Journal *of* African Development (2016) 18:31–60

Home vs. School Deworming and Meal Programs: Evidence from Rural Senegal*

Théophile T. Azomahou† • Fatoumata L. Diallo

Abstract

This study uses primary data to assess the impact of deworming and school feeding programs on pupil's educational outcomes measured by test scores. In particular we elaborate on deworming at home vs. school. Relying on Propensity Score Matching and selection models we find that meals program significantly increase pupils' test scores. The impact is more important for pupils who are in schools having association of mothers. Deworming at school has a positive effect on pupils' performance while deworming at home has a negative impact. This result indicates that the use of widely spread traditional deworming medicines should be discouraged. Among the determinants of test scores, class size displays a U-shaped relation pointing to the crucial role of group effect in shaping pupils' knowledge. The impact of programs does vary depending on the gender of pupils. Policy analysis including cost-effectiveness is also provided.

Keywords: Selection models, school feeding, deworming, school performance.

JEL Classification: C31, C52, I15, I28

*We gratefully acknowledge insightful comments from Wladimir Raymond and Abdoulaye Diagne. This work was supported by the FERDI (Fondation pour les Etudes et Recherches sur le Développement International) and the Agence Nationale de la Recherche of the French government through the program "Investissements d'avenir ANR-10-LABX-14-01."

Théophile T. Azomahou
United Nations University (UNU-MERIT), Maastricht University, University of Auvergne and CERDI
azomahou@merit.unu.edu

Fatoumata L. Diallo
University Cheikh-Anta-Diop of Dakar (UCAD) and Consortium pour la Recherche Economique et Sociale (CRES)
faatdiallo@yahoo.fr

1. Introduction

Several studies have examined the impact of school feeding and deworming programs on education outcomes in developing countries. However, scholarly contributions on Africa are spotty, despite the fact that program evaluation is an important tool to inform policy makers about the efficient allocation of resources and improvement of existing policies.[1] Adelman et al. (2008) provided a review of assessment of Food for Education (FFE) programs on education outcomes. The authors showed conflicting evidence on the ability of school meals to improve pupils' cognitive development test performance in mathematics and language. Some studies found that school meal programs lead to a significant increase in academic performance as measured by test scores, while other studies found a negative effect or no impact. Vermeerch and Kremer (2004) used data from a randomized breakfast program in kindergartens in Kenya. The authors showed that the program improved student learning, but only for those whose teachers were more experienced at the beginning of the program. Cueto and Chinen (2008) examined the impact of an experimental program of school breakfasts in primary schools in Peru. The results showed that children in the treatment group, who are in 'multigrade' schools and 'simple flow' performed better in coding tests, arithmetic and reading.

The impact of deworming on school performance is even rarer. Kvalsvig et al. (1991) examined the impact of deworming in South Africa and concluded that the drug treatments had some effects on school performance, though not significant. Miguel and Kremer (2004) studied the impact of an experimental deworming program, among others, on students scores in Kenyan primary schools. The authors showed that there is no evidence that deworming increases the scores of students. The study also indicates that not taking into account deworming externalities may underestimate the effects of the deworming program.

To the best of our knowledge, there is no study that sought to measure both the impact of a lunch and a deworming program, while estimating the determinants of school performance in a joint framework.[2] Our study uses a unique and recent data set from a school lunch and deworming program in rural Senegal to assess their effects on pupils' achievement measured by test scores while elaborating on the determinants of test scores. To this end, we use observational data on 162 schools from an experimental program called PIERI.[3] This program was led by the 'Consortium for Social and Economic Research' (CRES) and the Senegalese Ministry of Education. The sample contains data on school and teacher characteristics, pupils' achievements (score in French, Maths, aggregate score), pupils' and household characteristics, community characteristics, shock variables, geographical location and treatment indicators.

In this study we use both Propensity Score Matching (PSM) and Selection Models (structural specification) approaches for two reasons. Firstly, for the purpose of comparison, we wanted to have the most robust results. Indeed, the PSM allows us to estimate the Average Treatment effect on the Treated (ATET) while the structural specification informed about the average treatment effect on the entire sample. Secondly, we are not only interested in the determinants of pupils' performance but also in the joint estimation of the determinant of pupils' performance and the selection process (or the probability to benefit from the programs). In that context, the structural model was suited better to our purpose. Indeed, with such a framework we can specify a two system equation: one equation for the outcome (performance equation) and another for the treatment (selection equation). It is clear that the

[1] Policy makers are seen here as governments, aid donors and the development community at large.

2 Throughout the paper, school performance means the school performance of pupils and is measured by scores.

[3] PIERI (Policy Impact Evaluation Research Initiative).

program evaluation approach (in particular the propensity score matching) cannot do that as it is only designed for the purpose of evaluation.

In addition, this study distinguishes between deworming at home and deworming in school. We have one indicator of meal program (33.73% of pupils receive meals program) and three indicators of deworming: the school or home deworming (25.88% of pupils were dewormed in school or at home) which we call deworming program throughout the text, the school deworming (11.80% of pupils) which we call deworming in school and the house deworming (16,08% of pupils) which we call deworming at home. It is important to emphasize that the deworming program is not the sum of the deworming in school and at home. Furthermore, we do not take the intersection of deworming in school and at home into account.

Relying selection models, the main results that emerge from this study are the following: i) The results show that the meal program can significantly increase pupils performance. Also the impact is more significant for pupils who are in school with association of mothers. ii) The deworming in school has a positive effect on pupils performance while the deworming at home has a negative impact on pupils' scores. Our reading is that it is essential to assist rural households to deworm children at home. This will discorage the use of traditional medicines (that have adverse consequences on pupils's performance) for deworming children. In addition, the deworming program increases significantly the scores of pupils who are in school with association of mothers or school project. iii) Among the determinants of test scores (schools and teacher, pupils', household and community characteristics and shock variables), school manual, class size, age of teacher and pupil's age show a non-linear relation with the scores. Particularly, class size displays a U-shaped relation, meaning that performance (aggregate and math scores) decreases with increasing class size which is consistent with the findings in the literature on performance. However, starting from a certain threshold, the performance increases with respect to class size. In the context of rural Senegal, large classes may impact positively test scores emphasizing the crucial role of group effect in shaping pupils' knowledge. Among other controls, education, expenditure, household size per equivalent adult, water point, koranic school, early childhood institution, existence of a high school in the village have a positive effect on pupil's performance, while controls such as disturbed courses, gender of teacher, absenteeism of teacher and parents school have a negative impact on scores. In terms of policy analysis, the results show that the impact of programs varies depending on the gender of the pupil, the existence of association around the school or the educational level of the pupil.

The rest of the paper is structured as follows. In section 2 we describe the data used. In section 3 we present the econometric specification. Section 4 summarizes estimation results. Section 5 presents a policy analysis before the conclusion in section 6. Detailed results of this paper are supplemented with online appendix.

2. Description of the data

School feeding or Food For Education (FFE) is the activity through which the World Food Program (WFP) supported the education sector in Senegal since the 1960s. This intervention aims at providing pupils with a regular diet and to promote children's access to basic quality education, especially girls. This intervention supports the government in achieving universal education for all children by the year 2015, which is one of the goals of the Ten-Year Education and Training Program (PDEF) and Millennium Development Goals (MDGs). WFP activities are implemented within the framework of two programs namely: the Country Program, which extends from 2007 to 2011 and Programme of Protracted Relief and Recovery Operation from 2008 to 2010. To be assisted by WFP, the school must meet the following conditions:

i) be located in an area of food insecurity or particularly affected by higher prices, be in a rural community or peri-urban area,
ii) preferably have a minimum size of 50 pupils and maximum of 600 pupils for easier management concerns,
iii) have an operational management committee comprising women, commitment of the local community to develop infrastructure such as storage, cooking facilities etc.,
iv) have an acceptable standard of hygiene.

The Ministry of Education, through the Directorate of General Administration and Equipment (DAGE) also funds a program for the setting up of school canteens. Its division of the medical school (DCMS), also leads to another program which focuses on deworming and medical monitoring in rural public schools.

Initially, WFP's school feeding program covered the regions of Fatick, Kaolack, Tambacounda and Matam (under the Country Program 2007-2011) and the regions of Ziguinchor, Kolda and Sedhiou (under the IPSR successively carried out since 2003 in Casamance Natural). Under the action plan, WFP strengthened the school feeding program in its initial intervention areas: Fatick, Kaolack, Kaffrine, Tambacounda, Kedougou, Matam, Ziguinchor, Kolda and Sedhiou considered as priority areas. Later on, the action plan was extended to other regions: Diourbel, Louga and Thies. The program involves a total of 12 regions out of 14 in Senegal.

In this study, we use primary data collected by the 'CRES' and the Ministry of Education in four rural areas in Senegal namely: Diourbel, Fatick, Kolda and Sedhiou. The sampling strategy involved first selecting a number of areas for experimentation based on the following criteria: i) poor areas, ii) existence of a standardized assessment, iii) location of intervention must be priority areas for partners (such as WFP) that support the experiment.[4] Then, schools eligible for the program were randomly selected based on the following criteria:

i) lack of a school canteen, ii) lack of access to a package of health services (deworming and iron supplementation), iii) have an enrollment of at least 50 pupils and less than 600, iv) have at least second and fourth year classes of primary school, v) have a functioning management committee, vi) be located in rural areas. Finally, a random selection of 20 students in each class was conducted in each selected school. Because of inaccurate information on schools meeting the selection criteria, it was found out that some schools in the sample survey have had a deworming and/or canteens program for years.

We thus have an observational database of 162 schools out of which we have formed three groups: a group of pupils that benefit from a meal program (meaning have a meal at school's canteen or bring a meal to school), a group of pupils that are dewormed (at school or at home), and the control group that had nothing.

