household's survival strategies and v_j is the specific block characteristics.

Cereal consumption of the day before the interview is expressed in logarithm and per individual. This implies that households who have not eaten cereals were excluded from the regressions. Dummy variables for ethnicity and the day of the week a household was interviewed have been included to control for diets habits or customs related to a specific day of the week. The time interval is either expressed in days or with a simple dummy variable that split the sample into two parts whether the interview took place during the first or last part of the survey. The average number of days from distribution differs from one camp to another; this implies that equation (2.3) is estimated with a fixed-effect model at the block level. Through the random assignment of each household to one group or the other, our design ensures the orthogonality of time to unobservable camps or blocks characteristics. T_{ij} can be considered as the main variable of interest; a significant β_1 would contradict the assumption of households smoothing consumption.

Including variables related to survival strategies in the model could lead to endogeneity issues since they could respond to the current household nutritional situation. When cereal stocks become critically low, a household may engage in daily activities or sell assets. Having borrowed money or sold a part of the ration since last distribution are clearly strategic behaviors in response to an urgent need. Those variables are included as control variables and should be interpreted carefully when included in the regressions. However, farming, livestock rearing or IGA are activities with long-term return that may not be entirely questioned by occasional food shortages, even if endogeneity cannot be ruled out[19].

Table 2.7 shows the regression analysis results. All other things held constant, the time interval from distribution has a negative and significant effect on food consumption. On average, caloric intake from cereal consumption the day before the interview is 11% lower for households interviewed during the last part of survey. Considering the number of days, each day away from the distribution is associated with an average loss of 1% in cereal consumption, suggesting that households do not smooth their consumption between two distributions. Moreover, this relationship between time and consumption may not be linear; indeed, the square of the number of days is also significant, but positive. This result suggests that the decrease in consumption is higher at the beginning of the period. Results are robust to the inclusion of survival strategies variables. The role of agricultural production is a significant determinant of the present daily food intake (through self-consumption). Similarly, being engage in income generating activities also has a positive and significant impact on consumption while being a borrower or having sold a part of the ration is associated with a lower

[19] The cereal production variable refers to the previous harvest. IGA refers to any IGA done during the previous twelve months. Cattle size reflects the number of cows owned at the time of the survey.

consumption.

It is also worth noting than the number of years since a household settled in a camp is associated with poorer consumption, which could question the integration strategy. However, this result could also reflect a self-selection bias resulting from the long-term dynamics of refugees' presence in the camps. Under the assumption that better off households are more likely to leave the camps to find an alternative solution, older refugees (in terms of number of years spent in camps) differ from the younger generation on average. In other words, this variable may partially catch the effect of unobservable characteristics that influenced the decision to stay in the camp.

Household's size has a negative effect on the individual calories intake: each additional member decreases the average individual cereal consumption between 1.8% and 2.3%. The coefficient should be interpreted carefully since household's size is used in the computation of the dependent variable and appears to be sensitive to the use of adult equivalent ratio to adjust for household size.

Results are robust to alternative specifications presented in Table 2.8. Firstly, the use of camp fixed effect instead of blocks fixed effects gives similar coefficients as well as similar significance. More interestingly, the effect of time remains significant when the sample is restricted to households from the same camp: the effect is larger in Amboko (-1.8%) and similar in Gondjé (-1.1%), but not significant in Dosseye. Moreover, the use of equivalent adults ratios to adjust household size to the relative needs of household members gives higher estimates of time effects.[20] Finally, we use raw individual caloric intake as an alternative dependent variable and find similar results. This measurement allows us to include households that have not eaten cereals the day before the interview. As expected, the effect size is reduced but remains significant.

2.6.2 Results on malnutrition

Results on consumption indicate that preference for present prevents households from smoothing their consumption, creating monthly period of food shortage while waiting for the next distribution. In this section, we assess if this cycle has any negative impact on nutrition through the reduction in consumption. In order to examine whether the time interval since the last distribution affects children nutritional status, we estimate the following empirical model:

$$M_{ij} = \beta_0 + \beta_1 T_{ij} + \beta_2 CH_{ij} + \beta_3 H_{ij} + \beta_4 S_{ij} + v_j + \varepsilon_{ij} \tag{2.4}$$

where M_{ij} is the nutritional status of child i in block j; T_{ij} is the same measure of time since

[20] In our computation, each member aged five to sixteen and children under five counted respectively for 0.5 and 0.3 adult.

distribution used previously; CH_{ij} a vector of child characteristics; H_{ij} is a vector of fixed household characteristics, S_{ij} is a vector of households' survival strategies and v_j captures the block specific characteristics. Nutritional status is expressed by short-term anthropometric indicators (MUAC/H, MUAC/A and WHZ). As the previous one, this model is estimated with fixed-effect at the block level and ethnicity dummies.

Results are presented in Tables 2.9 and 2.10. Gender, age and illness control variables show expected signs: girls have on average significantly better nutritional status than boys; age is associated with better MUAC but lower HAZ; and recent illness lowers short-term malnutrition. Secondly, all things held constant, time since distribution has a negative and significant effect on MUAC/H and MUAC/A. This result is robust to the inclusion of household cereal consumption of the previous seven days. Moreover, the time variable remains significant if we restrict the regressions to households settled in Amboko or Dosseye only, despite the small sample size (Table 2.11). The alternative short-term malnutrition indicator, WHZ, which is less sensitive to short-term changes in the child's diet, is not significantly determined by time.

If children who experience a poor short-term nutritional status at present time are more likely to suffer long-term malnutrition, our results might be driven by a long-term dynamic rather than being the consequence of the present household's consumption. While our random sampling strategy prevent our analysis from this potential issue, we rule out this interpretation by implementing a falsification test, using a long-term malnutrition indicator (HAZ) as the dependent variable. As expected, the interval of time since distribution has no impact on long-term nutritional status (Table 2.10), suggesting that the impact of time on short-term malnutrition is not linked to the previous nutritional status but to the present situation.

As for the effect of time on consumption, the relationship between malnutrition and time appears to be non-linear. The number of days since the distribution and its squared version are jointly significant on both MUAC z-scores. Our results confirm the detrimental effect of time on household's welfare. Households' inability to smooth their consumption over the interval of time between two distributions creates monthly food shortages that weaken children current nutritional status. Since catch-up growth appears to be limited in the African context (Hoddinott and Kinsey (2001); Rieger and Wagner (2014)), we can expect this results to have a permanent impact on children's long-term abilities.

2.7 Discussion and policy recommendations

While their living standards are converging as the result of the integrated development policy, native and refugee households are facing food shortages with a different timing. For a typical rural household, whose life is punctuated by the annual agricultural season, the most critical period is the lean-season when one is waiting for the next harvest while the stock of cereals from the previous harvest is over. The monthly food aid could be seen as a safety net protecting refugees from the lean season. The absence of large periods of food shortage for refugee households may explain the relatively better long-term nutritional status of their children as compared to neighboring natives. However, the prevalence of malnutrition remains worrying in camps and, at the time of the survey, short-term malnutrition was significantly higher among refugees. Our results indicate that households do not (or cannot) smooth consumption during the interval of time between two distributions. They face short cycles of food shortage (akin to lean season) that have detrimental consequences on children malnutrition. The great occurrence of borrowing, mainly for food, days before the next distribution is consistent with our results.

Why households do not smooth their consumption? Which constraints do they face? Firstly, it should be kept in mind that a large part of the refugee population remains very vulnerable. Considering the whole population that fled after each crisis, our sample is composed of those who have decided to settle in camps rather than coming back home or seeking better opportunities elsewhere. As an illustration, we observe in our sample a large share of small households without members in the prime of life that is out of proportion as compared to the native households. The food aid is designed not to cover the basic needs fully. Therefore households who can't complete food aid with their own ressources, might find difficult to smooth their consumption simply because their own source of income is not sufficient. Secondly, despite the routine of food distribution, refugee households keep living in uncertainty, which makes decisions about survival strategies more difficult to take. From a longer term view, uncertainties about the duration of their settlement make any sizable investment insecure. In the short-run, there is a lack of information about the availability and contents of aid as well as the duration of intervals between distributions that could result in incapacity to manage food aid efficiently.

Our analysis does not go without limitations. The main concern is about the duration of the time interval between two distributions. Our survey took place within a two months interval after several one-month periods. Our results could therefore be interpreted as the transition from a one-month period to a two-months period. We cannot ruled out this assumption because we observe the situation within a single period identical to every households. However, it should be noticed that it was not the first time refugees had to deal with a two-months period. As the planning is mainly driven by logistical issues, the composition of the ration and the length of the period it is expected to cover can vary. Secondly, our results suggest that the diminution in cereal consumption is larger at the beginning of the period. Panel data would be useful to study the dynamics of refugee recovery and their consumption behavior under different time interval between distributions.

Our findings have important implications and argue for a set of policy recommendations. The way food aid is distributed has a direct effect on refugees' welfare. A first recommendation would be to reduce the time interval between two distributions in order to avoid long period of food shortage: the shortest the interval, the easiest the food stock management would be for households. The logistical implication of such a change may however be its main obstacle: the decision to expand the interval from one to two months in southern Chad was mainly driven by the need to lighten the logistical cost of food delivery. Switching to a vouchers strategy would probably relieve the pressure on the delivery date. A more simple recommendation is to reduce uncertainty on the allocation of aid to the minimum. Thirdly, failure to smooth consumption is often associated with a poor access to credit market. As it is emphasized in the paper, borrowing is closely related to consumption behavior. Better access to credit and the guarantee of future aid would probably help households better manage their resources over time. Finally, as our results suggest a large dependency to food aid, the main challenge is to support survival strategies in the long run and promote self-sufficiency.

Bibliography

ABDULAI, A., BARRETT, C. B., AND HODDINOTT, J. 2005. Does food aid really have disincentive effects? New evidence from Sub-Saharan Africa. *World Development* 33:1689–1704.

ABEBAW, D., FENTIE, Y., AND KASSA, B. 2010. The impact of a food security program on household food consumption in Northwestern Ethiopia: A matching estimator approach. *Food Policy* 35:286–293.

ADELMAN, S., GILLIGAN, D., AND LEHRER, K. 2013. Childhood interrupted: How conflict-induced shocks shape human capital formation. Working paper, International Food Policy Research Institute.

ALIX-GARCIA, J. AND SAAH, D. 2010. The effect of refugee inflows on host communities: Evidence from Tanzania. *The World Bank Economic Review* 24:148–170.

BAEZ, J. E. 2011. Civil wars beyond their borders: The human capital and health consequences of hosting refugees. *Journal of Development Economics* 96:391–408.

BERKLEY, J., MWANGI, I., GRIFFITHS, K., AHMED, I., MITHWANI, S., ENGLISH, M., NEWTON, C., AND MAITLAND, K. 2005. Assessment of severe malnutrition among hospitalized children in rural Kenya: Comparison of weight for height and mid upper arm circumference. *Jama* 294:591–597.

BERN, C. AND NATHANAIL, L. 1995. Is mid-upper-arm circumference a useful tool for screening in emergency settings? *The Lancet* 345:631–633.

DE ONIS, M. AND HABICHT, J.-P. 1996. Anthropometric reference data for international use: recommendations from a world health organization expert committee. *The American Journal of Clinical Nutrition* 64:650–658.

DEATON, A. 1992. Saving and income smoothing in Cote d'Ivoire. *Journal of African Economies* 1:1–24.

Dercon, S. and Krishnan, P. 2005. Food aid and informal insurance. *Insurance against Poverty* p. 305.

Frisancho, A. R. 1974. Triceps skin fold and upper arm muscle size norms for assessment of nutritional status. *The American journal of clinical nutrition* 27:1052–1058.

Gilligan, D. O. and Hoddinott, J. 2007. Is there persistence in the impact of emergency food aid? Evidence on consumption, food security, and assets in rural Ethiopia. *American Journal of Agricultural Economics* 89:225–242.

Hoddinott, J. and Kinsey, B. 2001. Child growth in the time of drought. *Oxford Bulletin of Economics and statistics* 63:409–436.

Jacobsen, K. 2005. The economic life of refugees. Kumarian Press.

Jacoby, H. G. 2002. Is there an intrahousehold 'flypaper effect'? Evidence from a school feeding programme. *The Economic Journal* 112:196–221.

Nieburg, P., Person-Karell, B., and Toole, M. 1992. Malnutrition-mortality relationships among refugees. *Journal of Refugee Studies* 5:247–256.

Paxson, C. H. 1993. Consumption and income seasonality in Thailand. *Journal of Political Economy* pp. 39–72.

Quisumbing, A. R. 2003. Food aid and child nutrition in rural Ethiopia. *World Development* 31:1309 – 1324.

Reed, B. and Habicht, J. 1998. Sales of food aid as sign of distress, not excess. *The Lancet* 351:128–130.

Rieger, M. and Wagner, N. 2014. Child health, its dynamic interaction with nutrition and health memory–evidence from senegal. *Economics & Human Biology* .

Seal, A. J., Creeke, P. I., Mirghani, Z., Abdalla, F., McBurney, R. P., Pratt, L. S., Brookes, D., Ruth, L. J., and Marchand, E. 2005. Iron and vitamin a deficiency in long-term African refugees. *The Journal of Nutrition* 135:808–813.

Toole, M., Nieburg, P., and Waldman, R. 1988. The association between inadequate rations, undernutrition prevalence, and mortality in refugee camps: Case studies of refugee populations in eastern Thailand, 1979–1980, and eastern sudan, 1984–1985. *Journal of Tropical Pediatrics* 34:218–224.

Townsend, R. M. 1995. Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *The Journal of Economic Perspectives* pp. 83–102.

UNHCR 2013. Global appeal 2012 - 2013. Technical report, UNHCR.

UNHCR & WFP 2004. Rapport de la mission d'evaluation conjointe PAM / HCR / Gouvernement du Tchad / ONGs, des besoins des populations réfugiées centrafricains au sud du Tchad. Technical report, UNHCR - HCR.

Werker, E. 2007. Refugee camp economies. *Journal of Refugee Studies* 20:461–480.

Yamano, T., Alderman, H., and Christiaensen, L. 2005. Child growth, shocks, and food aid in rural Ethiopia. *American Journal of Agricultural Economics* 87:273–288.