2.1. Variables

The data available are on the schools and teachers' characteristics, pupils and households' characteristics, communities where pupils live characteristics and geographical location. The variables that will be used in the estimation are: i)the outcome or school performance variables (aggregate score, French and math scores), ii) the determinants of performance or control variables that can broadly be grouped as follows: those related to schools and teacher

[4] The Senegalese education system implements, for nearly a decade, a major innovation in assessing the performance of elementary students. This approach called 'standardized assessment' is a logical progressions based on a harmonized process for formative assessment.

characteristics, pupils characteristics (gender, age, number of hours, grade, koranic school, early childhood institution, snack and pupil eat to their satisfaction), household characteristics (Food and education expenditure, household size, livestock, gender of household head, literacy of household head, marital status and land), community characteristics (”high school” and ”parents' school”) and shock variables (disturbed courses, absenteeism, number of pregnancies, number of marriages, sick in last three months preceding the survey and health expenditure of household). Additionally, we also have the treatment dummies and region dummies. The definition of variables are documented in the online appendix. The control variables in the equation of performance were chosen based on both their availability in the database and their relevance in terms of the literature on the determinants of academic performance. Some control variables are specific to the context of the study and are therefore of particular interest. We will revisit them during the description of variables.

School characteristics. Generally, it is assumed that schools with better infrastructure would record the best performance. School and teacher characteristics taken into account are: the number of classes in temporary shelters ”Temporary shelters”, the number of students per manual ”School manual”, the class size, the distance between the school and the student's home ”Distance to school”, existence of a water point in school ”Water point”, existence of a playground ”Playground”, the possibility for students to eat near their school ”Meal near school”, existence of a first aid in the school, gender of teacher, teacher's age, continuing training received by the teacher, professional and academic certificate of teacher. Among these variables, those that contextualize our study are temporary shelters and the opportunity for students to eat near their school. For variable temporary shelters, we expect a negative effect on scores. In Senegal, to overcome the lack of classrooms in some areas, temporary shelters have been set up as classrooms. It is usually makeshift straw constructions which become unusable during the rainy season. For ”Meal near school,” we expect a positive effect on student performance. The variable school manual and its square were also introduced to take into account a possible non-linearity that may exist between the variable and school performance. Indeed, it is generally accepted that the best students benefiting from textbooks in schools perform better. The variable class size and its square were introduced as control variables to take into account the non-linearity that may exist between performance and class size. Indeed, research and experiments on the effects of a reduction of class sizes on student performance have not led to consistent conclusions. While some studies show that students in small classes have better academic results (Piketty, 2004), others show that the reduction of classrooms still leaves uncertainty (Davies, 2003). Glass and Smith (1979) showed that the effect of reducing class sizes is not linear. We expect a positive relationship between academic performance and students who live less than 1 km from the school. So we expect a positive effect of the variables ”Water point” and ”playground” on student performance.

For teacher characteristics, in the Senegalese context, we expect that a female teacher will have a positive influence on students particularly girls who identify themselves more easily with a female teacher. Regarding the age of the teacher, we assume a positive effect. We introduced the square of the teacher's age to take into account of possible non-linearity that may exist between this variable and scores. Indeed, if we assume that the conduct of classroom instruction is positively affected by the maturity of the teachers, we can expect a positive correlation between teacher's age and school performance, Schwille et al., (1991). It is important to remember that in the literature, there is no consensus on the relationship between the age of teacher and school performance. Indeed, the teacher's age can interact with both experience and education level of the teacher. This means that some teachers having the same age but with different levels of education and different experience do not necessarily have the same effects on student achievement.

Concerning academic qualification of teachers, we made a distinction between those who have the National certificate (Brevet) and those who have the High school diploma and higher (bacplus). Regarding the impact of academic certificate of the teacher on student performance, studies so far have mixed results (Clotfelter et al., 2006). For continuing training of the teacher, we expect a positive effect on scores because training improves teaching skills of teachers.

Also, in the literature, there is no consensus on the effect of teacher's professional certificate on scores. In fact, according to CONFEMEN (1999), the teachers trained for a year in Burkina Faso and Cameroon led to less satisfactory results in 5th year than those who had no professional training. In Senegal, the same phenomenon is recorded for teachers who received two years of training compared to those who received only one year. It seems that the initial training is the basis of this result (CONFEMEN, 1999).

Regarding the variable continuing training, we expect a positive effect on pupil's performance because continuing training should improve teaching methods of teachers. However, it is important to note that the impact of continuing training on the performance of pupils yields a positive effect if the training covers several periods, otherwise you can even record a negative effect (UNESCO, 2000).

Pupils' characteristics. Pupils' characteristics taken into account are gender, age, number of hours that the child spends on domestic work per day, attending the koranic school, attending an early childhood institution, bring a snack to school and eat to his/her fill.

For the gender of pupils some studies indicate that boys are better than girls in mathematics and science (Felouzis, 1997). Other studies show that gender differences are not statistically significant and sometimes mixed (Ma, 2007). We anticipate that boys will perform better than girls because girls in rural areas are more likely to be engaged in domestic work. In addition, some families always doubt the value of the education of girls.

We also introduced the pupil's age and its square to highlight the non-linearity that may exist between the age of the pupil and their performance. The literature mentions a positive effect of pupils's age on scores (Schwille et al., 1991) or a negative effect, depending on whether pupils enroll late in school or has had a lot of repetitions.

Regarding the variable "number of hours," we expect a negative effect on school performance. In fact, children who are forced to spend a high number of hours on domestic activities are exposed to poor school performance because it reduces their learning time.

Attending an early childhood institution and or a koranic school are specific to the context of our study. The koranic school, usually found in countries where islam is predominant, is an informal private educational structure that provides religious education based on memorization. In Senegal, 95.9% of the population are muslim. We now understand that this structure, albeit informal, is commonplace in the Senegalese society. This school is known for developing the capacity of learning and memory of students because children learn to memorize at an early age. It is expected that these two variables will positively affect student performance.

Two other control variables are the possibility of pupils eating to their fill and being able to take a meal near to school. These two variables are not frequently used in the literature. Recall that the study focuses on a rural population and these indicators tell us about the nutritional well-being of students. We anticipate that these two variables will have a positive effect on pupils performance.

Households' characteristics. Among the households' characteristics of students we have included: food and education expenditures, household size per adult equivalent, literacy of the household head, gender of household head, marital status, possession of a farmland and livestock owned by the household.

The variable food expenditure is an indicator of standard of living of the household. We expect a positive correlation to performance. On variable education expenditure, we also expect

a positive relationship with performance because the more parents invest in the education of their children, the more they will supervise the children's learning at home.

Household size per adult equivalent has also been used as a control variable. Studies on the impact of family size on schooling in developing countries have led to highly controversial results that do not allow for generalising conclusions. We expect a positive relationship between family size and pupils' performance. Indeed, we believe that in large size households, it is likely to have older children who are already in high school and who should be willing to guide the youngest. Literacy of household head is another important variable in the explanation of the performance of the student. In fact, a literate environment has a positive influence on the child's ability to learn to read, write and do arithmetic. Provided, however, the child is well supported at home.

The marital status of parents is also counted among the household characteristics. Since more than 95% of household heads are married, we have grouped the terms of this variable to married and unmarried (unmarried included singles, divorced, concubine and widowed). We expect that children living with married parents will have better performance, although individual experiences suggest that children living with a single widowed or divorced mother are generally more successful compared to children living with both parents in a large polygamous household. Two other important variables not found in the literature are the household ownership of arable land and the number of livestock owned by the household. Despite the fact that these two variables are indicators of wealth, we expect that they will have a negative effect on children's performance. Indeed, in terms of arable land, especially in the context of Senegal, rural households still continue to use their children for domestic and farm work. It is therefore likely that children living in households with farmland are asked to work in the fields, which will result in reducing the learning time of pupils. The same reasoning applies to the variable livestock owned by the household. Indeed, the larger size of livestock, the higher the probability that children in the household would be exposed to work.

Community characteristic. Studies examining the effects of community characteristics on pupils' performance are still limited, compared to those that examine the effects of individual, family and school variables. However, factors associated with neighborhood can have both positive and negative influences on school performance. In this research, the characteristics of the environment taken into account are: the existence of a high school in the village of the pupils, the fact of children living in communities where some of them do not go to school because parents are not interested in school.

Regarding the existence of a high school, we expect a positive effect on performance. Indeed, parents who do not have the means to help their children to pursue studies in a remote village after obtaining the primary certificate will invest little in their child's education, because they are convinced that he/she will leave school prematurely. In addition, living in a village where there is a high school implies that there are seniors who are in high school and can support the younger ones in their studies. Regarding the variable "Parents school", we expect a negative effect on performance. Indeed, living in a community where some children do not go to school because parents are not interested could negatively impact on the scores. In this kind of community, a lot of children generally assist their families. We believe that living in this context is likely to adversely affect the performance of children who go to school because there is a big chance of a ripple effect.

Shock variables. Among the shocks, we have those directly affecting the child, those affecting his school, his household and those acting on the environment in which he lives. Shock variables directly affecting pupils are: being sick during the last three months preceding the survey ("sick in last 3 month"). The health of students is a major element for explaining of the performance. Being sick for the last three months preceding the survey can reduce the time

of school attendance and learning ability hence the performance of pupil. We expect a negative effect of this variable on student performance.

Shock variables affecting school of pupils are: the number of pregnancies "Pregnancies" and early marriages registered during the school year, teacher absenteeism and the existence of disturbances causing delays in starting courses.

For the variable number of pregnancies and number of marriages, we expect a negative effect on performance. Indeed, girls in rural areas are generally exposed to early marriage and pregnancy. If the marriage or pregnancy does not result automatically in a complete and simple dropout of pupils, it becomes extremely difficult for them to focus on their studies due to family responsibilities.

Regarding the variable absenteeism, we expect a negative effect of this variable on performance, because if the teacher is often absent, the learning time will reduce, as well as performance.

The existence of disturbances having led to delays in the start of the course also reduces the learning time of pupils. We hypothesized a negative effect of this variable on pupils performance. A shock variable affecting the household of the pupil is health expenditure. We expect a negative effect of this variable on performance of the pupil. In fact, a high health expenditure shows that there have been cases of illness in the household. Indeed, the illness of the pupil has a negative effect on the learning time. In addition, the illness of a family member (parent, etc.) can psychologically affect the pupil. It is important to note that low health expenditure does not necessarily signify that there are no cases of illness in the household concerned. It may be that the household is unable to provide medical care for their members.