Table 2.1: Camp's characteristics

	Amboko	Gondjé	Dosseye	Total
Creation of the camp	2003	2005	2006	
Date of first food distribution	Jul. 2003	Jan. 2006	Feb. 2007	
End of full ration	Dec. 2006	Dec. 2006	Apr. 2010	
Survey	Sep. 2010	Sep. 2010	Sep. 2010	
Number of households sampled	323	321	333	977
Household's size	6.59	5.85	5.61	6.01
Proportion of female household's head	0.48	0.44	0.44	0.45
Number of days since distribution	26.55	31.13	37.42	31.76
Total cereal consumption (in kcal)	1522	1727	1491	1576
Number of children sampled	198	245	261	704
Height for age z-score	-1.36	-1.34	-1.42	-1.38
Weight for age z-score	-0.95	-0.88	-1.43	-1.10
Weight for height z-score	-0.16	-0.16	-0.65	-0.34
MUAC for height z-score	-1.12	-1.06	-1.10	-1.09
MUAC for age z-score	-1.21	-1.11	-1.21	-1.17

Table 2.2: Summary statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Number of days since distribution	31.8	7.3	17	46	964
Household's size	6.01	3.12	1	27	977
Number of years in camp	5.07	1.56	1	7	970
Assets and housing index	0.00	2.24	-2.88	13.21	970
Gender of household head (female=1)	0.45	0.50	0	1	969
Age of household head	40.7	13.4	3	94	966
=1 if one member speak french	0.46	0.50	0	1	977
Number of cows	0.41	1.27	0	11	977
Practice an income-generating activity	0.45	0.50	0	1	977
=1 if household has sold a part of ration	0.55	0.50	0	1	958
=1 if household borrowed	0.34	0.47	0	1	977
Cereal production (kilograms)	309	430	0	2700	972
log(Cereal consumption)	7.18	0.68	0.54	8.27	817
Cereal consumption among sample	1379	870	0	3930	917
Cereal consumption among cereal consumer	1580	745	306	3930	798

Table 2.3: Mean-comparaison test by period of survey

	First half	**Second half**	**t-stat**	**N**
Cereal consumption among consumers	1596	1503	*0.0834*	816
Cereal production	309	311	*0.9237*	970
Household's size	6.50	5.54	*0.0000*	975
Number of person during morning meal	5.96	6.32	*0.1887*	975
Number of person during midday meal	3.89	3.73	*0.6041*	975
Number of person during evening meal	5.91	5.95	*0.8837*	975
Assets and housing index	-0.05	0.04	*0.5471*	968
Gender of household head (female=1)	0.45	0.45	*0.9607*	967
Age of household head	40.8	40.6	*0.8508*	964
=1 if one member speak french	0.48	0.45	*0.3632*	975
Number of cows	0.43	0.40	*0.7183*	964
Practice an income generating activity	0.42	0.47	*0.1081*	975
=1 if household has sold a part of ration	0.58	0.52	*0.0907*	956
=1 if household borrowed	0.35	0.34	*0.5857*	975
Height for age z-score	-1.32	-1.45	*0.3734*	651
Weight for age z-score	-1.08	-1.15	*0.5328*	652
Weight for height z-score	-0.40	-0.25	*0.2269*	647
MUAC for height z-score	-0.98	-1.25	*0.0034*	629
MUAC for age z-score	-1.06	-1.33	*0.0012*	620
Children's gender	0.47	0.44	*0.5650*	609
Children's age in months	31.0	30.8	*0.8424*	609
% of children sick	0.36	0.33	*0.5520*	609
% of disabled children	0.01	0.01	*0.4283*	609

Table 2.4: Mean of malnutrition indexes among refugee children above five years old

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Height for age z-score	-1.38	1.75	-5.97	3.80	651
Weight for age z-score	-1.11	1.54	-5.50	4.82	652
Weight for height z-score	-0.34	1.50	-4.77	4.77	648
MUAC for height z-score	-1.10	1.17	-5.10	2.63	601
MUAC for age z-score	-1.18	1.03	-4.73	2.61	579
Stunting (=1 if HAZ<-2)	0.37	0.48	0	1	651
Underweight (=1 if WAZ<-2)	0.26	0.44	0	1	652
Wasting (=1 if WHZ<-2)	0.12	0.32	0	1	648
MUAC for height z-score<-2	0.20	0.40	0	1	601
MUAC for age z-score<-2	0.19	0.40	0	1	579

Table 2.5: Mean of malnutrition z-score for refugee and native children

	Age	**Refugees**	**Natives**	**t-stat**	**N**
Height for age z-score	Age$<$2	-1.10	-0.63	*0.0128*	363
	Age$\geq$2	-1.51	-1.93	*0.0030*	666
Weight for age z-score	Age$<$2	-0.82	-0.57	*0.1758*	375
	Age$\geq$2	-1.25	-1.28	*0.8003*	667
Weight for height z-score	Age$<$2	-0.35	-0.05	*0.0588*	399
	Age$\geq$2	-0.33	-0.14	*0.0933*	650
MUAC for height z-score	Age$<$2	-1.23	-1.28	*0.7585*	312
	Age$\geq$2	-1.03	-1.18	*0.0927*	673
MUAC for age z-score	Age$<$2	-1.29	-1.30	*0.9251*	289
	Age$\geq$2	-1.13	-1.31	*0.0197*	686

Table 2.6: Correlation between anthropometric indexes

	MUAC for height	MUAC for age	WHZ	WAZ
MUAC for age	0.9571***			
	(0.000)			
WHZ	0.4576***	0.3904***		
	(0.000)	(0.000)		
WAZ	0.3653***	0.5198***	0.6102***	
	(0.000)	(0.000)	(0.000)	
HAZ	0.0567*	0.2991***	-0.0884***	0.6086***
	(0.076)	(0.000)	(0.004)	(0.000)

Notes: P-values in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.7: Regressions of cereal consumption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Second part of survey dummy	-.111***			-.119***			-.117***		
	(.038)			(.038)			(.038)		
Time since distribution		-.009***	-.054**		-.010***	-.049**		-.010***	-.050**
		(.003)	(.023)		(.003)	(.023)		(.003)	(.023)
Time since distribution^2			.0007*			.0006*			.0006*
			(.0004)			(.0004)			(.0004)
Household's size	-.023***	-.023***	-.023***	-.023***	-.023***	-.023***	-.018***	-.018***	-.018***
	(.007)	(.006)	(.006)	(.007)	(.007)	(.007)	(.007)	(.007)	(.007)
Years in camp	-.095**	-.09**	-.079**	-.100**	-.095**	-.084**	-.104***	-.099**	-.089**
	(.039)	(.039)	(.039)	(.039)	(.039)	(.039)	(.039)	(.039)	(.039)
Gender of household's head	.032	.03	.036	.041	.039	.044	.053	.051	.055
	(.038)	(.038)	(.038)	(.039)	(.039)	(.039)	(.039)	(.039)	(.039)
Age of household's head	.001	.001	.001	.002	.002	.002	.001	.001	.001
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Member speaks french	.052	.049	.048	.048	.045	.045	.051	.048	.047
	(.042)	(.042)	(.042)	(.042)	(.042)	(.042)	(.042)	(.042)	(.042)
Assets and housing index	.016*	.016*	.016*	.010	.010	.011	.008	.008	.009
	(.009)	(.009)	(.009)	(.009)	(.009)	(.009)	(.009)	(.009)	(.009)
Number of cows				.012	.012	.012	.013	.013	.014
				(.014)	(.014)	(.014)	(.014)	(.014)	(.014)
Income-generating activity				.072*	.073*	.072*	.060	.061	.060
				(.039)	(.039)	(.039)	(.039)	(.039)	(.039)
Cereal production				.009**	.009**	.008*	.010**	.010**	.009*
				(.005)	(.005)	(.005)	(.005)	(.005)	(.005)
Sold ration							-.069*	-.071*	-.074*
							(.039)	(.039)	(.039)
Borrowing							-.09**	-.089**	-.087**
							(.039)	(.039)	(.039)
Number of blocks	10	10	10	10	10	10	10	10	10
N	777	777	777	777	777	777	769	769	769
R^2	.038	.038	.042	.048	.048	.052	.061	.06	.064

Notes: Clustered standard errors are in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%. Ethnicity and day dummies included in every specification.

Table 2.8: Regressions of cereal consumption (alternative specifications)

	(1) Camp FE	(2) Camp FE	(3) Amboko	(4) Gondjé	(5) Dosseye	(6) Adult Equivalent	(7) Consumption in calories	(8) Consumption in calories
Time since distribution	-.009***	-.05**	-.018***	-.011*	.0003	-.012***	-16.126***	-8.036
	(.003)	(.023)	(.006)	(.006)	(.006)	(.004)	(4.816)	(5.152)
Time since distribution^2		.0006*						
		(.0004)						
Household's size	-.019***	-.019***	-.025**	.018	-.049***	.014*	-33.937***	-17.012
	(.007)	(.007)	(.011)	(.013)	(.011)	(.007)	(9.747)	(10.514)
Years in camp	-.072**	-.062*		-.097	-.075*	-.074*	-119.862**	-97.843
	(.035)	(.035)		(.093)	(.045)	(.043)	(58.064)	(64.079)
Gender of household's head	.050	.055	.013	.056	.029	.114***	72.4	67.33
	(.039)	(.039)	(.067)	(.073)	(.063)	(.042)	(57.703)	(62.196)
Age of household's head	.0009	.0009	-.004*	.001	.006***	-.003*	1.472	1.246
	(.001)	(.001)	(.002)	(.003)	(.002)	(.002)	(2.053)	(2.206)
Member speaks french	.046	.047	.025	.067	.073	-.004	37.995	6.289
	(.041)	(.041)	(.074)	(.071)	(.074)	(.046)	(62.202)	(66.083)
Assets and housing index	.008	.008	.017	-.006	.012	.004	10.336	19.286
	(.009)	(.009)	(.019)	(.018)	(.014)	(.01)	(13.666)	(15.05)
Number of cows	.014	.014	.073*	-.036	.032*	.009	23.313	17.69
	(.014)	(.014)	(.039)	(.029)	(.019)	(.016)	(21.508)	(23.297)
Income-generating activity	.063	.062	.063	.109	.076	.063	110.982*	195.62***
	(.039)	(.039)	(.066)	(.071)	(.069)	(.042)	(57.73)	(61.83)
Cereal production	.010**	.009**	.019**	.016**	-.004	.007	14.67**	16.841**
	(.005)	(.005)	(.008)	(.008)	(.008)	(.005)	(6.909)	(7.535)
Number of groups	3	3	4	3	3	10	10	10
Ethnicity dummies	yes	yes	yes	yes	yes	yes	yes	yes
Days dummies	yes	yes	yes	yes	yes	yes	yes	yes
N	769	769	245	247	285	776	777	885
R^2	.06	.064	.142	.069	.142	.05	.049	.048

Notes: Clustered standard errors are in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.9: Regressions of MUAC for height and MUAC for age z-scores

	(1) MUAC/H	(2) MUAC/H	(3) MUAC/H	(4) MUAC/A	(5) MUAC/A	(6) MUAC/A
Time since distribution	-.025***	-.079	-.025***	-.023***	-.072	-.022***
	(.008)	(.059)	(.008)	(.007)	(.052)	(.007)
Time since distribution^2		.0008			.0008	
		(.0009)			(.0008)	
Cereal consumption			.012			.014*
			(.01)			(.008)
Gender	.09	.08	.085	.144*	.134	.143*
	(.093)	(.093)	(.094)	(.083)	(.083)	(.084)
Age in months	.008***	.008***	.008***	.006**	.006**	.006**
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)
Illness	-.29***	-.293***	-.284***	-.24***	-.242***	-.225**
	(.099)	(.099)	(.1)	(.088)	(.088)	(.089)
Handicap	-.573	-.596	-.577	-.747*	-.766*	-.739*
	(.438)	(.439)	(.441)	(.418)	(.419)	(.42)
Household size	.008	.007	.01	.0006	.0005	.004
	(.018)	(.018)	(.018)	(.016)	(.016)	(.016)
Number of years in camp	-.064	-.054	-.058	.057	.066	.063
	(.104)	(.105)	(.105)	(.091)	(.092)	(.092)
Gender of household head	.151	.161	.15	.149	.157	.152
	(.107)	(.107)	(.108)	(.095)	(.096)	(.096)
Age of household head	.002	.002	.002	.003	.003	.003
	(.004)	(.004)	(.005)	(.004)	(.004)	(.004)
Member speaks french	.353***	.363***	.358***	.319***	.327***	.331***
	(.105)	(.106)	(.107)	(.094)	(.094)	(.095)
Assets and housing index	.002	.0003	.001	.004	.002	.003
	(.022)	(.022)	(.023)	(.02)	(.02)	(.02)
Number of cows	-.06	-.056	-.063	-.031	-.029	-.035
	(.038)	(.039)	(.039)	(.034)	(.034)	(.035)
IGA	.004	.002	-.007	-.003	-.005	-.016
	(.097)	(.097)	(.099)	(.087)	(.087)	(.089)
Cereal production	.009	.007	.008	.012	.01	.012
	(.013)	(.013)	(.013)	(.012)	(.012)	(.012)
Number of groups	10	10	10	10	10	10
Ethnicity dummies	yes	yes	yes	yes	yes	yes
Number of observations	616	616	607	607	607	598
R^2	.094	.095	.095	.096	.097	.101

Notes: Clustered standard errors are in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.10: Regressions of WHZ and HAZ