In addition, to take into account some specific characteristic area of intervention for which information is not available, we introduced the indicator regions among the control variables. Those are dummies variables for Diourbel Fatick, Kolda and Sedhiou. The reference region is Fatick which has the largest number of pupils. Moreover, for the choice of control variables in the selection equation of the meal program at school, we used the criteria set by WFP and the division of school canteens. Among the criteria there is information which is not available in the database such as being located in an area of food insecurity or particularly affected by rising prices of foods. The variables are the total number of students in the school, the distance between the school and pupils' homes, the existence of a management committee in the school, an association of parents, a school cooperative, a grant from the Rural Council, a water point, disturbances that delay the start of courses, a storage, and gender of the pupil. We assume that the long distance to school negatively influences the probability of benefiting from the program, while the forms of associations and infrastructure that exist within the school will act positively on the probability of having the program. Similarly, we anticipate disturbances to the course and having more boys in the school to act negatively on the probability of benefiting from the meal program. For the choice of control variables in the selection equation of deworming program, we drew on the criteria listed by the Ministry in the meal program and our intuition. The variables include: the total number of students, the existence of a management committee for the school, a parents association, a school cooperative, a water point, a first aid, gender of the pupil and literacy of the household head.

2.2. Descriptive statistics

We summarize the statistics of variables in the appendix, respectively for meal and deworming programs. The statistics are provided by group (treatment and control group). Additionally, the

last column of each block reports the test whether variables are statistically different across treatment and control groups.

For the meal program, on average, the treatment group performed better than the control group in terms of the scores (aggregate, French and math). Indeed, the mean aggregate, French and math, scores for the treatment group are 42.445, 41.877 and 42.890 respectively, whereas they are 37.587, 38.228 and 36.949 for the control group. There is a significant difference between the two groups as shown by the t-test. We also note that there is no significant difference between the two groups for the following variables: total pupils, number of hours, food expenditure, household size, school cooperative, meal near school, professional qualification (CAP), other professional qualification, academic qualification (High school diploma), academic qualification (national certificate), gender of pupil, grade, koranic school, sick in last 3 months, pupils eat to their fill, literacy of household head, marital status, land and parents school. On average, there are about 187 pupils per school in each group. Pupils spend, on average, 3 hours per day to do household work, also, the household size is about 11 persons on average and we note the same level of food expenditure in each group. In each group, about 56% of teachers have a ”Brevet” as academic qualification, while only 44% have a ”High school diploma” as academic qualification. In the treatment group, about 48% of the pupils are boys, while in the control group, there are about 50% of boys. But there is no significant difference between the two averages.

In the treatment group, approximately 27% of pupils had become sick during the last three months preceding the survey while it is 24.4% in the control group. But there is no significant difference between the two averages. Also, on average, in each group, 16% of household heads are literate, 95% are married and 95% have their own arable land. In each group, 49% of children live in a community where some children do not go to school because their parents are not interested in sending them to school.

For some variables such as temporary shelters, school manual, pregnancies, teacher’s age, education expenditure, health expenditure, distance to school, playground, management committee, association of parents, association of mothers, rural council grant, water point, disturbed courses, storage, professional qualification CEAP, gender of household head, Fatick and Sedhiou, the treatment group shows the highest average with a significant difference. However, if we take the variables class size, marriages, pupil’s age, livestock, school project, first aid, gender of teacher, no professional diploma, continuing training, absenteeism, early childhood institution, snack, high school, Diourbel and Kolda, the control group displays the highest average with a significant difference.

Indeed, around 92% of schools in the treatment group have a playground while it is around 75% for schools in the control group. Also about 90% of schools in the treatment group delayed in starting the courses because of disturbances. This figure is about 74.5% for schools in the control group. Also, one may note that 20% of teachers in the treatment group have no professional diploma compared to about 32% in the control group. We noted that 35% of teachers in the treatment group had a CEAP as professional qualification while there were 19% in the control group. For continuing training, about 65% of teachers had benefited from continuing training in the control group compared to 45% in the treatment group. It is important to note that 8% of the teachers in the control group are often absent while it was only around 4% in the treatment group. Also about 4% of pupils in the treatment group have attended an early childhood institution. This figure is about 7% for pupils in the control group. Also about 5% of pupils in the treatment group bring a snack to school. This figure is about 10% for pupils in the control group. About 24% of schools in the control group are located in a village having a high school while, the figure is only about 13% for schools in the treatment group. The majority of treated schools are located in Fatick while the majority of schools in

the control group are in Kolda. In sum, the treatment group performs better in terms of scores.

Regarding the deworming program, the summary statistics shows that on average, for score variables, now there are no significant differences between the treatment and control groups as documented by the t-test. In fact, in the treatment group one notes 37.495, 37.680 and 37.161 respectively for aggregate, French and Math scores. In the control group it is 37.587, 38.228 and 36.949 respectively for aggregate, French and Math scores. Also, for a number of control variables such as health, class size, marriages, food expenditure, household size, meal near school, first aid, gender of teacher, professional qualification CEAP, other professional qualification, gender of pupil, grade, koranic school, gender of household head, marital status, land and Sedhiou, there is no significant difference between the two groups. In fact, in each group, the average class size is about 36% of pupils, the household size is about 11 persons and about 12% of schools have a first aid. Also, around 49% of pupils are boys in the treatment group and 51% in the control group, 27.6% of pupils attended a koranic school in the treatment group and 29.6% in the control group. We observe that around 95% of households have married parents. In each group also, about 94% of household have an arable land. However, for a set of controls (temporary shelters, total pupils, pregnancies, teacher's age, education expenditure, distance to school, playground, management committee, school cooperative, association of parents, association of mothers, rural council grant, water point, storage, professional qualification CAP, High school diploma, absenteeism, early childhood institution, snack, sick in last 3 months, pupils eat to their fill, literacy of household head, parents school and Fatick), the control group displays the lowest average with a significant difference. While the treatment group displays the lowest average with a significant difference for all remaining variables, for example, about 85.6% of schools are in less than 1 km from pupil's home for the treatment group, only 78% of schools in the control group report the same. Also, about 52.6% of schools in the treatment group have a water point while only about 39.5% of schools in the control group have.

3. Estimation strategy

Observational studies are usually subject to non random samples. This is typically the case in non experimental data. In these situations where the sample is either fully or partly based on values taken by the outcome response variable, parameter estimates are likely to be inconsistent unless corrective measures are applied. Such samples are broadly defined as selected samples. Two types of specifications suited this purpose: i) the dummy endogenous selection model originally developed by Heckman (1979) and which is fully integrated into the evaluation literature, ii) the Roy model which dated from Roy (1951). In what follows, we apply these two specifications to our study.

Given the very influential debate by Deaton (2010), Heckman (2010) and Imbens (2010), it is clear that there is no unanimity on the superiority of various approaches: randomization or quasi-randomization vs. structural models. However, this strong debate sheds light on the methodological choices with regard to data availability, collection, and the aim of the study. As previously outlined, we are interested not only in assessing the effectiveness of meal and deworming programs, but also in the study of determinants of outcomes. In this regard, we found the structural model most appropriate for our purpose. Indeed, with such a framework, we can specify a two system equation: one equation for the outcome (performance equation) and another for the treatment (selection equation). The system can be estimated using the Full Information Maximum Likelihood (FIML) and the Control Function (CF) approach. Moreover, as the structural specification imbeds a treatment (selection) equation which is the

basic material of the PSM, we only present the structural model. The generic model combines a Probit selection equation (treatment equation) which is assumed to stem from an unobserved latent variable, and an intensity equation (performance equation).

3.1. The Heckman selection framework

This specification relies on an exclusion restriction, meaning that a variable determines participation in the programme but not the outcome of the program itself. Contrarily to 'matching' which is viewed as 'selection on observable', the Heckman approach enables 'selection on unobservables'. As outlined by Blundell and Costa Dias (2008), a comparison of both approaches turns to be very informative in understanding the advantages and limitations of these methods. Firstly, we assume a latent Probit selection mechanism (selection equation):

$$\begin{aligned} T_i^* &= \mathrm{w}_i'\alpha + u_i \text{ with } i = 1,\ldots,N \\ T_i &= 1_{[T_i^* > 0]} \end{aligned} \quad (1)$$

where w_i stands for the selection controls, and $T_i = 1$ if individual i is treated, u_i is the disturbance $u_i \sim \mathrm{N}[0,\sigma]$. Secondly, we assume a linear regression with sample selection (performance equation):

$$y_i = \mathrm{x}_i'\beta + \delta T_i + \varepsilon_i \quad (2)$$

We assume that equations (1) and (2) are linked by a bivariate normal distribution:

$$[u_i, \varepsilon_i] \sim \mathrm{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma & \rho \\ \rho & 1 \end{pmatrix}\right] \quad (3)$$

We use the Full Information Maximum Likelihood (FIML) and the Control Function (CF) which relies on Heckman (1976, 1979) two-step procedure to estimate the Average Treatment Effect (ATE) of the programs.

3.2. The Roy model

In the previous framework, selection is considered only on unobservables. Moreover, we are only in the position to identify a homogenous treatment effect namely the average treatment effect. This assumption is rather strong as it means that the effect of the program does not vary across populations. The Roy model has the advantage of considering selection on both observables and unobservables while relaxing the assumption of homogeneity of treatment by introducing heterogenous treatment effects. Therefore, we are able to identify and estimate not only the average treatment effect, but also treatment effect on the treated and on the untreated. The model with endogenous treatment and different outcome equations is

$$T_i^* = \gamma'\mathrm{w}_i + u_i, \quad T_i = 1_{[T_i^* > 0]}, \quad i = 1,\ldots,N \quad (4)$$

$$y_{i1} = \beta_1'\mathrm{x}_i + \varepsilon_{i1}, \quad (5)$$

$$y_{i0} = \beta_0'\mathrm{x}_i + \varepsilon_{i0}, \quad (6)$$

where T_i^* in equation (4) is a latent variable, the observed counterpart of which is T_i, and 1[ ] denotes the indicator function which takes on the value 1 if the corresponding latent variable is positive, and 0 otherwise. Relations (5) and (6) are respectively the outcome equations for participants and non-participants, γ and β's denote vectors of parameters to be estimated, and

μ_i, ε_1 and ε_0 denote error terms in the corresponding equations. Our estimation strategies are the maximum likelihood estimation and the two-step control function approaches.