	(1) WHZ	(2) WHZ	(3) WHZ	(4) HAZ	(5) HAZ	(6) HAZ
Time since distribution	.015	-.135*	.014	-.006	.013	-.007
	(.011)	(.075)	(.011)	(.013)	(.09)	(.013)
Time since distribution^2		.002**			-.0003	
		(.001)			(.001)	
Cereal consumption of the week			-.009			.025
			(.013)			(.016)
Gender	.304**	.278**	.301**	.142	.145	.175
	(.119)	(.12)	(.12)	(.143)	(.144)	(.142)
Age in months	.003	.003	.004	-.026***	-.026***	-.025***
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)
Illness	-.199	-.206	-.185	-.259*	-.258*	-.194
	(.126)	(.126)	(.127)	(.151)	(.152)	(.151)
Handicap	.221	.167	.181	.069	.076	.076
	(.57)	(.569)	(.568)	(.723)	(.724)	(.729)
Household size	.036	.036	.032	.008	.008	.005
	(.023)	(.023)	(.023)	(.028)	(.028)	(.028)
Number of years in camp	-.23*	-.212	-.234*	.075	.072	.14
	(.133)	(.133)	(.133)	(.16)	(.16)	(.156)
Gender of household head	.059	.088	.056	.023	.02	.073
	(.14)	(.141)	(.14)	(.168)	(.169)	(.163)
Age of household head	.002	.001	.003	-.004	-.004	-.001
	(.006)	(.006)	(.006)	(.007)	(.007)	(.007)
Member speak french	.304**	.334**	.271**	.142	.138	.145
	(.134)	(.135)	(.135)	(.161)	(.162)	(.161)
Assets and housing index	.02	.015	.016	-.035	-.034	-.036
	(.029)	(.029)	(.029)	(.034)	(.034)	(.034)
Number of cows	-.079	-.068	-.07	.022	.02	.034
	(.049)	(.049)	(.049)	(.058)	(.059)	(.058)
IGA	-.152	-.159	-.165	.119	.119	.152
	(.124)	(.124)	(.125)	(.15)	(.15)	(.151)
Cereal production	.009	.003	.004	.012	.013	.012
	(.017)	(.017)	(.017)	(.02)	(.02)	(.019)
Number of groups	10	10	10	10	10	10
Ethnicity dummies	yes	yes	yes	yes	yes	yes
Number of observations	634	634	625	608	608	629
R^2	.065	.071	.064	.076	.076	.074

Notes: Clustered standard errors are in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2.11: Regressions of MUAC for age z-score (alternative specifications)

	(1) MUAC/A	(2) MUAC/A	(3) MUAC/A Amboko	(4) MUAC/A Gondjé	(5) MUAC/A Dosseye	(6) MUAC
Time since distribution		-.025***	-.026*	-.012	-.028**	-.03***
		(.007)	(.013)	(.013)	(.014)	(.01)
Second part of survey	-.247***					
	(.086)					
Gender	.147*	.13	.178	-.15	.347**	-.148
	(.083)	(.083)	(.136)	(.146)	(.152)	(.109)
Age in months	.006**	.006**	-.0008	.006	.013**	.03***
	(.003)	(.003)	(.005)	(.005)	(.005)	(.003)
Illness	-.237***	-.236***	-.248*	-.053	-.436**	-.32***
	(.088)	(.089)	(.142)	(.147)	(.175)	(.117)
Handicap	-.75*	-.723*	-.313	-2.239***	.27	-.56
	(.419)	(.42)	(.531)	(.731)	(1.109)	(.511)
Household size	.001	-.003	.048**	-.012	-.034	.003
	(.016)	(.016)	(.024)	(.031)	(.032)	(.021)
Number of years in camp	.051	.004		-.145	.05	.085
	(.092)	(.085)		(.203)	(.118)	(.119)
Gender of household head	.152	.102	.362**	.039	.046	.267**
	(.095)	(.094)	(.166)	(.168)	(.177)	(.127)
Age of household head	.003	.003	.007	-.013	.006	.003
	(.004)	(.004)	(.007)	(.009)	(.007)	(.005)
Member speaks french	.321***	.337***	.088	.449***	.213	.411***
	(.094)	(.093)	(.161)	(.16)	(.181)	(.125)
Assets and housing index	.004	.002	.029	-.02	.001	-.006
	(.02)	(.02)	(.04)	(.037)	(.034)	(.027)
Number of cows	-.033	-.036	.005	.021	-.086	-.011
	(.034)	(.034)	(.081)	(.061)	(.058)	(.045)
IGA	-.006	-.008	.16	.025	-.086	.051
	(.088)	(.088)	(.147)	(.157)	(.171)	(.115)
Cereal production	.011	.011	-.02	.01	.035	.007
	(.012)	(.012)	(.022)	(.019)	(.023)	(.015)
Number of groups	10	3	4	3	3	10
Ethnicity dummies	yes	yes	yes	yes	yes	yes
Number of observations	607	607	176	206	225	686
R^2	.094	.097	.212	.135	.152	.212

Notes: Clustered standard errors are in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%.

Chapter 3

Cognitive achievement production in Madagascar: a value-added model approach

This chapter was written in collaboration with David E. Sahn.[1]

3.1 Introduction

The rise in enrollment rates experienced by most developing countries over the previous decade has shifted the debate from the quantity to the quality side of education. Despite the paucity of test score data, a direct measure of achievement, the little evidence available from developing countries suggests that the gain in academic knowledge from schooling is quite limited (Jones et al. (2014)). This finding that there are small potential returns to schooling on learning have led some to suggest a limited response of earning to schooling, and conversely, that more attention needs to be given to the role of cognitive skills, rather than educational attainment per se, in determining labor market outcomes Boissiere et al. (1985). In the same way, Hanushek and Woessmann (2008) point out that focusing on grade attainment (instead of skills) would mislead our understanding of the relation between education and economic development. As Glick and Sahn (2009) and Behrman et al. (2008) emphasize, grade achievement is not sufficient to describe human capital accumulation: for a given grade, knowledge acquisition may differ depending on school and teaching quality as well as parental investment and the home environment which affects knowledge acquisition.

It is thus the issue of the extent to which schooling translates into cognitive achievement, or more specifically, what is the gain in knowledge of an additional year of schooling in the low

[1] Professor at Cornell University

income island nation of Madagascar, that motivates our paper. Glick and Sahn (2009) argue that the quantity of schooling, rather than family background, is the main factor in academic and non-academic skill formation in Senegal. Their research thus, suggests that parental inputs and school quality only have a limited and indirect effect. Behrman et al. (2008) also found a positive effect of schooling on reading skills in Guatemala, but demonstrate that pre-school (childhood nutrition) and post-school (employment) events also contribute to the accumulation of skills. They conclude that not controlling for extra-school factors would lead to overestimate the effect of grade attainment. More generally, it is crucial to understand how schools inputs and extra-school factors interact with schooling on the production of cognitive skills.[2]

Estimating the long and complex process of human capital accumulation requires a full history of parental and school inputs, as well as detailed information on children characteristics. Some of these factors, like parents' preferences or individual innate abilities, remain unobservable. Others, like the timing of schooling or attendance, result from a choice made by parents (or the student himself), based on their observation of their child's potential, the opportunities in the labor market and the quality of school infrastructure. Under the reasonable assumption that unobserved heterogeneity is correlated with the use and quality of specific inputs, ordinary least squares estimations would not provide consistent estimates (Todd and Wolpin (2003)). Therefore, the lack of data remains an important obstacle to the estimation of a cognitive production function, and thus much of the literature on the determinants of human capital is dedicated to the identifying ways to avoid biases and draw causal inference.

Panel data offers perhaps the best opportunity to model the dynamic dimension of learning. Among the most noteworthy efforts in this regard is the work of Cunha et al. (2010) who argue that it exists multiple "developmental phases" in a child's life, and exploit panel data to identify optimal periods of intervention, i.e., the periods when the use of specific inputs would be more efficient or, conversely, when negative shocks would be the most detrimental. Moreover, the availability of panel test scores allows for the use of individual fixed effects thus capturing time-invariant unobservable characteristics. This allows for the estimation of a value-added model (hereafter, VAM), where a measure of prior achievement is introduced in the model with the aim of controlling for the complete history of observable and unobservable inputs (Hanushek (1986)). Controlling for prior achievement changes the nature of the estimation: what is estimated is no longer the factors of final achievement but the determinants of the achievement gain since the baseline measure. In other words, the model isolates

[2]As Todd and Wolpin (2003) noted, the literature on the production of cognitive achievement focus either on the impact of school inputs (like class size (Angrist and Lavy (1999)), availability of textbooks (Glewwe et al. (2009)), infrastructure quality (Glewwe and Jacoby (1994))) or on the effect of households' characteristics. Das et al. (2013) demonstrate however that households adjust their own educational expenditure when they anticipate future school inputs, suggesting that schools' and household's inputs are substitutes.

the value-added of a unit of a specific input, all other thing held constant.[3]

While VAMs certainly helps capturing the effect of past unobserved inputs, this approach is not sufficient to get unbiased estimates. As Boardman and Murnane (1979) argued, a VAM can only be estimated without bias under the (strong) assumption that unobserved characteristics are not correlated with covariates while controlling for prior achievement and covariates. This issue of selection on unobservable covariates is also a major concern for Rothstein (2010) who pointed out that non-randomization in the allocation of school inputs (assignment of teachers in his study) would mislead the interpretation of the value-added of these inputs. In other words, the identification strategy would fail if the allocation of inputs were related to unobserved pupils' heterogeneity that is not captured by the lagged score. In our work, we thus are concerned that the schooling decision is likely to respond to unobserved children ability and, as noted by Singh (2013), an earlier test score doesn't capture the entire set of information parents observe before taking a decision on future investments in their children's human capital. A final concern about the dynamics of the education production function was raised by Andrabi et al. (2011) who added that introducing a measure of prior achievement only controls for time invariant effects of children innate ability whereas children with better endowment may learn faster.

In this paper, we are cognizant of all these concerns and formulate a strategy to take into account for the endogenity of additional schooling of young men and women since adolescence. More specifically, the questions we ask are: *(i)* How does overall schooling affect cognitive skills of young adults; *(ii)* How well do test scores of adolescents predict eventual cognitive achievement when they are young adults; and *(iii)* What is the contribution of additional years of education of a cohort of adolescents in Madagascar who we revisit some eight years later when they are young adults?

The remainder of the paper is structured as follows: the next section presents the conceptual framework, section 3.3 details the context and data, section 3.4 outlines our estimation strategy and section 3.5 presents the results.

3.2 Model

Following Todd and Wolpin (2003) and Sass et al. (2014), we consider a simple model of knowledge acquisition where the school achievement of a student i ($A_i(t)$) is expressed as

[3]VAM-based studies have been widely used in the United States to assess the value-added of a specific teacher by estimating the evolution of student's scores over an academic year (See Hanushek (1979), Hanushek (1986) or Harris and Sass (2011)). Rothstein (2010) argues however that the lack of randomization in the sorting of students into schools is likely to introduce a bias in the estimates. In developing countries, Hanushek et al. (2008) is a recent and elegant use of a VAM approach. They compute a school quality index by estimating the average value-added of a school on pupils achievement gain and find that pupils enrolled in better schools are less likely to drop out.

a function of school inputs ($S_i(t)$ and $Q_i(t)$), family inputs ($F_i(t)$) and the time-invariant student endowment (μ_{i0}). In the same manner as Behrman and Birdsall (1983) or Glick and Sahn (2009), school inputs can be decomposed into years of schooling ($S_i(t)$) and school characteristics ($Q_i(t)$). The subscript t refers to the specific period in the life of a student when achievement is measured and the time when inputs enter into production of cognitive skills. At each period, achievement at time t can be expressed as a function of past inputs:

$$A_{it} = g_t(S_i(t), Q_i(t), F_i(t), \mu_{i0}) \tag{3.1}$$

Family and school inputs result from choices made by teachers, parents, as well as the student himself. It can be depicted as:

$$F_{it} = \phi_t(A_i(t), W_i(t), Q_i(t), \mu_{i0}) \tag{3.2}$$

$$S_{it} = \theta(A_i(t), W_i(t), F_i(t), Q_i(t), \mu_{i0}) \tag{3.3}$$

$$Q_{it} = \omega(A_i(t), W_i(t), \mu_{i0}) \tag{3.4}$$

where $W_i(t)$ represents the household resources.[4] Schooling decisions and family inputs depend on the return to education expected in the labor market. With the aim of maximizing their children's cognitive achievement (and future earnings), parents choose the optimal level of inputs given the household budget constraint and depending on their observation of the child's achievement or endowment. School characteristics also play an important role in the allocation of family inputs and schooling decision. Firstly, parents can adjust their own investment according to the school quality (by compensating for the lack of textbooks for instance). Secondly, the schooling decision is presumably taken according to the expected achievement gain resulting from an additional year of schooling, which greatly depends on school quality, as observed by parents. School characteristics are also related to the household's resources as parents choose the school their child is attending, and in some cases, can have a direct impact on school quality. School inputs may also respond to the student's achievement if, for instance, better teachers were assigned to less endowed students or if they allocated more of their time to them.

Three main points should be noticed from this simple model of human capital accumulation. Firstly, inputs have both a direct and an indirect effect on achievement, through schooling duration and adjustments of family inputs. The impact of school quality on achievement, as measured by equation (3.1) is expressed net of its impact on schooling and family inputs. Secondly, the estimation of equation (3.1) requires the entire history of inputs. Not taking

[4]In this simple model, household resources are considered exogenous but this assumption may be violated, especially in developing countries, if the household income depends on child labor. Todd and Wolpin (2003) include the household's permanent resources rather than the time-varying income.

past inputs into account would lead to inconsistent estimates of the impact of contemporaneous inputs. Thirdly, child endowment, plays a crucial role in the production of cognitive achievement as well as in the allocation of inputs, raising concerns about potential endogeneity issues. Methods to address these concerns are presented in section 3.4.

3.3 Context and data

3.3.1 The Malagasy context

The challenges faced by the education sector in Madagascar are in many ways similar to the rest of Sub-Saharan Africa. Following the abolition of school fees in 2002, which was soon followed by a policy to provide school supplies to all primary school children, the gross primary enrollment rate has substantially increased (from 99% in 2000 to 145% in 2012, according to the World Development Indicators) and is now largely above the regional mean (110%). This increase in enrollment was closely linked to the large increase in the number of primary schools.[5] Additionally, there has been a trend toward a growing number of private schools, implying that parents may face more than one primary schooling option, even in rural communities. In contrast, the number of lower and higher secondary schools remains very limited, especially in remote areas, which has contributed to a much slower increase in secondary school enrollment.

The good news about improvement in enrollment numbers, especially at the primary level, is tempered by the heterogeneous quality of school infrastructure. Multi-grade classes, overcrowded classrooms and schools with an incomplete cycle are common. Also, the associations of parents have been given license to hire a teacher locally when needed. While this policy has facilitated the expansion of primary schools in remote areas, despite the shortage of trained teachers, the educational and pedagogical skills of those hired at the local level, often without any formal credentials, is worrying. As Michaelowa (2001) emphasize, schooling in Madagascar is characterized by high repetition and dropout rates. As a consequence, the primary completion rate remains worryingly low, only 60%; and only a small proportion of those who do complete primary school enroll in secondary school (World Bank (2008)). In part this may be attributed to the inaccessibility of secondary schools. But it also reflects the poor performance of many primary school completers on the national test which needs to be passed in order to advance to secondary school.

[5] According to the Malagasy Ministry of Education, there were 21'877 public primary schools registered during the school year 2010/2011, whereas there were only 12'730 ten years before.

3.3.2 Data

The data used in our analysis are from the *Madagascar Life Course Transition of Young Adults Survey* in 2011-2012, which follows a cohort of Malagasy adults born in the late 1980s, most of whom were 21 to 23 years old when they were interviewed in this latest survey round. The cohort was last surveyed in 2004-2005 as part of the *Progression Through School and Academic Performance in Madagascar Study* (EPSPAM), at which time they were adolescents.[6] The national sample were collected in 50 districts/fivondrana (over 116) from the six regions of the country. The sample design involved returning to 48 communities from a larger education study, the *Programme d'Analyse des Systèmes Educatifs de la CONFEMEN* (PASEC), which was conducted in 1998 in 120 clusters defined by the catchment areas of primary schools in which scholastic aptitude tests took place. While our sample was selected randomly from these clusters, the original PASEC communities themselves were randomly selected from schools with at least 20 students in both the second and fifth grades. Given schools are typically small in rural areas, the rural PASEC clusters thus tended to be larger-than-average communities. To partially address this issue, the 2004 survey supplemented the 48 PASEC clusters with an additional 12 clusters randomly selected from rural communities with small primary schools.[7] These additional schools were randomly selected from the list of schools in the education ministry database after stratifying on province. In each of the original and new clusters, we also did a complete enumeration of all the children in the cohort's age range, and randomly selected an equal number of students who were not in the original PASEC sample. This was to make sure that we did not exclude those who never attended school, or enrolled very late, which is not an uncommon occurrence in Madagascar. Thus, the 2004-5 and 2011-12 samples includes cohort members who would not have been selected by the original school-based survey (e.g. CM who dropped out of school early or who never attended) and is therefore more nationally representative of cohort.

Besides being a long-term panel, the other characteristic of this dataset that makes it well suited for our analysis is that we collected detailed retrospective information about the cohort members, their households and the availability and quality of school infrastructures in both the 2004-5 and 2011-12 survey rounds. Additionally, cohort members were given tests to assess their knowledge in mathematics, French and their 'life skills'. Our sample consists of 989 young adults who took at least one of the three tests during each of the two rounds of survey. Additionally, the household surveys were accompanied by both school and community based surveys that provide details on the nature and characteristics of existing infrastructure. In total, 140 schools were interviewed in 2011-12 in the 73 clusters (i.e., communities) where the survey was conducted. Information was gathered on the experience and credentials of the

[6]See Glick et al. (2011a) and Glick et al. (2011b) for a detailed description of the 2004-2005 survey.

[7]"Small" was defined as a school having fewer students than the national median of about 140.

principal and management of the school, as well as the number of teachers, their qualifications and pedagogical practices, and building and classroom conditions.

For the aim of this study, we focus only on the last two rounds of survey (2004-2005 and 2011-2012) and consider the former as our baseline.[8] Nearly 89% of the targeted cohort members were interviewed in the last round, which makes the attrition rate (11%) remarkably small as compared to other individual-based panel data collected in Sub-Saharan Africa.

As shown in Table 3.1, the panel encompasses the period of time in the life-course when many are deciding whether to drop out of school: while 83% of cohort members in our sample were enrolled at school in 2004, only 23% were still enrolled in 2011. In 2004, the average level of schooling was 5.3 years. Only 45% of the cohort members had reached a grade beyond primary school and less than 2% of them made it to higher secondary school. The average number of additional years of schooling is 2.8 years, making the average higher completed grade increase to 8.1 years and the standard deviation rise from 2.1 in 2004 to 3.5 in 2011.

As documented by Glick et al. (2011b), academic achievement is positively related with the household's wealth and parent's education (Table 3.1). The gradients between both wealth and parental education, and grade attainment, have persisted over the two rounds of the survey. For instance, cohort members whose mother has no education are likely to have experienced 1.8 additional years of schooling between the two surveys while cohort members whose mother has reached secondary school (at least) have experienced 4.1 additional years of schooling. Although this relationship cannot be interpreted as causal, it suggests a strong intergenerational transmission of disparities through the production of education.

Finally, there is a lack of gender bias in the educational outcomes in contrast to what is usually observed in the African context. For example, there is no difference between male and female cohort members in terms of their highest grade completed. Interestingly, the dynamic in the grade progression is not identical: male cohort members, who were enrolled in lower grade on average in 2004, caught up with the female members during their transition to adulthood. This finding is consistent with the high prevalence of early pregnancy amongst Malagasy teenagers leading to early dropout (Herrera and Sahn (2013)).

3.3.3 Cognitive tests

As previously noted, grade achievement is not sufficient to analyze the formation of human capital and, when available, tests scores offer a better option since they are a more direct and precise measure of knowledge and skills. In the case of our study, cohort members were tested in 2004-5 and 2011-12 for their knowledge in mathematics and French. Numeracy and literacy are the two main academic outputs of schooling, but in addition, we tested the life

[8]Restricting our sample to the cohort members interviewed three times would considerably reduce its statistical power and would raise issues about representativeness and selectivity.

skills of the cohort members using a set of questions designed to capture less academic and practical knowledge.[9] The tests were designed by the Ministry of Education with the aim of obtaining a comparable measure of cognitive achievement among a heterogeneous population. Tests were identical for each cohort member, regardless of their grade achievement, starting with questions on basic knowledge and getting progressively more complicated. They were composed of an oral and a written part, so illiterate cohort members (a common situation even after several years of schooling) were also given the test. Tests were administered to cohort members at home, which mitigates somewhat the concern that the scores may be biased by intra-classroom correlation. Internal consistency of each test, as expressed by the Cronbach's alpha, was above the recommended standard: ranging from 0.82 to 0.92 for academic skills, and from 0.74 to 0.82 for life skills.

Test scores were standardized with zero mean and standard deviation equal to one. Table 3.3 presents the baseline and follow-up raw test scores as well as the standardized scores by family background. Scores follow the same patterns observed with the progression through grades: they are positively correlated with parents' education and household's wealth, but negatively correlated with the remoteness of the cluster in which the children reside. Moreover, the correlation between scores within rounds is strong and highly significant. In other words, cohort members who performed better at math tend to have a better score in French too. This result is partly driven by heterogeneity in grade attainment; but correlation coefficients remain high while controlling for final grade (respectively 0.56, 0.59 and 0.57 for sixth, tenth and thirteenth grade). Not surprisingly, the correlation between life skills and academic scores is lower but still significant.

We also observe a strong intertemporal correlation between scores. As shown in Table 3.4, cohort members who were in the lowest baseline score quartile performed poorly at the follow-up test. This result could either reflect a persistent effect of past knowledge or a higher grade progression for better students. Cohort members from the highest baseline score quartile have experienced an average 3.7 additional years of schooling while cohort members from the lowest quartile have only experienced 1.7 years (Table 3.2). Moreover, higher grade attainment is associated with better cognitive achievement both in 2004 and 2011, and for all the skills that we measure.

[9] The life skills test features questions on behavior related to health, nutrition or civic issues.

3.4 Estimation strategy

3.4.1 Three models of knowledge acquisition

With the aim of taking advantage of the panel dimension of our database, we consider two periods in the production of cognitive skills of the cohort member i : the first period ($t = 1$) corresponds to the accumulation of skills that happens before 2004, the time prior to when the cohort member was interviewed at teenage years; the second period ($t = 2$) corresponds to the transition period from the teenage years to young adulthood (including additional schooling since 2004). As depicted by equation (3.1), family inputs, years of schooling and school quality enters the production function at each period. We present three distinct models to estimate the determinants of school achievement at adulthood (A_{i2}), as measured by our four test scores. Our first model is to estimate a naive model where the final score is expressed as a function of grade achievement (G_{it}), family inputs (F_{it}) and school inputs (Q_{it}), while controlling for the cohort member's own characteristics (C_{it}). The equation is given by:

$$A_{i2} = \beta_1 G_{i2} + \beta_2 F_{i1} + \beta_3 F_{i2} + \beta_4 Q_{i1} + \beta_5 Q_{i2} + \beta_6 C_{j1} + \beta_7 C_{j2} + \mu_{i0} + \varepsilon_{i2} \qquad (3.5)$$

where μ_{i0} is the cohort member's time-invariant (unobservable) endowment and ε_{i2} a random disturbance term. Estimation of β_1 gives the direct contribution of grade achievement to the formation of skills, all other things held constant. However, its estimation raises two serious concerns, both of which plague the vast majority of empirical studies that model education attainment and cognitive outcomes. First, even with the retrospective panel data we collect, if the allocation of inputs over time is serially correlated (i.e. if inputs are not distributed randomly at each period), estimation of the effect of contemporaneous inputs will be inconsistent, as they would partly reflect the allocation of past inputs. Contemporaneous parents' pedagogical support, for instance, would capture the effect of past support that may be of relatively greater importance. Neglecting the effect of past inputs in the production of cognitive achievement is only conceivable under the strong assumption that they have no long-lasting effect on achievement at time t. Second, inputs are allocated with regard to the student's characteristics and endowment. If inputs and endowments are positively correlated, failure to control for μ_{i0} would lead the estimates of inputs' parameters to be biased upward. In our case, if parents who observe a high potential in their child invest more than other parents, the effect of grade achievement on skills acquisition would be overestimated.

In order to control for unobserved heterogeneity, we use a VAM approach by introducing a measure of prior achievement (A_{i1}), measured at the end of the first period, into equation (3.1). A_{i1} is expected to capture the effect of all inputs, observable or not, that enter into its production. Using prior achievement transforms the production function as it estimates the *value-added* of the allocation of a specific input during the time interval between the two tests. In other words, for a given baseline achievement, the gain in knowledge is only attributed to the contemporaneous inputs. Our intent, and strategy, differ somewhat from the traditional VAM. First, our major concern is not with a specific input or school or classroom characteristic, as is often the case in the schooling literature in developed countries, but instead we are interested in the more general question of whether grade attainment and schooling contributes to skills and human capital accumulation. The second, and related difference, is that we are not focused on the value added of a specific year of schooling, or the progression, for example, from 9th to 10th grade, but more generally, the impact of additional years of schooling during the intervening period between our two surveys. Thus, we estimate two additional models. First, we express final knowledge as a function of prior achievement (final grade being excluded) to assess how path-dependent knowledge acquisition is. This model is given by:

$$A_{i2} = \beta_2 F_{i1} + \beta_3 F_{i2} + \beta_4 Q_{i1} + \beta_5 Q_{i2} + \beta_6 C_{j1} + \beta_7 C_{j2} + \beta_8 A_{i1} + \mu_{i0} + \varepsilon_{i2} \tag{3.6}$$

A parameter β_8 significantly different from zero would indicate a persistent impact of baseline achievement, suggesting that past inputs have long-lasting effects on learning.[10] In our second value-added model we add to model (3.5) an additional regressor, (S_{i2}), that captures years of schooling since baseline achievement in 2004-05:

$$A_{i2} = \beta_1 S_{i2} + \beta_2 F_{i1} + \beta_3 F_{i2} + \beta_4 Q_{i1} + \beta_5 Q_{i2} + \beta_6 C_{j1} + \beta_7 C_{j2} + \beta_8 A_{i1} + \mu_{i0} + \varepsilon_{i2} \tag{3.7}$$

As noted by Todd and Wolpin (2003), the conditions under which the estimation of equation (3.6) and (3.7) provides consistent estimates of inputs' parameters by OLS are likely to be violated. Firstly, OLS estimation requires inputs to be uncorrelated with the student's unobserved endowment, conditional on prior achievement (orthogonality condition). The enrollment decision is, however, not random and likely to be taken in accordance with the student's performance and innate ability, e.g. parents would be more prone to promote additional schooling for high-ability students. Rothstein (2010) and Singh (2013) likewise point out that baseline achievement may not capture the entire factors that lead the inputs allocation i.e. the schooling decision. A second concern is the potential for correlation between the lagged score and the error term, as the latter includes the unobservable student's endow-

[10] In the literature, the lagged score is sometimes introduced in the left-hand side of the equation ($A_{it} - A_{it-1}$). This specification is more restricted as it posits that $\beta_8 = 1$, which implies that prior achievement has a complete persistent effect on final achievement, i.e. that past inputs' effect doesn't decay with time. Moreover, Hanushek et al. (2008) argue that this restriction is not appropriate if the two tests are not on the same scale, as in our case.

ment. Under the assumption that the students' endowment enters the production function, the lagged score should not be taken as exogenous. And finally, it is worth noting that measurement errors in test scores are expected to play a role in the estimation by biasing downward the parameter of lagged achievement.

3.4.2 Endogeneity issues

We address the biases introduced by OLS by employing an instrumental variable strategy to control for the endogeneity in all three models. Specifically, in the first model we instrument final grade, in the two value added models we instrument baseline achievement, as well as additional schooling in the second VAM in which it is included. The challenge is identifying instruments that have a direct impact on the three endogenous variables but no direct impact on final achievement, conditional on the other covariates included in the estimation. It is instructive to consider what has been done in the literature, despite the paucity of studies that make a serious attempt to deal with these problems. Perhaps the most similar analysis is the work of Behrman et al. (2008) in which they consider final grade (along with nutrition and work experience) as endogenous to the adult cognitive skills production function. They use the student's school characteristics and economic shocks as instruments to predict higher completed grade. One option we thus considered was to similarly use, the characteristics of the classroom's attended as instrument, but concluded that this would not be appropriate since we suspect school quality to have a direct and long-lasting impact on learning. Better teachers may use specific pedagogical methods which can affect the way students will learn in future periods, that is to say, how efficiently students will take advantage of their own endowment. Thus, the variables are simply included as covariates, serving as controls in the models.