3.2a. Maximum Likelihood (ML)

We assume joint normality for the three disturbances:

$$\begin{pmatrix} \mu_{i0} \\ \varepsilon_{i1} \\ \varepsilon_{i0} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & & \\ \rho\mu_{\varepsilon_1}\sigma_{\varepsilon_1} & \sigma^2_{\varepsilon_1} & \\ \rho\mu_{\varepsilon_0}\sigma_{\varepsilon_0} & 0 & \sigma^2_{\varepsilon_0} \end{pmatrix} \right]$$

We observe here that the zero correlation term corresponds to $\text{Cov}(\varepsilon_0, \varepsilon_1) = \rho_{\varepsilon_0\varepsilon_1}\,\sigma_{\varepsilon_0}\,\sigma_{\varepsilon_1}$ as the outcome equations are assumed not to be related. The likelihood function is then given by

$$\begin{aligned} L &= \prod_{i=1}^{N} \left[\int_{-\infty}^{-\gamma' \mathrm{w}_i} f_2(\mu_i, y_{i0})\, \mathrm{d}\mu_i \right]^{1-T_i} \left[\int_{-\gamma' \mathrm{w}_i}^{\infty} f_2(\mu_i, y_{i1})\, d\mu_i \right]^{T_i} \\ &= \prod_{i=1}^{N} \left[f_1(y_{i0}) \int_{-\infty}^{-\gamma' \mathrm{w}_i} f_1(\mu_i | y_{i0})\, \mathrm{d}\mu_i \right]^{1-T_i} \left[f_1(y_{i1}) \int_{-\gamma' \mathrm{w}_i}^{\infty} f_1(\mu_i | y_{i1})\, d\mu_i \right]^{T_i} \end{aligned} \tag{7}$$

where the functions $f_2(\cdot)$ and $f_1(\cdot)$ stand for bivariate and univariate normal density functions respectively. By replacing these functions by the analogous populations, the likelihood function becomes

$$L = \prod_{i=1}^{N} \left[\frac{1}{\sigma_{\varepsilon_0}} \phi(\zeta_0)\, \Phi\left(\frac{-\gamma' \mathrm{w}_i - \rho_{\mu_i\varepsilon_{i0}}\zeta_0}{\sqrt{1-\rho^2_{\mu_i\varepsilon_{i0}}}} \right) \right]^{1-T_i} \left[\frac{1}{\sigma_{\varepsilon_1}} \phi(\zeta_1)\, \Phi\left(\frac{-\gamma' \mathrm{w}_i - \rho_{\mu_i\varepsilon_{i1}}\zeta_1}{\sqrt{1-\rho^2_{\mu_i\varepsilon_{i1}}}} \right) \right]^{T_i} \tag{8}$$

where $\zeta_k = \frac{y_{ik} - \beta_k' x_i}{\sigma_{\varepsilon_k}}$, with $k = (0, 1)$, and $\phi(\cdot)$ and $\Phi(\cdot)$ denote respectively the standard normal probability density and the normal cumulative distribution functions. The parameter vector estimates are obtained by maximizing the log-likelihood ln L using standard numerical procedures (e.g. Newton-Raphson). Standard error estimates are obtained using the inverse Hessian or outer product gradient.

3.2b. Two-step

In order to estimate the Roy model using the Heckman's two-step method applied to the truncated means, we write the regression function for each subpopulation as

$$E(y_{i1} | \mathrm{T} = 1, \mathrm{x}, \mathrm{w}) = \beta_1' \mathrm{x}_i + \text{Cov}(\varepsilon_{i1}, \mu_i)\lambda_1(\gamma' w_i) \tag{9}$$

$$E(y_{i0} | \mathrm{T} = 0, \mathrm{x}, \mathrm{w}) = \beta_0' \mathrm{x}_i + \text{Cov}(\varepsilon_{i0}, \mu_i)\lambda_0(\gamma' w_i), \tag{10}$$

where $\lambda_1(\gamma' \mathrm{w}_i) = \frac{\phi(\gamma' \mathrm{w}_i)}{\Phi(\gamma' \mathrm{w}_i)}$ and $\lambda_0(\gamma' \mathrm{w}_i) = -\frac{\phi(\gamma' \mathrm{w}_i)}{1 - \Phi(\gamma' \mathrm{w}_i)}$ are the inverse Mills ratios. In terms of parameters to be estimated, these regressions can be rewritten as:

$$y_{i1} = \beta_1' \mathrm{x}_i + \rho_{\mu\varepsilon_1}\sigma_{\varepsilon_1}\lambda_1(\gamma' w_i) + \eta_{i1} \tag{11}$$

$$y_{i0} = \beta_0' \mathrm{x}_i + \rho_{\mu\varepsilon_0}\sigma_{\varepsilon_0}\lambda_0(\gamma' w_i) + \eta_{i0}, \tag{12}$$

where $E(\eta_{i1} | \mathrm{x}_i, \lambda_1) = E(\eta_{i0} | \mathrm{x}_i, \lambda_0) = 0$. In relations (11) and (12), $\lambda_1(\gamma' \mathrm{w}_i)$ and $\lambda_0(\gamma' \mathrm{w}_i)$ do enter as additional controls, the parameters of which $\rho_{\mu\varepsilon_1}\sigma_{\varepsilon_1}$ and $\rho_{\mu\varepsilon_0}\sigma_{\varepsilon_0}$ have to be estimated in addition to parameters vector β_1 and β_0. The estimation proceeds as follows:

1. Obtain consistent and efficient (under normality) estimates for γ by estimating a probit using maximum likelihood. Compute $\lambda_1(\gamma' w_i)$ and $\lambda_0(\gamma' w_i)$ given the predictions.
2. Use $\lambda_1(\gamma' w_i)$ and $\lambda_0(\gamma' w_i)$ as additional controls along side x_i and apply OLS to equations (11) and (12). Since we use the estimates of the λ's, the conventional standard errors are not valid and need to be corrected by using techniques of simulation or bootstrap.

3.2c. Treatment effects

The computations of the treatment effects are summarized in Table 1, where the covariance term is given by $\text{Cov}(\varepsilon_{i1} - \varepsilon_{i0}, \mu_i) = (\rho_{\mu i \varepsilon i1}\sigma_{\varepsilon i1} - \rho_{\mu i \varepsilon i0}\sigma_{\varepsilon i0})$. For the LATE(x), we apply the formulæ for double truncation to obtain the term $\lambda_{01} = \frac{\phi(\gamma' w_i) - \phi(\gamma' \tilde{w}_i)}{\Phi(\gamma' \tilde{w}_i) - \Phi(\gamma' \tilde{w}_i)}$ where $\tilde{w}$ stands for a 'policy instrument' (Heckman 2007). For these effects, the associated parameters ATE, ATET, ATENT and LATE can be retrieved by simply integrating out or averaging over the sample.

Table 1: Treatment parameters (single treatment Roy model)

Parameter	Definition	Assumptions (model)
ATE(x)[a]	$E(y_{i1} - y_{i0}\|x)$	$x_i(\beta_1' - \beta_0') := \vartheta(x)$
ATET(x, w)[b]	$E(y_{i1} - y_{i0}\|T = 1, x)$	$\vartheta(x) + \text{Cov}(\varepsilon_{i1} - \varepsilon_{i0}, \mu_i)\lambda_1$
ATENT(x, w)[c]	$E(y_{i1} - y_i 0\|T = 0, x)$	$\vartheta(x) - \text{Cov}(\varepsilon_{i1} - \varepsilon_{i0}, \mu_i)\lambda_0$
LATE(x, $\tilde{w}$, w)[d]	$E(y_{i1} - y_{i0}\|T(w) = 0, T(\tilde{w}) = 1, x)$ $E(y_{i1} - y_{i0}\| -\gamma' \tilde{w}_i \leq \mu_i < \gamma' w_i, x)$	$\vartheta(x) + \text{Cov}(\varepsilon_{i1} - \varepsilon_{i0}, \mu_i)\lambda_{01}$

(*a*) Average treatment effect
(*b*) Average treatment effect on the treated
(*c*) Average treatment effect on the untreated
(*d*) Local average treatment effect

4. Results

As indicated above, we used both PSM and the structural model. The latter allows to estimate jointly the performance equation (2) and the selection equation (1) while estimating the effect of the programs. Both equations are related via the correlation term ρ between errors. If the correlation is zero, the two equations are unrelated and the problem reduced to one estimable by OLS and the effect of the program is δ. Moreover, the comparison of PSM and the structural model results should be taken with caution. Indeed, some of the controls in the selection equation of the structural model are not included in the PSM specification. This is due to the restrictive nature of the PSM as it requires, among others, that i) only variables that simultaneously influence the participation decision and the outcome variable should be included, ii) only variables that are unaffected by participation (or the anticipation of it) should be included in the specification of the model.[5] In our framework, the two specifications do not exactly inform about the same result in term of causal effect. If the PSM allows to have the

[5] Moreover, the unsettled nature of the PSM as usually pointed out in the literature also prevents from using all the desired controls.

ATET, the structural approach can be used to produce various effects (ATE, ATET, ATENT and LATE).

4.1 Determinants of school achievement: A summary of findings

For the determinants of performance, we conducted several sets of estimates. For example, taking the meal program, we have made estimates with DESM (FIML and two step) and estimates with the Roy model (FIML and two step). With the deworming program, we also did estimations with DESM (FIML and two step) and the Roy model (FIML and two step). For reasons of space, we present only estimates on the meal program (FIML) and deworming program (FIML). Detailed tables of estimates are provided in the appendix.

The LR test on ρ supports our modeling strategy that both equations should be estimated jointly. The sign of the coefficients in the selection equation are the same for the three scores. The same applies in terms of significance apart from the variables 'total pupils' which is not significant for the aggregate score and 'distance to school' which is not significant for the math score. For example, total pupils, association of mothers, rural council grant, disturbed courses, storage, distance to school, impact positively the probability for a pupil to benefit from a meal program. Similar results are found when using the two step estimation. These results are of interest as we earlier mentioned in the descriptive statistics, these controls are specific indicators used by the World Food Program (WFP) and the Senegalese Ministry of Education. However, in the same line, one should expect that having an operational management committee and a water point will be the determinants of selection, which is not surprisingly the case. When we compare these findings to those obtained from the PSM estimation. We have quite similar results except for variable water point which is positive and significant now.