Instead, we use instruments information on the availability of school infrastructure and the economic conditions of the community, gathered in the 2004 survey. That is, the details on the schools and communities is from the period that precedes the measurement of the cohort member's baseline knowledge in 2004, and is from where the cohort member was residing at that time. To be more precise, we use the number of secondary schools established in the commune and the number of years since the first primary school opened as instruments in our first two models, in which only educational attainment is instrumented. In 2004, before the first round of tests, a high-secondary school was available in only 42% of the sampled communes. In 24% of them, at least two secondary schools were available. The average number of years since a primary school was available in the commune was 43 years. The two instruments don't reflect the quality of the specific school attended, which would be a problem in terms of the exclusion restriction criteria; but instead, the instruments capture the availability of school infrastructures, which has a direct impact on schooling decision. For a cohort member, the lack of a secondary school in his commune means higher distance to the nearest school, and thus a higher cost to reach it and attend secondary school. It can have a large disincentive impact on enrollment and progression even through earlier grades, especially in a Malagasy context where parents might be reluctant to let their children leave the household. That being said, the information on the availability of secondary schools is unlikely to have a direct impact on knowledge acquisition. Similarly, the date of first presence of a primary school in the commune could be considered as exogenous to cohort members' learning achievement as they were usually constructed prior to the MCs entering school.[11] It can be seen as a proxy for the government investment in the area or the conditions the parents of cohort members were facing as students, since we know that there was little migration among the parents.

In the value added model with two endogenous variables are instrumented, we employ additional instruments: the share of households in the community who declare having a deposit and the share of those who had a loan in 2004. These two variables serve as proxies for the development of the local credit and insurance market. Jacoby and Skoufias (1997), in the context of rural India, and Sahn et al. (2012) and Gubert and Robilliard (2008) in a Malagasy context, have shown that, in the presence of an imperfect credit markets, unanticipated income shocks lead to lower school attendance through the use of child labor. The credit market conditions that cohort members' parents were facing can therefore have an impact on the learning and grade progression of their children.

All four instruments are capturing the commune level characteristics, not choices of the individual households. The one potential draw-back of our IV strategy is that it relies on the assumption that parents have not chosen to settle in (or leave their child to a household leaving in) a particular commune according to the availability of school infrastructure. At the time of the baseline survey, a large majority of cohort members' household (89%) have always lived in the commune where they were interviewed, but we cannot rule out that the remaining 11% moved in a new commune in order to get access to secondary school infrastructures. Moreover, amongst the 9.4% cohort members who were not living with their parents in 2004, only 22% of them (less than 2% of the sample) have lived in more than one commune since their birth, suggesting than child fostering related to the access of school infrastructures is not common in our sample.

[11]In only one commune (1% of our sample), the creation of the first primary school, as reported by our community questionnaire, is subsequent to official age at first entry of the cohort we study.

3.5 Results and discussion

The three models are estimated for each of the four pairs of score (written French, written math, oral math and life skills). Estimation results are presented in Tables 3.5 to 3.12. Time-varying family background variables (wealth index, gender of the household head, the share of household members under sixteen years old) and the location details (urban area, remoteness index) refers to the baseline situation.[12] About half cohort members were found living in a new household in 2011. As this choice cannot be considered as exogenous, details about the cohort members' current situation (marital status or fertility) were not included. The effect of primary school quality is captured by a facility and teaching material index and by the proportion of teachers with more than lower secondary education. Because parents can choose which school their child attends, schools characteristics are expressed as the commune mean value, reflecting the average quality of the available schools.[13] As some of the key covariates and instruments are expressed at the community level, we don't use a fixed effect method to control for the unobserved heterogeneity. However, standard errors are clustered by localities.

Student's height (measured during the baseline survey and standardized by age and gender) is introduced to control for nutritional deficiencies that could have affected the student's growth during early childhood. This parameter should be interpreted cautiously as we don't control for its potential endogeneity (Behrman et al. (2008)).

OLS estimations of the first model, presented in Tables 3.5 and 3.6, show a large and strong effect of grade attainment on the acquisition of cognitive skills, whatever skill considered. Each additional grade is associated with a 0.19 standard deviation increase in the French score (which corresponds to a 12.4% increase in the average score). Interestingly, higher grade is also associated with the so-called "life skills", a measure of non-academic knowledge. Consistent with Glick et al. (2011b), the effect of family background appears modest in comparison to the impact of schooling: the effect of father's education is not significantly different from zero; the wealth index, which could be considered as a proxy for the parental pedagogical inputs, only shows a direct effect on the French score. When grade attainment is instrumented the estimated parameter associated with grade gets even larger (for the three academic scores but not for life skills). Each additional grade is associated with a 17.7% increase in the written French score. This result indicates that OLS estimates un-

[12]Following the methodology of Sahn and Stifel (2003), the wealth index is based on the ownership of durable goods (such as a radio, TV, bicycle or a lamp), the source of drinking water and type of toilet facilities. The remoteness index is based on the availability of health services, banks, post offices, schools, markets and the access to roads.

[13]The facility and teaching material index was created using a factor analysis. The characteristics used for the computation of the index included the number of rooms, the share of classrooms with a blackboard, whether there are latrines, an infirmary or a sport playground, whether electricity and drinking water are available.

derestimated the real impact of grade attainment on cognitive skills. Conversely, the direct effect of mother's education becomes non-significant. While family background appears to have nearly no direct effect on the acquisition of cognitive skills once controlled for grade attainment, parents' education and household's wealth are strong predictors of the final grade.

The two instruments have a large and significant effect on final grade attainment. As expected the presence of a higher secondary school in the commune play a positive role the grade progression. The number of years that a primary school is available, a proxy for the government investment in the commune school infrastructures, is also positively correlated with grades. The instruments are valid according to the Hansen J-test statistics and the F-statistics on excluded instruments.

The estimation of the third model, presented in Tables 3.9 to 3.12, show the contribution of additional years of education on the accumulation of cognitive skills. For a given level of cognitive skills acquired in 2004, one additional year of schooling corresponds to an 11.5% increase in the written French score according to the OLS estimates, which is very similar in amplitude to the results of the first model. Once we control for endogeneity, the impact of an additional grade is twice as large (+23%), suggesting that schooling is the main factor of the acquisition of skills.

Beside the availability of schools, which we use as instruments, the quality of school infrastructures and the qualification of teachers have a persistent effect on the production of cognitive skills. However, the positive effects of primary schools quality and more educated teachers disappear in the instrumented third model, which could indicate that the persistent effect of primary school fades out with time.

The effect of economic shocks on the acquisition of skills is particularly interesting. The number of bad harvests experienced by households before and after the baseline test has a negative impact on both written test scores. Good harvests are also associated with better scores in French and life skills. This effect is present in the three models. This should be put in line with the findings of Gubert and Robilliard (2008) who highlight that unexpected reduction in households' resources increase the likelihood to drop out of school, arguing that parents use child labor to cope with shocks. Then, the likely reduction of a child attendance at school over the year, as a result of a cut in parents' resources, is one possible explanation for the direct effect of shocks on the production of cognitive skills. This suggestive evidence of household's sensitivity to economic shocks could also explained the positive effect the standardized height on mathematics and life skills scores, though this relationship could not be interpreted as causal without a correction for its potential endogeneity.

Finally, the parameter associated with lagged score parameter in the estimation of model 2 and 3 show how do lagged test scores predict eventual cognitive achievement when they are young adults. Tables 3.7 and 3.8 present the estimation of model 2 where no indicator of grade attainment is included. OLS estimates indicate a positive effect of the lagged score, suggesting that knowledge acquisition is path dependent. In other words, we observe acceleration in learning: adolescents who had accumulated more knowledge in 2004 performed better in 2011. This result is confirmed and accentuated for the both written tests once we instrument lagged achievement, but not for the oral mathematics score and life skills. Once we control for grade attainment in model 3, the effect of prior achievement is mixed. The parameter remains positive and significant when we regress the model with OLS but is only significant for the French score in the IV estimates.

Bibliography

ANDRABI, T., DAS, J., KHWAJA, A. I., AND ZAJONC, T. 2011. Do value-added estimates add value? Accounting for learning dynamics. *American Economic Journal: Applied Economics* 3:29–54.

ANGRIST, J. D. AND LAVY, V. 1999. Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics* 114:533–575.

BEHRMAN, J. R. AND BIRDSALL, N. 1983. The quality of schooling: Quantity alone is misleading. *The American Economic Review* pp. 928–946.

BEHRMAN, J. R., HODDINOTT, J. F., MALUCCIO, J. A., SOLER-HAMPEJSEK, E., BEHRMAN, E. L., MARTORELL, R., RAMIREZ-ZEA, M., AND STEIN, A. D. 2008. What determines adult cognitive skills? Impacts of pre-schooling, schooling and post-schooling experiences in Guatemala. IFPRI Discussion Paper 826, International Food Policy Research Institute.

BOARDMAN, A. E. AND MURNANE, R. J. 1979. Using panel data to improve estimates of the determinants of educational achievement. *Sociology of Education* pp. 113–121.

BOISSIERE, M., KNIGHT, J. B., AND SABOT, R. H. 1985. Earnings, schooling, ability, and cognitive skills. *The American Economic Review* pp. 1016–1030.

CUNHA, F., HECKMAN, J. J., AND SCHENNACH, S. M. 2010. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78:883–931.

DAS, J., DERCON, S., HABYARIMANA, J., KRISHNAN, P., MURALIDHARAN, K., AND SUNDARARAMAN, V. 2013. School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics* 5:29–57.

GLEWWE, P. AND JACOBY, H. 1994. Student achievement and schooling choice in low-income countries: Evidence from Ghana. *Journal of Human Resources* pp. 843–864.

GLEWWE, P., KREMER, M., AND MOULIN, S. 2009. Many children left behind? Textbooks and test scores in Kenya. *American Economic Journal: Applied Economics* 1:112–135.

Glick, P., Handy, C., and Sahn, D. E. 2011a. Schooling, marriage, and childbearing in Madagascar. *Cornell Food and Nutrition Policy Program Working Paper* .

Glick, P., Randrianarisoa, J. C., and Sahn, D. E. 2011b. Family background, school characteristics, and children's cognitive achievement in Madagascar. *Education Economics* 19:363–396.

Glick, P. and Sahn, D. E. 2009. Cognitive skills among children in Senegal: Disentangling the roles of schooling and family background. *Economics of Education Review* 28:178–188.

Gubert, F. and Robilliard, A. S. 2008. Risk and schooling decisions in rural Madagascar: A panel data-analysis. *Journal of African Economies* 17:207–238.

Hanushek, E. A. 1979. Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources* pp. 351–388.

Hanushek, E. A. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature* 24:1141–1177.

Hanushek, E. A., Lavy, V. C., and Hitomi, K. 2008. Do students care about school quality? Determinants of dropout behavior in developing countries. *Journal of Human Capital* 2:69–105.

Hanushek, E. A. and Woessmann, L. 2008. The role of cognitive skills in economic development. *Journal of Economic Literature* pp. 607–668.

Harris, D. N. and Sass, T. R. 2011. Teacher training, teacher quality and student achievement. *Journal of Public Economics* 95:798–812.

Herrera, C. and Sahn, D. 2013. The impact of early childbearing on schooling and cognitive skills among young women in Madagascar. *Cornell University Food and Nutrition Policy Program Working Paper* .

Jacoby, H. G. and Skoufias, E. 1997. Risk, financial markets, and human capital in a developing country. *The Review of Economic Studies* 64:311–335.

Jones, S., Schipper, Y., Ruto, S., and Rajani, R. 2014. Can your child read and count? Measuring learning outcomes in East Africa. *Journal of African Economies* p. eju009.

Michaelowa, K. 2001. Primary education quality in francophone Sub-Saharan Africa: Determinants of learning achievement and efficiency considerations. *World Development* 29:1699–1716.

Rothstein, J. 2010. Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics* 125:175–214.

Sahn, D. E., Peter, G., Thomas, F., et al. 2012. Household shocks and education investment in Madagascar. CERDI Working Paper N° 201227, CERDI, Université d'Auvergne.

Sahn, D. E. and Stifel, D. 2003. Exploring alternative measures of welfare in the absence of expenditure data. *Review of Income and Wealth* 49:463–489.

Sass, T. R., Semykina, A., and Harris, D. N. 2014. Value-added models and the measurement of teacher productivity. *Economics of Education Review* 38:9–23.

Singh, A. 2013. Size and sources of the private school premium in test scores in india. *Young Lives Working Paper* 98.

Todd, P. E. and Wolpin, K. I. 2003. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal* 113:F3–F33.

World Bank 2008. The challenge of expanding secondary education and training in madagascar. Technical report, World Bank.

Table 3.1: Summary statistics

Variable	Mean	Std. Dev.	N
Cohort members' characteristics			
% Enrolled at school in 2004	0.830	0.376	941
% Enrolled at school in 2011	0.227	0.419	941
% of CM who reached lower secondary at least in 2004	0.458	0.498	941
% of CM who reached lower secondary at least in 2011	0.675	0.469	941
% of CM who reached higher secondary at least in 2004	0.014	0.117	941
% of CM who reached higher secondary at least in 2011	0.370	0.483	941
Highest grade attained in 2004	5.323	2.109	941
Highest grade attained in 2011	8.055	3.545	941
Grade progression	2.762	2.242	941
Gender (Male=1)	0.471	0.499	941
Age in 2011	21.832	1.196	941
Height in cm	159.801	7.947	921
Household's background			
Mother's education (in years)	4.850	3.582	935
Father's education (in years)	5.461	3.904	935
Asset index	0.060	0.984	941
Gender of household head	0.806	0.396	941
Share of household members under sixteen years old	0.513	0.179	941
Community and primary schools characteristics			
Urban dummy (urban=1)	0.244	0.430	941
Remoteness index	2.416	1.332	929
Primary school facility and teaching material index	−0.004	0.689	931
Proportion of teachers with more than lower secondary education	0.551	0.277	931
Shocks experienced by households			
Number of unusually good harvest before 2004	0.448	0.842	941
Number of unusually bad harvest before 2004	0.437	0.856	941
Number of unusually good harvest after 2004	0.606	0.900	941
Number of unusually bad harvest after 2004	0.746	1.044	941
Number of years of poor household enterprise performance after 2004	0.247	0.700	941
Number of years of good household enterprise performance after 2004	0.256	0.750	941
Number of years with livestock loss after 2004	0.326	0.579	941
Instruments			
Number of higher secondary schools in the commune (in 2004)	0.804	1.121	941
Number of years since first primary school opened	45.565	23.689	931
Share of households with a deposit (in 2004)	0.084	0.098	941
Share of households who have contracted a loan (in 2004)	0.085	0.077	941

All variables refers to the baseline survey (2004), unless otherwise stated.

Table 3.2: Progression through grades since teenage years
by family background

	Grade 2004	**Grade 2011**	**Grade progression**
Sample mean	5.323	8.055	2.762
Standard deviation	2.109	3.545	2.242
Gender			
Female	5.446	8.026	2.610
Male	5.185	8.088	2.932
Household wealth quartile			
1	3.541	5.095	1.641
2	5.152	7.667	2.532
3	5.704	8.728	3.037
4	6.734	10.461	3.734
Mother's education			
No education	3.854	5.606	1.818
Some primary	4.821	7.070	2.279
Completed primary	5.204	7.885	2.701
Some lower secondary	5.976	9.373	3.405
Completed lower secondary	6.681	10.245	3.617
Some higher secondary	6.718	10.846	4.128
Completed higher secondary	7.541	11.622	4.081
Above secondary	8.200	13.333	5.133
Father's education			
No education	3.853	5.618	1.873
Some primary	4.623	6.846	2.246
Completed primary	5.315	7.869	2.592
Some lower secondary	5.844	8.510	2.688
Completed lower secondary	6.299	9.907	3.617
Some higher secondary	6.400	9.750	3.367
Completed higher secondary	7.297	11.891	4.594
Above secondary	7.231	11.962	4.731
Remoteness index			
1	6.231	9.477	3.258
2	5.771	8.751	3.005
3	4.851	7.124	2.323
4	4.543	6.939	2.451
5	3.716	5.338	1.649
Baseline French score quartile			
1	3.900	5.560	1.735
2	5.076	7.571	2.510
3	5.843	8.975	3.152
4	6.769	10.443	3.686

The wealth quartiles are ranked from the poorest (1) to the wealthiest (4). The remoteness index ranks households from the less remote (1) to the more remote (5).

Table 3.3: Baseline and adult tests scores by family background

	Baseline test scores (2004)				**Follow-up test scores (2011)**			
	Written French	**Oral math**	**Written math**	**Life skills**	**Written French**	**Oral math**	**Written math**	**Life skills**
Raw score (% correct)	0.527	0.465	0.522	0.728	0.474	0.526	0.390	0.735
Standard deviation	0.192	0.249	0.220	0.201	0.284	0.305	0.223	0.185
N	863	897	807	908	881	892	898	939
Standardized score by:								
Wealth quartile								
1	−0.644	−0.585	−0.740	−0.591	−0.758	−0.672	−0.644	−0.542
2	−0.222	−0.128	−0.157	−0.155	−0.159	−0.104	−0.093	−0.021
3	0.165	0.229	0.199	0.208	0.181	0.253	0.219	0.266
4	0.558	0.464	0.537	0.466	0.715	0.602	0.560	0.465
Mother's education								
No education	−0.614	−0.412	−0.488	−0.561	−0.584	−0.515	−0.541	−0.366
Some primary	−0.202	−0.175	−0.202	−0.163	−0.225	−0.200	−0.199	−0.171
Completed primary	−0.116	−0.068	−0.124	−0.046	−0.068	0.046	0.039	0.107
Some lower secondary	0.385	0.358	0.285	0.359	0.326	0.387	0.346	0.320
Completed lower sec.	0.266	0.277	0.397	0.355	0.593	0.540	0.487	0.468
Some higher secondary	0.595	0.561	0.566	0.560	0.687	0.531	0.554	0.488
Completed higher sec.	0.913	0.696	0.761	0.553	1.139	0.947	0.967	0.664
Above secondary	1.715	1.041	1.116	0.870	1.509	1.113	1.278	0.889
Father's education								
No education	−0.640	−0.332	−0.326	−0.458	−0.489	−0.492	−0.450	−0.422
Some primary	−0.237	−0.199	−0.223	−0.170	−0.287	−0.191	−0.213	−0.163
Completed primary	−0.007	−0.050	−0.081	0.044	−0.066	−0.028	0.021	0.075
Some lower secondary	0.163	0.217	0.049	0.143	0.083	0.069	0.036	0.176
Completed lower sec.	0.246	0.298	0.280	0.192	0.484	0.545	0.525	0.395
Some higher secondary	0.254	0.261	0.342	0.277	0.557	0.488	0.325	0.399
Completed higher sec.	0.884	0.573	0.652	0.455	0.956	0.730	0.786	0.575
Above secondary	0.978	0.677	0.884	0.799	1.129	0.786	0.790	0.834
Remoteness index								
1	0.375	0.404	0.393	0.417	0.350	0.329	0.271	0.389
2	0.144	0.194	0.234	0.107	0.236	0.287	0.248	0.232
3	−0.412	−0.253	−0.535	−0.329	−0.187	−0.249	−0.148	−0.192
4	−0.485	−0.539	−0.397	−0.495	−0.181	−0.101	−0.139	−0.249
5	−0.296	−0.475	−0.393	−0.340	−0.910	−0.739	−0.661	−0.548

The wealth quartiles are ranked from the poorest (1) to the wealthiest (4). The remoteness index ranks households from the less remote (1) to the more remote (5).

Table 3.4: Baseline and adult tests scores by grade attainment

	Baseline test scores (2004)				Follow-up test scores (2011)			
	Written French	**Oral math**	**Written math**	**Life skills**	**Written French**	**Oral math**	**Written math**	**Life skills**
Baseline score quartile								
1					−0.618	−0.531	−0.549	−0.629
2					−0.200	−0.190	0.007	−0.027
3					0.216	0.323	0.101	0.191
4					0.663	0.416	0.544	0.446
Years of schooling (2004)								
0-1	−0.598	−0.753	−1.086	−0.808	−1.249	−1.323	−1.073	−0.949
2	−0.605	−0.389	−1.331	−0.842	−1.030	−0.953	−0.984	−0.714
3	−0.890	−0.925	−0.982	−0.989	−0.933	−0.742	−0.855	−0.681
4	−0.527	−0.451	−0.388	−0.225	−0.457	−0.416	−0.358	−0.210
5	−0.267	−0.049	−0.021	−0.029	−0.110	−0.028	−0.019	0.046
6	0.144	0.293	0.195	0.302	0.232	0.263	0.216	0.240
7	0.514	0.397	0.542	0.454	0.677	0.641	0.602	0.485
8	0.796	0.586	0.692	0.546	0.830	0.701	0.799	0.616
9 years and more	1.151	0.941	0.903	0.706	1.009	0.814	0.708	0.864
Years of schooling (2011)								
0-2	−0.632	−0.547	−0.942	−0.836	−1.321	−1.430	−1.282	−1.039
3	−0.670	−0.717	−1.051	−0.974	−1.082	−1.093	−1.048	−0.779
4	−0.670	−0.671	−0.901	−0.579	−0.939	−0.691	−0.835	−0.510
5	−0.546	−0.463	−0.528	−0.302	−0.738	−0.619	−0.609	−0.446
6	−0.421	−0.387	0.068	−0.060	−0.523	−0.260	−0.255	−0.071
7	−0.335	−0.019	−0.073	−0.044	−0.182	−0.215	−0.217	0.000
8	−0.171	0.090	0.002	−0.075	0.061	0.047	0.118	0.137
9	0.012	0.143	0.136	0.212	−0.007	0.146	0.080	0.251
10	0.238	0.146	0.244	0.207	0.491	0.396	0.375	0.417
11	0.285	0.448	0.396	0.275	0.437	0.454	0.487	0.525
12	0.505	0.479	0.476	0.478	0.855	0.799	0.745	0.541
13 years and more	1.218	0.817	0.920	0.679	1.380	1.113	1.227	0.834

Baseline score quartiles ranks cohort members by their achievement in mathematics, French and life skills from the lowest (1) to the highest achievement (4).

Table 3.5: Model 1 - Written French and math

	Written French		**Final grade**	**Written math**		**Final grade**
	OLS	**IV**		**OLS**	**IV**	
Final grade	0.190***	0.274***		0.184***	0.306***	
	(0.010)	(0.055)		(0.011)	(0.062)	
Gender	−0.011	0.000	−0.091	0.115**	0.129***	−0.067
	(0.039)	(0.044)	(0.191)	(0.046)	(0.047)	(0.191)
Height	0.006	−0.011	0.188*	−0.035	−0.057**	0.171*
	(0.025)	(0.026)	(0.098)	(0.023)	(0.026)	(0.095)
Mother's education	0.019**	−0.000	0.218***	0.022**	−0.005	0.222***
	(0.008)	(0.015)	(0.037)	(0.009)	(0.018)	(0.036)
Father's education	0.009	−0.008	0.183***	−0.002	−0.025	0.181***
	(0.007)	(0.014)	(0.030)	(0.008)	(0.016)	(0.030)
Wealth index	0.072**	0.029	0.470***	0.049	−0.016	0.499***
	(0.028)	(0.035)	(0.149)	(0.042)	(0.050)	(0.145)
Male-headed household	−0.130***	−0.139***	0.226	−0.090	−0.101	0.222
	(0.045)	(0.048)	(0.248)	(0.061)	(0.068)	(0.250)
Share of children in household	−0.045	0.128	−1.911***	−0.045	0.204	−1.878***
	(0.120)	(0.155)	(0.579)	(0.171)	(0.217)	(0.576)
Urban area	−0.271***	−0.191**	−0.886***	−0.424***	−0.304***	−0.907***
	(0.099)	(0.086)	(0.334)	(0.121)	(0.117)	(0.325)
Remoteness index	−0.038	0.001	−0.331***	−0.040	0.018	−0.344**
	(0.038)	(0.039)	(0.125)	(0.034)	(0.047)	(0.130)
School quality index	0.183**	0.136*	0.579**	0.125	0.057	0.571**
	(0.078)	(0.081)	(0.225)	(0.079)	(0.079)	(0.227)
Primary teacher's education	0.175	0.230*	−0.416	0.144	0.212*	−0.319
	(0.124)	(0.129)	(0.460)	(0.120)	(0.128)	(0.451)
Good harvests before 04	0.075*	0.078**	−0.026	0.052	0.054	−0.015
	(0.041)	(0.039)	(0.132)	(0.043)	(0.044)	(0.130)
Bad harvests before 04	−0.102***	−0.101***	−0.022	−0.082*	−0.084*	0.005
	(0.038)	(0.037)	(0.143)	(0.043)	(0.046)	(0.145)
Good harvests since 04	0.007	0.017	−0.090	0.018	0.030	−0.069
	(0.029)	(0.031)	(0.096)	(0.035)	(0.034)	(0.095)
Bad harvests since 04	−0.055**	−0.055**	0.009	−0.009	−0.008	0.003
	(0.026)	(0.025)	(0.113)	(0.024)	(0.023)	(0.109)
Good years of family business	0.034	0.038	−0.049	0.044	0.049	−0.046
	(0.030)	(0.030)	(0.171)	(0.031)	(0.032)	(0.168)
Bad years of family business	−0.008	−0.001	−0.079	−0.003	0.008	−0.099
	(0.036)	(0.033)	(0.116)	(0.030)	(0.029)	(0.120)
Livestock loss since 04	−0.023	−0.033	0.076	−0.071	−0.082	0.058
	(0.044)	(0.039)	(0.182)	(0.052)	(0.053)	(0.182)
Number of high schools			0.450***			0.454***
			(0.099)			(0.106)
Presence of primary school (years)			0.015**			0.015**
			(0.006)			(0.006)
Constant	−1.571***	−2.329***	8.326***	−1.422***	−2.508***	8.226***
	(0.165)	(0.477)	(0.886)	(0.193)	(0.580)	(0.872)
N	833	833	833	850	850	850
R^2	0.660	0.609	0.431	0.522	0.418	0.433
F	89.723	62.320	32.397	42.055	29.309	45.981
Underidentification test (p-value)		0.001			0.001	
Hansen J statistic (p-value)		0.744			0.509	

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Model 1 - Oral math and life skills