Regarding the determinants of performance as measured by test scores, as mentioned in the description of variables, we distinguished four categories of controls: those related to schools, pupils household characteristics, community and shock variables. Most of the control variables are of expected sign when significant. Some others are not intuitive. However, put in the specific context of rural Africa's experience, they become rather highly interesting. This is the case for class size as we will see later.

For controls related to school characteristics, as expected, increasing temporary shelters has a negative effect on pupils' performance (aggregate, French and math scores). However, the result is not significant for French score. Indeed, temporary shelters are not good for quality of learning because pupils are almost outside, and sometimes there are no suitable furniture, or blackboard that meet the requirement talk less of a roof.

For variables number of pupils per manual "School manual", class size and teacher's age as mentioned above, we included the linear and the square terms to account for possible nonlinearities. In fact, for school manual which vary between 0.2 to 7 pupils per manual, figures 1 and 2 show inverted U-shaped kind relations with scores, meaning that the number of pupils per manual increase the scores until a point from which the variable has a negative effect on performance. So turning point is 5 pupils for the aggregate score and 4 pupils for the French score. It is important to note that the nonlinear part of school manual is not significant for the math score although we observe an inverted U-shaped kind relation. Our reading is that even if school manual has a capacity to ameliorate the competence of pupils, when the number of pupils per manual becomes higher this impact negatively on scores because the learning time decreases with the corresponding increases of the number of pupils per manual. CONFEMEN (1999) shows that French and mathematics manuals have a positive effect on learning with a larger impact for the French manuals.

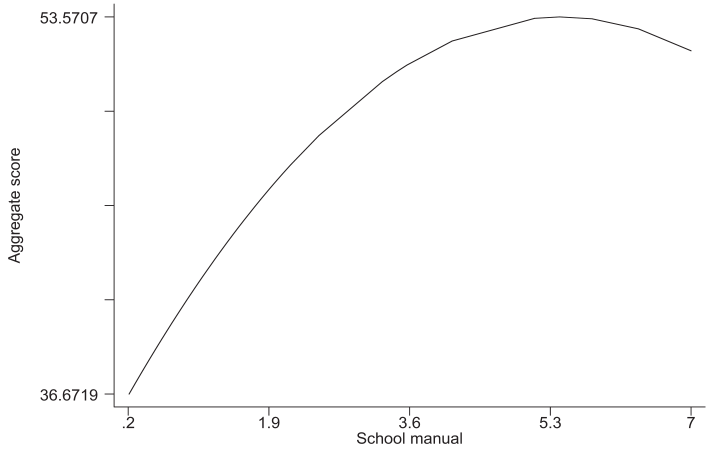

Figure 1: Meal: Relation between aggregate score and school manual.

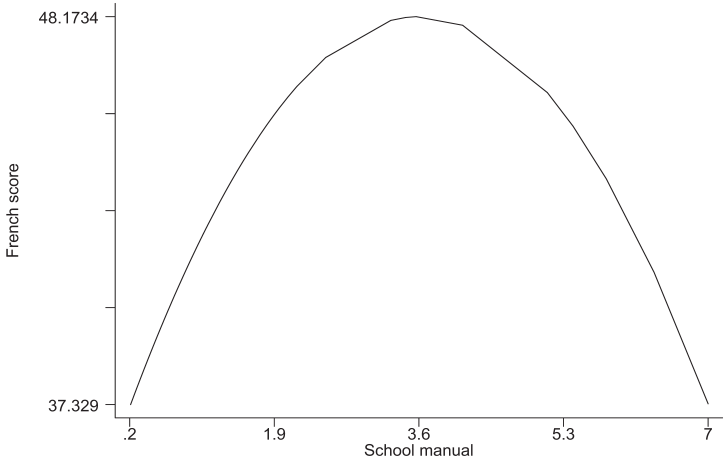

Figure 2: Meal: Relation between French score and school manual.

Class size displays a U-shaped relation with aggregate and math scores. Note that this variable ranges from 4 to 82 pupils. Firstly, the performance decreases with increasing class size which is consistent with some findings in the literature on performance. However, starting from 58 and 47 pupils respectively, for aggregate and math scores, the performance increases which is also consistent with other findings (see figures 3 and 4). It is important to note that the class size and its square are not significant for the French score. As mentioned above, the question as to whether a reduction of class size lead to improved performance remains inconclusive in the literature (Davies, 2003). How to explain that large classes can perform? In these specific rural areas, large classes may impact positively test scores thereby emphasizing the crucial role of group effect in shaping pupils' knowledge.

The teacher's age also shows a similar U-shaped picture. This variable ranges from 20 to 53. The linear term is negative while the square term is positive for scores. This means that the teacher's age affects negatively the scores up to a certain threshold at which additional age

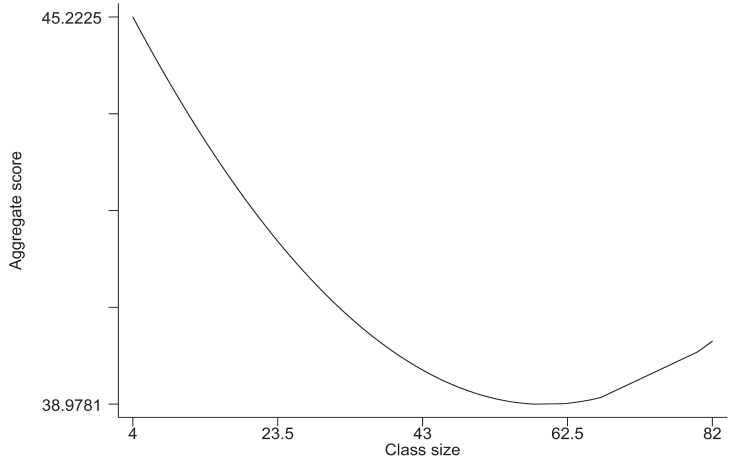

Figure 3: Meal: Relation between aggregate score and class size.

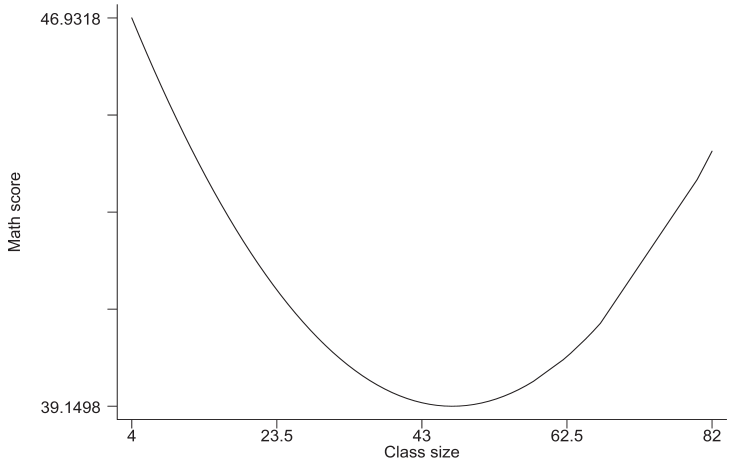

Figure 4: Meal: Relation between math score and class size.

positively affects performance. We observe that the turning point is 36, 37 and 33 respectively for aggregate, French and math scores (See figures 5 6 and 7).

Still on controls related to school characteristics, one can note that variable gender of teacher (male=1) and teacher absenteeism, as expected have a negative effect on scores. While the variables water point, professional qualification and academic qualification (High School Diploma) have a positive effect on pupils' performance. In fact, for the variable: gender of teacher, the result shows that pupils with female teachers perform better than those with male teachers. This result is consistent with the findings of Jarousse and Mingat (1989), who found similar results relying on data from Togo. For the variable absenteeism, the results are not surprising because when teachers are often absent, it reduces the learning time of pupils and their performance. For the variable water point, the result is intuitive because schools with better infrastructure will record the best performance.

Regarding professional qualification, the results show that pupils supervised by a teacher with the professional qualification CAP and CEAP perform better than those taught by a

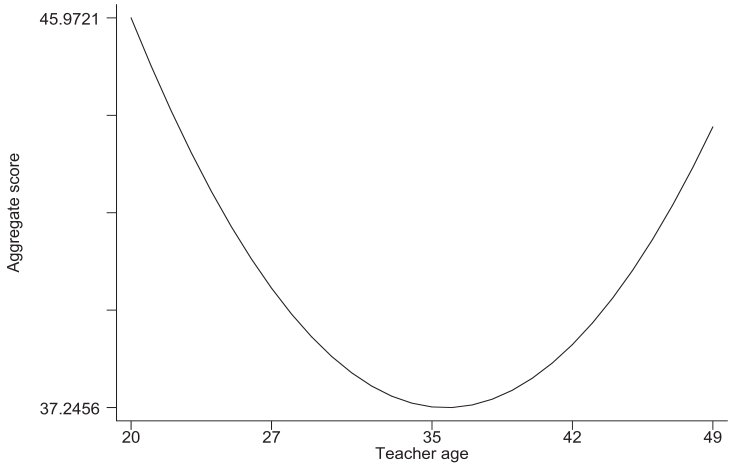

Figure 5: Meal: Relation between aggregate score and teacher age.