	Oral math		Final grade	Life skills		Final grade
	OLS	IV		OLS	IV	
Final grade	0.177***	0.270***		0.119***	0.114**	
	(0.010)	(0.074)		(0.010)	(0.048)	
Gender	0.141***	0.152***	−0.058	0.115*	0.114*	−0.089
	(0.043)	(0.047)	(0.193)	(0.060)	(0.059)	(0.182)
Height	0.068**	0.048*	0.197**	0.060**	0.061**	0.194**
	(0.026)	(0.029)	(0.092)	(0.024)	(0.026)	(0.091)
Mother's education	0.020**	−0.001	0.227***	0.016*	0.017	0.237***
	(0.009)	(0.019)	(0.035)	(0.009)	(0.016)	(0.035)
Father's education	−0.005	−0.022	0.177***	0.013*	0.014	0.176***
	(0.009)	(0.018)	(0.029)	(0.008)	(0.011)	(0.030)
Wealth index	0.045	−0.006	0.511***	−0.083**	−0.080**	0.487***
	(0.044)	(0.058)	(0.137)	(0.033)	(0.040)	(0.139)
Male-headed household	−0.085	−0.098*	0.261	−0.046	−0.046	0.227
	(0.053)	(0.057)	(0.252)	(0.057)	(0.056)	(0.252)
Share of children in household	0.119	0.303	−1.813***	−0.086	−0.097	−2.013***
	(0.147)	(0.209)	(0.567)	(0.138)	(0.163)	(0.593)
Urban area	−0.318**	−0.228*	−0.895***	−0.057	−0.062	−0.968***
	(0.125)	(0.135)	(0.325)	(0.094)	(0.089)	(0.330)
Remoteness index	−0.066**	−0.022	−0.346**	−0.059*	−0.061*	−0.399***
	(0.032)	(0.052)	(0.132)	(0.031)	(0.035)	(0.134)
School quality index	0.144*	0.090	0.589***	0.133*	0.136*	0.619***
	(0.083)	(0.078)	(0.203)	(0.067)	(0.071)	(0.221)
Primary teacher's education	0.193	0.249*	−0.379	−0.223**	−0.225**	−0.180
	(0.139)	(0.149)	(0.457)	(0.108)	(0.106)	(0.457)
Good harvests before 04	0.061	0.063	−0.025	0.097**	0.097**	0.070
	(0.041)	(0.041)	(0.129)	(0.048)	(0.047)	(0.137)
Bad harvests before 04	−0.059	−0.060	0.000	−0.073	−0.073	−0.073
	(0.040)	(0.041)	(0.144)	(0.050)	(0.050)	(0.147)
Good harvests since 04	0.008	0.018	−0.078	−0.009	−0.010	−0.076
	(0.030)	(0.031)	(0.096)	(0.030)	(0.030)	(0.095)
Bad harvests since 04	−0.011	−0.010	−0.005	−0.032	−0.032	−0.010
	(0.029)	(0.028)	(0.109)	(0.033)	(0.033)	(0.105)
Good years family business	−0.004	0.002	−0.062	0.088***	0.088***	−0.013
	(0.030)	(0.030)	(0.168)	(0.031)	(0.031)	(0.169)
Bad years family business	−0.010	−0.002	−0.095	−0.076	−0.076	−0.076
	(0.030)	(0.029)	(0.122)	(0.052)	(0.051)	(0.116)
Livestock loss since 04	−0.048	−0.058	0.081	0.045	0.045	0.069
	(0.051)	(0.049)	(0.181)	(0.037)	(0.037)	(0.176)
Number of high schools			0.442***			0.446***
			(0.106)			(0.111)
Presence of primary school (years)			0.014**			0.017***
			(0.006)			(0.006)
Constant	−1.382***	−2.207***	8.257***	−0.817***	−0.773*	8.090***
	(0.198)	(0.687)	(0.874)	(0.206)	(0.403)	(0.871)
N	844	844	844	888	888	888
R^2	0.524	0.465	0.443	0.361	0.361	0.457
F	55.303	25.722	42.427	20.653	15.971	43.621
Underidentification test (p-value)		0.001			0.000	
Hansen J statistic (p-value)		0.329			0.010	

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Model 2 - Written French and math

	Written French		**Baseline**	**Written math**		**Baseline**
	OLS	**IV**	**achievement**	**OLS**	**IV**	**achievement**
Baseline French score	0.234***	0.697*				
	(0.051)	(0.411)				
Baseline math score				0.217***	1.044**	
				(0.072)	(0.444)	
Gender	−0.010	0.040	−0.103	0.066	0.047	0.033
	(0.051)	(0.062)	(0.067)	(0.064)	(0.067)	(0.060)
Height	0.043	0.040	0.013	−0.002	−0.004	0.008
	(0.037)	(0.040)	(0.030)	(0.032)	(0.041)	(0.036)
Mother's education	0.044***	0.014	0.066***	0.056***	0.013	0.055***
	(0.011)	(0.029)	(0.011)	(0.015)	(0.035)	(0.013)
Father's education	0.037***	0.026*	0.024**	0.024**	0.011	0.015
	(0.010)	(0.015)	(0.010)	(0.011)	(0.013)	(0.009)
Wealth index	0.138***	0.068	0.138***	0.147***	0.111**	0.030
	(0.043)	(0.075)	(0.032)	(0.047)	(0.057)	(0.040)
Male-headed household	−0.101	−0.067	−0.055	−0.085	−0.071	0.007
	(0.071)	(0.095)	(0.069)	(0.073)	(0.082)	(0.078)
Share of children in household	−0.318	−0.151	−0.308*	−0.466**	−0.358	−0.063
	(0.194)	(0.274)	(0.172)	(0.222)	(0.279)	(0.187)
Urban area	−0.449***	−0.411***	−0.097	−0.561***	−0.606***	0.033
	(0.133)	(0.151)	(0.144)	(0.177)	(0.212)	(0.155)
Remoteness index	−0.116**	−0.082	−0.039	−0.110**	−0.024	−0.067
	(0.050)	(0.074)	(0.071)	(0.042)	(0.090)	(0.068)
School quality index	0.245**	0.120	0.284***	0.136	−0.118	0.311**
	(0.097)	(0.141)	(0.087)	(0.109)	(0.215)	(0.136)
Primary teacher's education	0.044	0.041	−0.027	0.043	0.166	−0.154
	(0.173)	(0.185)	(0.189)	(0.139)	(0.220)	(0.216)
Good harvests before 04	0.035	0.018	0.038	0.059	0.037	0.024
	(0.053)	(0.055)	(0.049)	(0.056)	(0.066)	(0.049)
Bad harvests before 04	−0.092*	−0.113*	0.046	−0.080	−0.093	0.021
	(0.052)	(0.060)	(0.049)	(0.056)	(0.068)	(0.055)
Good harvests since 04	−0.032	−0.058	0.061*	0.001	−0.051	0.068*
	(0.037)	(0.045)	(0.032)	(0.044)	(0.056)	(0.038)
Bad harvests since 04	−0.034	−0.011	−0.050	0.012	0.056	−0.055
	(0.037)	(0.043)	(0.042)	(0.037)	(0.061)	(0.052)
Good years family business	0.011	0.008	0.006	0.032	0.039	−0.009
	(0.056)	(0.055)	(0.038)	(0.060)	(0.081)	(0.049)
Bad years family business	−0.016	−0.014	−0.004	−0.029	−0.059	0.031
	(0.050)	(0.053)	(0.031)	(0.048)	(0.064)	(0.038)
Livestock loss since 04	0.004	0.011	−0.015	−0.061	−0.001	−0.077
	(0.067)	(0.075)	(0.074)	(0.073)	(0.104)	(0.064)
Number of high schools			0.098**			0.123**
			(0.042)			(0.051)
Presence of primary school (years)			−0.001			−0.000
			(0.002)			(0.002)
Constant	0.211	0.250	−0.218	0.280	0.208	−0.150
	(0.240)	(0.259)	(0.290)	(0.204)	(0.302)	(0.316)
N	776	776	776	750	750	750
R^2	0.429	0.281	0.332	0.316	−0.214	0.280
F	40.736	22.903	15.111	15.980	10.289	10.690
Underidentification test (p-value)		0.106			0.070	
Hansen J statistic (p-value)		0.002			0.102	

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: Model 2 - Oral math and life skills

	Oral math		Baseline	Life skills		Baseline
	OLS	**IV**	**achievement**	**OLS**	**IV**	**achievement**
Baseline math score	0.152***	0.164				
	(0.050)	(0.536)				
Baseline life skills score				0.222***	2.954	
				(0.038)	(4.726)	
Gender	0.144**	0.145**	−0.031	0.102	0.161	−0.022
	(0.055)	(0.057)	(0.068)	(0.068)	(0.182)	(0.066)
Height	0.092***	0.092**	0.028	0.089***	0.113	−0.006
	(0.033)	(0.036)	(0.038)	(0.029)	(0.125)	(0.035)
Mother's education	0.057***	0.056**	0.047***	0.032***	−0.103	0.050***
	(0.011)	(0.027)	(0.012)	(0.010)	(0.231)	(0.011)
Father's education	0.023**	0.023*	0.016	0.034***	0.007	0.010
	(0.010)	(0.013)	(0.010)	(0.009)	(0.041)	(0.010)
Wealth index	0.138***	0.137**	0.058	−0.033	−0.252	0.076
	(0.049)	(0.058)	(0.046)	(0.036)	(0.451)	(0.046)
Male-headed houseold	−0.067	−0.066	−0.107	−0.012	0.034	−0.015
	(0.071)	(0.091)	(0.066)	(0.061)	(0.273)	(0.088)
Share of children in household	−0.114	−0.112	−0.116	−0.274*	0.034	−0.095
	(0.174)	(0.196)	(0.173)	(0.141)	(0.663)	(0.172)
Urban area	−0.475***	−0.475***	−0.099	−0.161	0.047	−0.087
	(0.151)	(0.145)	(0.166)	(0.098)	(0.426)	(0.142)
Remoteness index	−0.143***	−0.141	−0.148**	−0.097***	0.261	−0.121*
	(0.040)	(0.088)	(0.058)	(0.031)	(0.696)	(0.069)
School quality index	0.197*	0.195	0.224*	0.157**	−0.388	0.206*
	(0.103)	(0.141)	(0.115)	(0.078)	(0.785)	(0.109)
Primary teacher's education	0.087	0.091	−0.360*	−0.229*	0.567	−0.306
	(0.160)	(0.226)	(0.212)	(0.117)	(1.793)	(0.222)
Good harvests before 04	0.052	0.051	0.020	0.092*	−0.146	0.088**
	(0.054)	(0.054)	(0.058)	(0.051)	(0.456)	(0.041)
Bad harvests before 04	−0.061	−0.061	0.042	−0.078	0.002	−0.028
	(0.054)	(0.055)	(0.066)	(0.051)	(0.203)	(0.043)
Good harvests since 04	−0.009	−0.010	0.036	−0.024	−0.157	0.048
	(0.035)	(0.038)	(0.032)	(0.031)	(0.258)	(0.036)
Bad harvests since 04	−0.008	−0.008	−0.022	−0.016	0.198	−0.078*
	(0.042)	(0.043)	(0.038)	(0.037)	(0.408)	(0.040)
Good years family business	−0.033	−0.034	0.063	0.090*	0.134	−0.016
	(0.046)	(0.056)	(0.043)	(0.047)	(0.159)	(0.044)
Bad years family business	−0.015	−0.014	−0.033	−0.091	−0.190	0.036
	(0.043)	(0.043)	(0.050)	(0.060)	(0.224)	(0.042)
Livestock loss since 04	−0.049	−0.049	−0.007	0.034	0.008	0.012
	(0.063)	(0.061)	(0.071)	(0.041)	(0.167)	(0.061)
Number of high schools			0.032			0.029
			(0.046)			(0.069)
Presence of primary school (years)			−0.003			−0.001
			(0.002)			(0.002)
Constant	0.153	0.147	0.587*	0.199	−0.412	0.208
	(0.179)	(0.284)	(0.343)	(0.190)	(1.673)	(0.427)
N	801	801	801	856	856	856
R^2	0.336	0.336	0.241	0.293	−5.972	0.237
F	18.592	14.945	7.324	16.424	1.224	9.556
Underidentification test (p-value)		0.391			0.807	
Hansen J statistic (p-value)		0.005			0.804	

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Model 3 - Written French

	Written French		**Grade**	**Lagged**
	OLS	**IV**	**progression**	**score**
Grade progression	0.197***	0.396***		
	(0.012)	(0.078)		
Baseline French score	0.150***	0.290*		
	(0.045)	(0.173)		
Gender	−0.062	−0.090	0.271**	−0.110*
	(0.047)	(0.058)	(0.119)	(0.066)
Height	0.030	0.016	0.062	0.008
	(0.030)	(0.029)	(0.072)	(0.030)
Mother's education	0.031***	0.003	0.084***	0.064***
	(0.008)	(0.014)	(0.029)	(0.011)
Father's education	0.024***	0.006	0.073**	0.026**
	(0.007)	(0.010)	(0.028)	(0.010)
Wealth index	0.103***	0.035	0.222**	0.086**
	(0.035)	(0.047)	(0.102)	(0.037)
male-headed household	−0.142**	−0.167**	0.199	−0.040
	(0.055)	(0.069)	(0.186)	(0.066)
Share of children in the household	−0.266*	−0.132	−0.331	−0.286
	(0.152)	(0.187)	(0.476)	(0.179)
Urban area	−0.359***	−0.251***	−0.394*	−0.173
	(0.103)	(0.080)	(0.204)	(0.140)
Remoteness index	−0.079*	−0.026	−0.114	−0.049
	(0.042)	(0.043)	(0.076)	(0.068)
School quality index	0.204**	0.102	0.314**	0.241**
	(0.086)	(0.087)	(0.151)	(0.108)
Primary teacher's education	0.145	0.245	−0.410	0.021
	(0.152)	(0.153)	(0.246)	(0.174)
Good harvests before 04	0.067	0.092*	−0.139	0.031
	(0.049)	(0.055)	(0.106)	(0.046)
Bad harvests before 04	−0.088**	−0.094*	−0.004	0.052
	(0.043)	(0.048)	(0.148)	(0.047)
Good harvests since 04	0.009	0.038	−0.161**	0.064*
	(0.031)	(0.035)	(0.081)	(0.032)
Bad harvests since 04	−0.071**	−0.097***	0.155*	−0.044
	(0.028)	(0.030)	(0.085)	(0.042)
Good years family business	0.027	0.043	−0.078	−0.006
	(0.036)	(0.036)	(0.159)	(0.038)
Bad years family business	−0.000	0.016	−0.076	−0.001
	(0.032)	(0.030)	(0.125)	(0.031)
Livestock loss since 04	0.008	0.016	−0.054	−0.013
	(0.049)	(0.049)	(0.146)	(0.073)
Number of high schools			0.254***	0.130***
			(0.071)	(0.039)
Presence of primary school (years)			0.011***	−0.001
			(0.003)	(0.002)
% of households with a deposit			−0.248	1.562**
			(1.111)	(0.665)
% of households with a loan			1.710	−1.319*
			(1.212)	(0.780)
Constant	−0.505**	−1.207***	3.371***	−0.290
	(0.191)	(0.337)	(0.618)	(0.276)
N	776	776	776	776
R^2	0.585	0.392	0.258	0.347
F	60.846	25.673	23.262	20.947
Underidentification test (p-value)		0.000		
Hansen J statistic (p-value)		0.495		