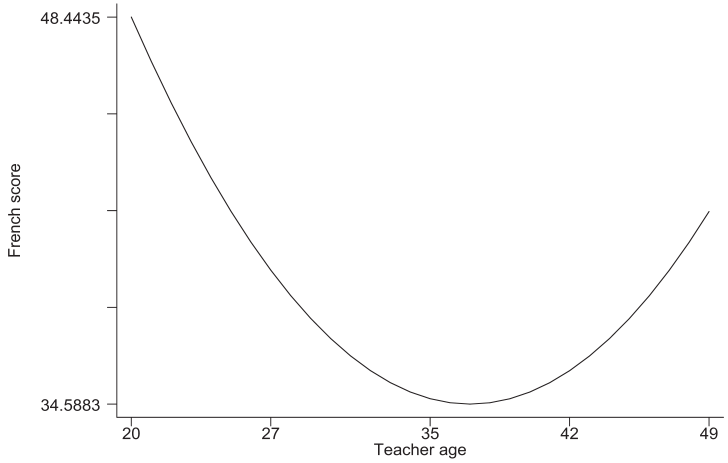

Figure 6: Meal: Relation between French score and teacher age.

teacher with another vocational degree which is the reference. This result is intuitive, in fact, in Senegal, the most two important professional qualifications are the CAP (Certificat d'Aptitude Pédagogique, meaning 'Pedagogical Aptitude Certificate') and CEAP (Certificat Elémentaire d'Aptitude Pédagogique, which is 'Basic Pedagogical Aptitude Certificate'). The surprising result is that pupils taught by a teacher with no professional diploma performs better than those supervised by a teacher with another professional degree. This result is consistent with a finding of CONFEMEN (2007) in Senegal. In rural areas, the result can be explained in two ways. Firstly, it may be that those teachers without professional diploma are awaiting to be officially graduated because most often they have already passed the written examination. Secondly, their level of motivation is much higher than teachers with professional qualification. In fact, teachers without professional qualification can be assessed at any time and this assessment is crucial for their career. So the performance of pupils under their care is reflected in the status of their job. As a result, it is likely that the motivation of those teachers is a push factor which

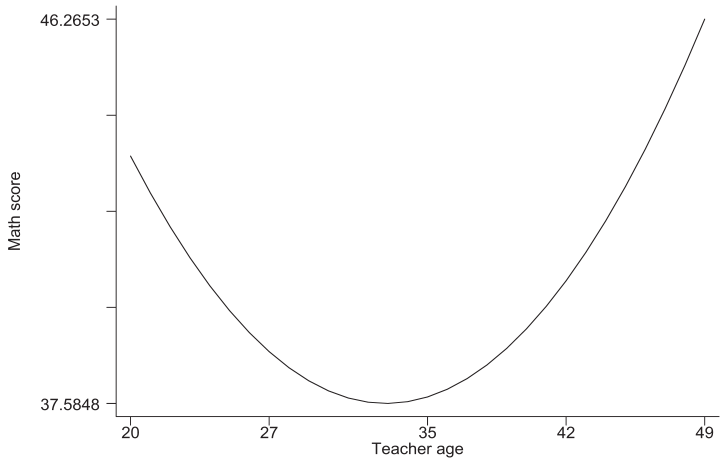

Figure 7: Meal: Relation between math score and teacher age.

leads them to teach pupils very well. This may also explain why pupils supervised by teachers without professional qualification perform better.

For the academic qualification of teachers, the results from the DESM model show that pupils supervised by a teacher with the High School Diploma perform better in mathematics compared to those supervised by a teacher with the national certificate. As pointed out by Rivers and Sanders (2002), this result suggests that teachers with a High School Diploma and more have a higher level of knowledge in mathematics than those with national certificate.

When one takes the deworming estimations, we always have the same result as in the meal regressions. If one takes the Roy model, one has the same results for untreated. However if one takes the treated, one observes a negative relation between the scores and the proportion of pupils who are supervised by a teachers with High School Diploma. This result is consistent with the findings of CONFEMEN (1999). The results also show that pupils supervised by teachers who have benefited from continuing training are less successful than those taught by teachers who have not benefited from such training, but the result is not statistically significant.

Regarding the variables related to pupils characteristics, pupil's age shows an inverted Ushaped relation. Meaning that, broadly, getting old has a negative effect on performance. In the sample, the variable varies between 6 to 15 years old. In fact, the results show that the age of the pupil has a positive effect on scores up to a certain threshold at which the age has a negative impact on scores. The thresholds are 13, 12 and 13 years old respectively for aggregate, French and math scores (see figures 8, 9 and 10). This finding is consistent with observations and can be related to several factors increasing the age of pupils during their primary school such as temporary dropout, repetition, adverse family conditions, etc. This result is observed for all estimations methods.

Other control variables related still on pupils' characteristic which includes gender of pupils, koranic school, early childhood institution, pupils eat to their fill positively impact performance. Indeed, for the gender of pupils, as expected, boys perform better than girls in scores and the results are significant for math and aggregate scores. These results are consistent with the findings of Felouzis (1997). Therefore, if we take the Roy's estimations, one observes that girls perform better in French scores in the group of pupils who benefit from the meal or the deworming program. Indeed, the presence of meal programs or deworming

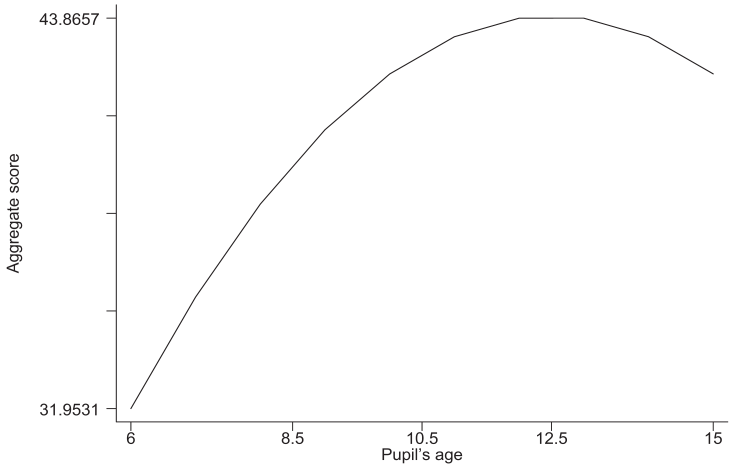

Figure 8: Meal: Relation between aggregate score and pupil age.

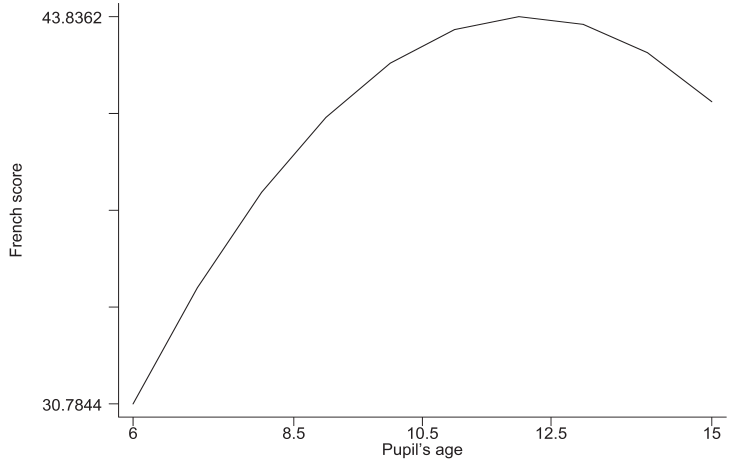

Figure 9: Meal: Relation between French score and pupil age.

improves the nutritional status of pupils. In addition, the presence of canteen allows girls to stay in school all day long. So they don't have to do domestic chores before going to school in the afternoon. Therefore they can concentrate better in the afternoon in the presence of a school meal.

For control variables "koranic school" which is an informal school and "early childhood institution" which is a modern pre-school, the results show a positive impact on scores. This result is intuitive and well known in the literature. Indeed, both types of schools are known to develop children's ability to learn at an early age. Particularly, koranic school is known for developing the ability of learning and memory because children learn by heart very early. For control "pupils eat fill", as anticipated, one observes a positive effect on scores. In fact the study focused on a population in rural areas and the fact to eat fill at home is an indicators telling us about the nutritional well-being of pupils in their household. So one can understand that this pupil performs better as compared to pupil that do not eat fill at home.

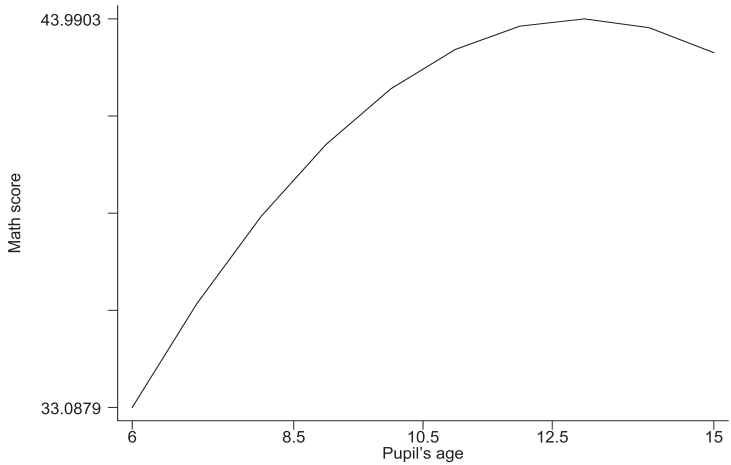

Figure 10: Meal: Relation between math score and pupil age.

Regarding household characteristics, the results show that education expenditure and household size, have a positive effect on performances. Indeed, for education expenditure, we observe that the more parents invest in the education of their children the more they will supervise their kids at home and more one could expect a better performance of pupils. For household size per adult equivalent as expected, one observes a positive effect on scores. As mentioned on the description of variables, we believe that large household size is likely to have senior students who are already in high school and are willing to support the younger ones. It is important to remember that studies on the effect of family size on education outcomes in developing countries have led to highly controversial results.

Another important variable is the literacy of household head. Contrary to our expectation, results show a negative impact of this variable on aggregate and math scores. In fact, as mentioned in the description of the variables, the fact of having a literate parent is not enough to get good results. To have good results pupils need monitoring at home. Therefore, one can think that children with literate parents are not well monitored by their parents at home.

For variable land, as expected, the result shows a negative impact on scores. The result is only significant for math scores. In fact, in Senegal, rural households still continue to use their children for domestic and farm work. So children living in households with a farmland are likely going to be asked to work in the fields, which will result in reducing the learning time of the pupils, hence their performances.

Regarding community characteristics, the control ”parents school” shows a negative effect on scores while high school shows a positive one. These two results are expected. For the control ”parents school”, living in a community where some children do not go to school because parents are not interested is likely to adversely affect the performance of children who go to school because there is a high chance of a ripple effect.

For the control ”high school”, the result can be viewed as an indirect effect. As mentioned in the description of variables, the result can be explained from two angles. Firstly, living in a community where there is a high school implies that parents are encouraged to invest in their children’s education because they know that children easily have the possibility to continue studying in the village after obtaining the primary certificate. Secondly, living in a village

that has a high school implies that there are more advanced seniors (brother or sister) who are in high school and who can guide the younger ones in their studies. However, we have no information on the fact that a pupil would have a brother or sister in a high school.

Regarding shock variables, the controls number of pregnancies registered in school, absenteeism of teacher, disturbed courses and health expenditure, impact scores negatively. These results are intuitive, for example the presence of absenteeism and disturbed courses will reduce the learning time of pupils as well as their performances. For regional dummies, the results show that pupils living in Diourbel perform better as compared to those living in Fatick which is the reference.

Finally the coefficient of the control meal program is positive and significant for French outcomes, meaning that having a meal program affects pupils' performance positively. Also, coefficient of the deworming program is negative for aggregate and math scores, but positive for French score. However it is worthwhile taking into consideration that these estimations are not causality results but rather correlations coefficients. We now turn to the causal relation as given by the treatment effects from the PSM and the structural models.

4.2. Treatments effects

As outlined earlier, our framework allows to identify and estimate the average treatment effects on the population (ATE); the average treatment effect on the treated (ATET); and the average treatment effect on the non-treated (ATENT) depending on the model used. In the Heckman selection framework, we can identify only the ATE. In the PSM we can identify the ATET and in the Roy model in addition to identifying the ATE and ATET, we can also calculate the ATENT. Regarding the ATENT, i.e, the average effect on the group of pupils who do not benefit from the meal program: the purpose is to inform policy makers on the rationale for extending the programs to this subpopulation of pupils. The combination of the average treatment effect on the treated (ATET) and the average treatment effect on the non-treated (ATENT) provides a more comprehensive view of the effectiveness of a given program and allows to choose a result from a wide range of options.

4.2a. School feeding program

From the foregoing, we computed the ATE of the meal program for the Heckman selection framework and for the Roy model, we used relations in Table 1 to compute ATE, ATET and ATENT. See Table 2 for results on various effect computed from Heckman selection framework and Roy model. For the PSM specification, the average treatment effect on the treated (Table 3) is based on various matching methods or algorithms in order to check the robustness of the estimates.[6] Table 2 shows a positive impact of meal programs on scores. We observe that the ATE on math score is higher than the ATE on aggregate and French scores.

From the Roy model, the ATE computed goes in the same direction as the ATE obtained from the DESM. However, it is important to note that the ATE from the Roy model are higher than those from de DESM. As mentioned above, in addition to the ATE, the Roy model allows us to compute treatment effect on treated and on non-treated. The results on average treatment effect on treated (ATET), show that the meal program has a positive impact on the pupils who benefit from the program. One can note that the ATE is higher than the ATET. This can be explained by the fact that the ATENT is higher than the ATET. For the ATENT, results show a

[6] In order to check the balancing between groups (treated vs. control), we also report the matching tests in the appendix.

Table 2: Treatment effect: Meal

	DESM[a]		Roy model (FIML[b])			Roy model (2-step[c])		
Outcome	ATE (FIMLa)	ATE (2-stepc)	ATE	ATET	ATENT	ATE	ATET	ATENT
Aggregate score	10.627***	10.478***	16.509***	7.596***	13.754***	22.233***	22.233***	22.026***
French score	7.086***	7.966***	13.881***	11.665***	13.197***	20.817***	21.774***	21.280***
Math score	12.993***	12.972***	21.658***	6.221***	16.858**	24.223***	22.085***	23.188***

Notes [a]: Dummy Endogenous Selection Model; [b]: Full Information Maximum Likelihood; [c]: Two step. ATE:Average treatment effect;
ATET:Average Treatment Effect on Treated; ATENT:Average Treatment Effect on Untreated
Significance levels (Bootstrap -100 replications-): * : 10% ** : 5% *** : 1%

positive effect of the meal program on all scores. The ATENT on math score is higher than the ATENT on aggregate and French score.

Results from the PSM specification (Table 3), show that ATET is positive and significant for aggregate and French scores using only the kernel matching and a ATET positive and significant using the K-NN, the K-NN(5) and the kernel procedure for the math score.

4.2b. Deworming programs: school vs. home

For the deworming program, as mentioned earlier, we distinguished between deworming program, school deworming program and home deworming program. It is important to remember that the deworming program is not the sum total of the two programs. Furthermore, we do not take into account the intersection of school deworming and home deworming. The effects of deworming on scores are reported in Table 5, for structural models and in Table 3, for the PSM method.

Table 3: Propensity score matching (ATET)

	Meal			Deworming		
Score	Coef.	Std. Err.	Boot.[a]	Coef.	Std. Err.	Boot.[a]
Aggregate						
K-NN[b]	2.879	2.804	2.153	−3.404	2.206	2.113
K-NN (5)[b]	2.928	1.425	1.832	−1.228	0.934	1.635
Kernel	3.757	0.77***	0.805***	−1.927	0.766**	0.802**
French						
K-NN[b]	−1.241	2.884	2.954	−5.874	2.390**	2.183**
K-NN (5)[b]	0.548	1.514	1.791	−2.762	1.224**	1.277**
Kernel	2.683	0.859***	0.888***	−1.680	0.85**	0.854**
Math						
K-NN[b]	6.965	3.232**	2.313***	0.358	2.437	2.081
K-NN (5)[b]	5.266	1.584***	1.872***	1.092	1.256	1.445
Kernel	5.062	0.841***	0.895***	−1.713	0.835**	0.789**

Note: [a]Bootstraping standard errors. [b]K-Nearest Neighbors. Significance levels: * : 10% ** : 5% *** : 1%

From the structural models, Table 5 gives the effect of the deworming in school, the deworming at home and the deworming program (in school or at home) in the first, second and last column of the table respectively. Due to a small number of observations regarding students' dewormed at school or at home, we have not been able to make estimates with the Roy model for these subgroups. That is why we do not have the ATET and ATENT for school and home deworming. Regarding school deworming Table 5 shows a positive impact on all scores both using FIML and two step. This result means that the deworming program in school has a positive effect on the entire population (ATE). The impact is more significant for the math score. If we take deworming at home, the table shows a negative impact on all scores (both with FIML and two step). This result could be explained by the nature of the drug used by rural households to deworm children. Indeed, they use either traditional or modern deworming

Table 5: Treatment effect: Deworming programs

	Deworming in school[a]		Deworming at home[a]		Deworming[b]		
Outcome	ATE (FIML*c*)	ATE (2-step*d*)	ATE (FIML*c*)	ATE (2-step*d*)	ATE	ATET	ATENT
Aggregate score	9.967***	8.845***	−1.112***	−1.020***	−31.211***	−13.155***	−23.785***
French score	5.780***	3.540***	−0.748***	−0.610***	−30.843***	−10.643***	−22.538***
Math score	10.191***	13.715***	−1.195***	−1.130***	−35.827***	−15.369***	−27.395***

Notes [a]: Dummy Endogenous Selection Model; [b]:Roy model, Full Information Maximum Likelihood; [c]: Full Information Maximum Likelihood; [d]: Two step.
Significance levels (bootstrap -100 replications-): * : 10% ** : 5% *** : 1%

drugs, and one knows that several cases of side effects in terms of fatigue or diarrhea have been reported. In addition, the use of traditional drugs are not without consequences because this could lead to the reduction of pupils' performance.

Regarding the deworming program, Table 5 shows a negative impact on all scores for the entire population, for the treated and for the non-treated. The ATET from the PSM (Table 3), shows a negative and significant effect on both the aggregate and math scores in the case of the Kernel matching methods. For the score in French, we have a negative impact for all algorithms.[7]

5. Policy implications

5.1. Impact heterogeneity

We have estimated the effects of the meal program, but it is also interesting to study these effects on some subgroups. This exercise allows us to emphasize the variability of the effects for different recipients. It is thought, for example, that the program's impact is different according to gender or according to whether there is an association around the school as we know that the operation of a school canteen requires a minimum degree of organization. In this subsection, using the structural approach only, we present the impact heterogeneity of the meal and deworming programs by gender, according to whether there is a school project in the school, an association of pupils' mother, and by grade or education level of the pupils. In Table 6, one can see that the impact of the meal program is positive on the scores for boys and girls. However, the math score of boys is higher than girls, while the French score of girls is higher than for boys. Thus one can say that the effect of the meal program is not uniform with regard to gender.

In addition, an analysis according to the existence of an association of mothers of pupils is conducted. We found that the meal program impact positively on all scores for pupils who are in schools with and without association of mothers. However, the impact of the meal program is higher in schools with association of mothers. For the grade, Table 6 shows that pupils in second grade (CP) perform better in math score while pupils in sixth grade (CE2) perform better in French score.

Table 6: Impacts heterogeneity: Meal

	Aggregate	French	Math
Gender of pupil			
Male	11.018***	7.570***	14.107***
Female	10.465***	8.905***	12.315***
Association of mother			
With	14.596***	17.640***	12.372***
Without	8.904***	6.994***	10.780***
Grade			
CE2	10.333***	9.823***	10.701***
CP	11.828***	8.104***	15.593***

Significance levels (bootstrap): * : 10% ** : 5% *** : 1%

7 The matching test reported in the appendix indicates that on average, the treated and control group show similar characteristics after matching in view of the variables used.

Table 7: Impacts heterogeneity: Deworming

	Aggregate	French	Math
Gender of pupil			
Male	−2.094***	−3.693***	−0.456***
Female	0.734***	0.123***	1.143***
School project			
With	0.668***	1.091***	0.249***
Without	−0.386***	−1.508***	0.679***
Association of mother			
With	1.659***	0.205**	5.096***
Without	−0.175***	−0.22***	−0.203***
Grade			
CE2	−0.253***	−0.704***	0.109***
CP	−0.593***	−1.769***	0.582***

Significance levels (bootstrap): * : 10% ** : 5% *** : 1%

Regarding the deworming program, Table 7 shows that the scores of girls are positive while the scores of boys are negative. This means that the program has a positive effect on girls performance unlike for boys. In addition, the table shows that the deworming program impact positively on the pupils' performance in school with association of mothers and school project. We can thus summarize by saying that having an association of mothers and a school project would lead to positive effect respectively on meal and deworming program on pupils scores. This implies that even if a meal or deworming program is implemented at school, if a number of conditions are not met, the program cannot improve pupils' performance.