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.10: Model 3 - Written math

	Written math		**Grade**	**Lagged**
	OLS	**IV**	**progression**	**score**
Grade progression	0.189***	0.392***		
	(0.016)	(0.103)		
Baseline math score	0.144**	0.308		
	(0.058)	(0.276)		
Gender	0.033	−0.008	0.306**	0.038
	(0.059)	(0.063)	(0.116)	(0.063)
Height	−0.014	−0.028	0.028	0.006
	(0.027)	(0.027)	(0.069)	(0.036)
Mother's education	0.039***	0.008	0.089***	0.056***
	(0.012)	(0.019)	(0.026)	(0.012)
Father's education	0.010	−0.008	0.080***	0.016*
	(0.008)	(0.012)	(0.026)	(0.009)
Wealth index	0.108**	0.055	0.250***	0.019
	(0.048)	(0.052)	(0.094)	(0.039)
Male-headed household	−0.124*	−0.162**	0.230	0.008
	(0.067)	(0.078)	(0.177)	(0.080)
Share of children in the household	−0.298	−0.087	−0.604	−0.065
	(0.196)	(0.245)	(0.472)	(0.188)
Urban area	−0.485***	−0.417***	−0.429**	−0.005
	(0.147)	(0.123)	(0.205)	(0.162)
Remoteness index	−0.078**	−0.018	−0.120	−0.072
	(0.035)	(0.044)	(0.078)	(0.066)
School quality index	0.118	0.024	0.392***	0.359**
	(0.093)	(0.119)	(0.141)	(0.167)
Primary teacher's education	0.089	0.174	−0.238	−0.117
	(0.127)	(0.143)	(0.260)	(0.205)
Good harvests before 04	0.082	0.100*	−0.050	0.016
	(0.054)	(0.061)	(0.101)	(0.051)
Bad harvests before 04	−0.091*	−0.106*	−0.058	0.029
	(0.053)	(0.060)	(0.129)	(0.054)
Good harvests since 04	0.030	0.046	−0.091	0.071*
	(0.035)	(0.038)	(0.074)	(0.038)
Bad harvests since 04	−0.014	−0.029	0.104	−0.053
	(0.029)	(0.039)	(0.078)	(0.052)
Good years family business	0.037	0.044	−0.040	−0.016
	(0.037)	(0.039)	(0.149)	(0.048)
Bad years family business	−0.011	0.000	−0.078	0.030
	(0.029)	(0.030)	(0.124)	(0.038)
Livestock loss since 04	−0.054	−0.028	−0.071	−0.080
	(0.056)	(0.060)	(0.139)	(0.063)
Number of high schools			0.251***	0.140***
			(0.076)	(0.051)
Presence of primary school (years)			0.011***	0.000
			(0.003)	(0.002)
% of households with a deposit			−0.669	0.135
			(1.144)	(0.756)
% of households with a loan			1.050	−1.092
			(1.224)	(0.950)
Constant	−0.388*	−1.126***	3.245***	−0.122
	(0.195)	(0.429)	(0.604)	(0.316)
N	750	750	856	750
R^2	0.458	0.248	0.267	0.285
F	25.300	16.592	28.637	12.006
Underidentification test (p-value)		0.105		
Hansen J statistic (p-value)		0.824		

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Model 3 - Oral math

	Oral math		**Grade**	**Lagged**
	OLS	**IV**	**progression**	**score**
Grade progression	0.182***	0.396***		
	(0.014)	(0.126)		
Baseline math score	0.090**	0.067		
	(0.043)	(0.337)		
Gender	0.085*	0.017	0.365***	−0.043
	(0.049)	(0.071)	(0.128)	(0.067)
Height	0.086***	0.077**	0.041	0.026
	(0.029)	(0.031)	(0.073)	(0.037)
Mother's education	0.040***	0.019	0.096***	0.045***
	(0.010)	(0.024)	(0.029)	(0.011)
Father's education	0.010	−0.006	0.073***	0.016
	(0.009)	(0.017)	(0.027)	(0.010)
Wealth index	0.097**	0.046	0.251**	0.019
	(0.048)	(0.056)	(0.103)	(0.046)
Male-headed household	−0.106*	−0.145**	0.189	−0.102
	(0.058)	(0.063)	(0.185)	(0.066)
Share of children in the household	−0.052	0.028	−0.321	−0.098
	(0.149)	(0.193)	(0.472)	(0.175)
Urban area	−0.392***	−0.291**	−0.382*	−0.140
	(0.134)	(0.128)	(0.214)	(0.161)
Remoteness index	−0.107***	−0.057	−0.144*	−0.152***
	(0.033)	(0.073)	(0.080)	(0.057)
School quality index	0.151	0.085	0.360**	0.141
	(0.091)	(0.113)	(0.149)	(0.139)
Primary teacher's education	0.151	0.242	−0.387	−0.365*
	(0.148)	(0.218)	(0.265)	(0.192)
Good harvests before 04	0.072	0.094*	−0.092	0.023
	(0.048)	(0.056)	(0.104)	(0.059)
Bad harvests before 04	−0.053	−0.046	−0.032	0.041
	(0.045)	(0.049)	(0.144)	(0.067)
Good harvests since 04	0.020	0.052	−0.132*	0.038
	(0.028)	(0.035)	(0.079)	(0.032)
Bad harvests since 04	−0.038	−0.073*	0.155*	−0.021
	(0.034)	(0.039)	(0.082)	(0.038)
Good years family business	−0.009	0.016	−0.092	0.059
	(0.033)	(0.039)	(0.119)	(0.042)
Bad years family business	−0.005	0.007	−0.073	−0.031
	(0.032)	(0.037)	(0.111)	(0.050)
Livestock loss since 04	−0.046	−0.042	−0.040	−0.004
	(0.050)	(0.050)	(0.147)	(0.069)
Number of high schools			0.232***	0.044
			(0.077)	(0.049)
Presence of primary school (years)			0.009***	−0.003
			(0.003)	(0.002)
% of households with a deposit			−0.802	1.281
			(1.172)	(0.871)
% of households with a loan			1.629	0.024
			(1.249)	(0.735)
Constant	−0.446**	−1.178**	3.379***	0.489
	(0.170)	(0.532)	(0.618)	(0.350)
N	801	801	801	801
R^2	0.465	0.284	0.271	0.248
F	23.027	14.518	26.916	7.347
Underidentification test (p-value)		0.291		
Hansen J statistic (p-value)		0.363		

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.12: Model 3 - Life skills

	Life skills		**Grade**	**Lagged**
	OLS	**IV**	**progression**	**score**
Grade progression	0.108***	0.191		
	(0.013)	(0.126)		
Baseline life skills score	0.194***	0.666		
	(0.034)	(0.633)		
Gender	0.074	0.063	0.306**	−0.026
	(0.065)	(0.077)	(0.116)	(0.067)
Height	0.085***	0.086**	0.028	−0.007
	(0.028)	(0.038)	(0.069)	(0.035)
Mother's education	0.023**	−0.009	0.089***	0.049***
	(0.009)	(0.036)	(0.026)	(0.011)
Father's education	0.025***	0.014	0.080***	0.010
	(0.008)	(0.013)	(0.026)	(0.010)
Wealth index	−0.058*	−0.117	0.250***	0.054
	(0.033)	(0.086)	(0.094)	(0.047)
Male-headed household	−0.034	−0.043	0.230	−0.010
	(0.057)	(0.076)	(0.177)	(0.088)
Share of children in household	−0.206	−0.098	−0.604	−0.086
	(0.137)	(0.197)	(0.472)	(0.173)
urban area	−0.108	−0.029	−0.429**	−0.115
	(0.092)	(0.135)	(0.205)	(0.145)
Remoteness index	−0.077**	0.003	−0.120	−0.125*
	(0.030)	(0.112)	(0.078)	(0.069)
School quality index	0.125*	0.002	0.392***	0.179
	(0.072)	(0.140)	(0.141)	(0.119)
Primary teacher's education	−0.196*	−0.026	−0.238	−0.290
	(0.109)	(0.274)	(0.260)	(0.213)
Good harvests before 04	0.100*	0.063	−0.050	0.086**
	(0.051)	(0.077)	(0.101)	(0.041)
Bad harvests before 04	−0.072	−0.053	−0.058	−0.026
	(0.053)	(0.063)	(0.129)	(0.044)
Good harvests since 04	−0.011	−0.025	−0.091	0.049
	(0.030)	(0.045)	(0.074)	(0.035)
Bad harvests since 04	−0.031	−0.003	0.104	−0.075*
	(0.034)	(0.066)	(0.078)	(0.039)
Good years family business	0.094**	0.106**	−0.040	−0.021
	(0.036)	(0.045)	(0.149)	(0.044)
Bad years family business	−0.082	−0.093	−0.078	0.037
	(0.053)	(0.060)	(0.124)	(0.042)
Livestock loss since 04	0.039	0.038	−0.071	0.013
	(0.039)	(0.053)	(0.139)	(0.061)
Number of high schools			0.251***	0.041
			(0.076)	(0.067)
Presence of primary school (years)			0.011***	−0.001
			(0.003)	(0.002)
% of households with a deposit			−0.669	0.684
			(1.144)	(0.588)
% of households with a loan			1.050	−0.391
			(1.224)	(0.636)
Constant	−0.164	−0.556	3.245***	0.170
	(0.190)	(0.603)	(0.604)	(0.435)
N	856	856	856	856
R^2	0.343	0.108	0.267	0.239
F	21.268	13.825	28.637	8.878
Underidentification test (p-value)		0.546		
Hansen J statistic (p-value)		0.039		

All regressions include age dummies. Robust standard errors in parentheses are clustered by localities.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

General conclusion

This doctoral thesis aimed at answering three different questions. In the first chapter, I assessed to what extent malnutrition can be an obstacle to school achievement in the context of rural Madagascar. I used rainfall shocks experienced during early childhood as exogenous instruments to estimate the effect of the height for age z-score on mathematics and reading test scores. The results suggest that stunting has a direct detrimental effect on the acquisition of basic knowledge and that naive OLS estimates are likely to underestimate this relationship.

The second chapter focused on the effect of food distribution cycle on refugee households' food consumption and the nutritional status of their children. We used an exogenous variation in the time interval between the distribution and the interview to estimate the effect of time on cereal consumption. The more distant the food distribution, the less important is household's consumption of cereals. This pattern has direct consequences on child's nutritional status: each day away from the distribution reduces the average mid-upper arm circumference for age of children under five years old.

Finally, the third chapter aimed at measuring the contribution of schooling in the acquisition of cognitive skills among a cohort of young adults. We take advantage of a panel database to estimate a value-added model by including a measure of prior achievement in the production function. We combine this approach with an instrumental variable strategy by using the availability of school infrastructures and characteristics of the credit market to instrument grade attainment and baseline achievement. Our results demonstrate that schooling is by far the main factor in the acquisition of cognitive skills.

The issues discussed in the three chapters of this thesis are varied but summarizing the findings can give a larger picture of the obstacles to the accumulation of human capital in developing countries. First of all, consistent with the literature, results from the first two chapters confirm that early childhood is a crucial period in the formation of human capital. In rural Madagascar and in refugee camps from southern Chad, children nutritional status is very sensitive to the variation of their parents' resources. Moreover, stunting resulting from early childhood nutritional deficiencies has a detrimental effect on children's learning. Therefore, we can expect programs orientated toward the protection of the nutritional status of children under five years old to have large multiplier effects.

Secondly, households' welfare is particularly sensitive to unexpected variation in their resources. The most vulnerable of them show limited capacities to cope with income shocks and to smooth their consumption when times are hard. Considering the large and permanent effect of shocks on the accumulation of human capital, it is essential to protect the most vulnerable. Poor households would greatly benefit from safety net programs and the development of the local credit market.

Lastly, the level of schooling is worrying in Madagascar. We have reported suggestive evidence that the weakest children leave school before completing primary school and that knowledge acquisition is very limited over the first grades. As grade attainment appears to be the main determinant of skills acquisition and as we observe acceleration in learning, it is essential to keep children at school (the availability of secondary schools have an incentive effect on schooling) and make schooling more efficient by improving primary school quality as well as teachers' qualification.

Abstract

This thesis contributes to the literatures on food consumption, child malnutrition and school achievement in developing countries. The first chapter aims at estimating the causal relationship between malnutrition and school achievement among Malagasy children enrolled in primary school. Rainfall shocks during the first years of a child's life are used as exogenous instruments to predict the long-term nutritional status. Results indicate that stunting is a significant obstacle to learning. The second chapter investigates the effect of food distribution cycle on refugee households' consumption. Results suggest that the more distant the food distribution, the less important is household's consumption of cereals. The effect is sufficiently large to impact children short-term nutritional status. The third chapter takes advantage of a panel database of a cohort of Malagasy young adults to estimate a cognitive skills production function with a value-added approach. Results highlight the large role of schooling in the acquisition of cognitive skills.

Keywords: Malnutrition, school achievement, value-added model, refugee camps, food aid, food consumption.

Résumé

Cette thèse contribue à la littérature sur la consommation alimentaire, la malnutrition infantile et la réussite scolaire dans les pays en développement. Le premier chapitre a pour but d'estimer la relation causale entre la malnutrition et la réussite scolaire parmi un échantillon d'enfants malgaches inscrits dans le cycle primaire. Les chocs de pluie endurés lors de la petite enfance sont utilisés comme instruments exogènes pour expliquer le statut nutritionnel de long terme. Les résultats indiquent que les retards de croissance sont un obstacle important pour l'apprentissage scolaire. Le deuxième chapitre examine l'effet des cycles de distribution alimentaire sur la consommation alimentaire de ménages réfugiés. Les résultats suggèrent que la consommation moyenne de céréales diminue au fur et à mesure que l'on s'éloigne du jour de la distribution alimentaire. Cet effet est suffisamment important pour avoir un impact sur le statut nutritionnel de court terme des enfants de notre échantillon. Le troisième chapitre exploite des données de panel sur une cohorte de jeunes adultes malgaches afin d'estimer une fonction de production de capacités cognitives grace à l'utilisation d'un modèle de valeur ajoutée. Les résultats soulignent le rôle essentiel de la scolarisation dans l'acquisition de capacités cognitives.

Mots clefs: Malnutrition, réussite scolaire, modèle de valeur ajoutée, camps de réfugiés, aide alimentaire, consommation alimentaire.