5.2. LATE as policy instrument

The local average treatment's effect (LATE), introduced by Imdens and Angrist (1994), is an interesting parameter in terms of economic policies, Blundell and Costa Dias (2008). It represents the average gain to participate in a program for individuals called to receive the program due to a change in one variable that Imbens and Angrist (1994) called instrument. In the context of our study, the LATE is the average impact of benefiting from the school canteen for a group of pupils that is passed from the status of non-beneficiaries to a status of beneficiaries while providing an input school called instrument. The instruments used are: management committee, school project, school cooperative, association of parents, association of mothers, eliminating absenteeism of teachers and providing enough textbook to pupils. In fact, the management committee supports staff management, implementation and monitoring of the budget, and implements all measures conducive to the proper functioning of a school. The management committee has three specialized committees responsible for food management, community involvement and cooking in the school. Consequently the presence of a functional management committee in a school is important for the proper operation of the school canteen.

The school project is a management system tool that can improve the quality of learning. The management of a school project is in the responsibility of a management committee. The school cooperative is also another association around the school. It's a group of adults and pupils who must implement an educational project based on the practice of a community life

and cooperative. Parents of pupils are grouped around associations such as the association of parents or mothers pupils. These associations include ensuring the smooth operation of school canteens. The parents also help to supply firewood for cooking meals and providing support to the construction of kitchens and warehouses for food storage.

Table 8: Local Average Treatment Effect (LATE)[a]: Meal

	Aggregate score		French score		Math score	
Policy instruments	FIML	2-step	FIML	2-step	FIML	2-step
Management committee	33.371***	30.671***	24.146***	30.135***	44.668***	33.568***
School project	30.311***	27.281***	21.502***	26.179***	42.585***	32.306***
Cooperative School	40.349***	36.630***	34.825***	37.062***	53.155***	44.035***
Association of mothers	23.382***	20.177***	14.542***	19.587***	34.359***	23.194***
Association of parents	45.292***	43.828***	37.095***	41.081***	57.166***	50.639***
Absence	22.515***	23.290***	13.083***	13.916***	37.830***	38.798***
Textbooks	24.951***	22.451***	15.160***	21.061***	35.932***	25.349***

Note [a]: Estimate from the Roy model.
Significance levels (bootstrap -100 replications-): * : 10% ** : 5% *** : 1%

Elimination of teacher absenteeism is important. Indeed, knowing that teacher absenteeism could reduce the learning time of students and consequently their results, it is important to eliminate the absenteeism of teachers if one wants to have a better result in terms of pupils' performance. For school manuals, it is generally accepted that pupils in schools with enough textbooks perform better compared to pupils in schools with few textbooks.

From the relationship contained in Table 1 , the LATE was calculated for each instrument. Table 8 shows results on various LATE depending on the instrument used.

The association of parents gives the greatest LATE for all scores. After the association of parents, we have the school cooperative, management committee, then the school project, eliminating absenteeism of teachers and finally the provision of textbooks in sufficient quantity. In contrast, to know the best option it is important to link the cost of each input to the corresponding average effect i e the corresponding LATE. This analysis can be made through a cost-effectiveness analysis, which is a form of economic analysis that compares the relative costs and outcomes (effects) of two or more courses of action. It is important to note that the cost effectiveness analysis made here is not about comparing the canteen program with another option with the aim of increasing pupils' performance. It is a question of comparing several options in the presence of the canteen program in order to obtain the best result at a lower cost. To conduct the analysis we need the cost of the meal program and the cost related to each input. For the cost of the canteen, based on information from the WFP, the annual cost of the canteen per student is approximately 13099.966 CFA. For the school project, based on information on the region of Matam from the Ministry of Education, we estimated an approximate cost per pupil for the year 2010. The total cost for the school project for the year is 78,978,651 CFA and the total number of pupils in the region was 87,971 in 2010. So the cost of a school project per year is 897.780 CFA. Although Matam is not part of regions covered in our study, we still felt that the cost of a school project in Matam can be a rough average cost for other regions. Then, the annual cost of a canteen combined with a school project per student is 13,997.746 CFA. From Diagne (2012, page 220), we have the cost of increasing the effective learning time of pupils by the establishment of an administrative system to monitor the attendance of teachers. This cost is 111 CFA per year and per pupil. So the annual cost

of a canteen combined with monitoring the attendance of teachers is about 13,210.966 CFA. Also, Diagne (2012, page 220) gives us the cost of improving the staffing of pupils' textbooks. He estimated that the cost of a textbook that can be used for two years is 2305 CFA. In this study, we divided this cost by 2 to get the annual cost of a textbook. So, the cost of a book per year is 1152.5 CFA and the cost of a canteen combined with improving the staffing of pupils' textbooks is 14,252.066. For management committee, school cooperative, association of parents and mothers, we will only use the cost of the canteen (13099.966 CFA). Indeed, for these forms of association, a local organization can allow to make them operational without a budget from the government.

To assess the cost-effectiveness, in Table 9 we used first, the cost of each option (cost of the canteen plus the cost of the instrument) divided by the percentage of additional score which is here the corresponding LATE.

Table 9: Policy analysis

	Aggregate	French	Math
Cost			
Management committee	13,099.966	13,099.966	13,099.966
School project	13,997.746	13,997.746	13,997.746
Cooperative school	13,099.966	13,099.966	13,099.966
Association of mothers	13,099.966	13,099.966	13,099.966
Association of parents	13,099.966	13,099.966	13,099.966
Absence	13,210.966	13,210.966	13,210.966
Textbook	14,252.066	14,252.066	14,252.066
Percentage of additional score			
Management committee	33.371	24.146	44.668
School project	30.311	21.502	42.585
Cooperative school	40.349	34.825	53.155
Association of mothers	23.382	14.542	34.359
Association of parents	45.292	37.095	57.166
Absence	22.515	13.083	37.830
Textbook	24.951	15.160	35.932
Cost per percentage of additional score			
Management committee	392.555	542.531	293.274
School project	461.804	650.997	328.701
Cooperative school	324.666	376.165	246.448
Association of mothers	560.258	900.836	381.267
Association of parents	289.233	353.146	229.156
Absence	586.762	1009.781	349.219
Textbook	571.202	940.109	396.639

Table 9 shows that the most cost effective option is the combination of the canteen and the association of parents. If we take the aggregate and French scores, the results of the analysis show that after association of parents, one observes school cooperative, the management committee, the school project, the association of mothers, the textbook and the eliminating of absenteeism of teachers. Meaning that eliminating teacher absenteeism is the less profitable option when one wants to increase the aggregate and French scores of pupils to 1 point as compared to other options. If one takes the math score, the less profitable option is improving the stock of pupils' textbooks.

6. Conclusion

In this study, we have assessed the effects of school feeding and deworming programs on school performance measured through aggregate score, French score and mathematic score. We use an observational database of 162 schools collected in four rural areas in Senegal: Diourbel, Fatick, Kolda and Sedhiou. We use both programs evaluation PSM and selection frameworks. The structural framework has the advantage of allowing for the joint estimation of a performance equation and a selection mechanism while estimating the effect of the program. Our study distinguishes between deworming at home and deworming in school. It is important to emphasize that the deworming program is not the sum total of the deworming in school and at home. Furthermore, we do not take into account the intersection of deworming in school and at home. Also, in terms of policy analysis, this study uses the local average treatment effect (LATE) as an instrument to analyze economic policy.

The results show that the meal program can significantly increase pupils performances. Also the impact is more important for pupils who are in school with association of mothers. Deworming in school has a positive effect on pupils performance while deworming at home has a negative impact on pupils' scores. Our reading is that it is essential to assist rural households to deworm children in households. This will avoid the use of traditional medicines (that have adverse consequences on pupils' performance) for deworming children. In addition, the deworming program increases significantly the scores of pupils who are in school with association of mothers or school project. Among the determinants of test scores, school manual, class size, age of teacher and pupil's age display a non-linear relation with the scores. Particularly, class size displays a U-shaped relation meaning that the performance (aggregate and math scores) decreases with increasing class size which is consistent with the findings in the literature on performance. However, starting from a certain threshold, the performance increases with respect to class size. In the context of rural Senegal, large classes may positively impact test scores thereby emphasizing the crucial role of group effect in shaping pupils' knowledge. Among other controls, education spending, household size per equivalent adult, water point, koranic school, early childhood institution, existence of a high school in the village have a positive effect on pupil's performance, while controls such as disturbed courses, gender of teacher, absenteeism of teacher and parents school have negative impact on scores. In terms of policy analysis, results show that the impact of programs varies depending on the gender of the pupil, the existence of association around the school and the educational level of the pupil. Relying on the LATE as policy instrument, it appears that the combination of school canteens and the association of parents is the most cost effective option for improving scores.

References

Adelman, S.W., D.O. Gilligan and K. Lehrer (2008) How Effective are Food for Education Programs? A Critical Assessment of the Evidence from Developing Countries. Food Policy Review 9, International Food Policy Research Institute IFPRI.

Blundell, R. and M.C. Dias (2008) Alternative Approaches to Evaluation in Empirical Microeconomics. The Institute for Fiscal Studies Working Paper 26.

Clotfelter, C.T., H. Ladd and J. Vigdor (2006) Teacher-Student Matching and the Assessment of Teacher Effectiveness. NBER Working Papers 11936.

CONFEMEN (1999) Les Facteurs de l'Efficacité dans l'Enseignement Primaire: Les Résultats du Programme PASEC sur Neuf Pays d'Afrique et de l'Océan Indien. Secrétariat Technique Permanent de la CONFEMEN.

CONFEMEN (2007) Evaluation PASEC Sénégal. Secrétariat Technique Permanent de la CONFEMEN.

Cueto, S. and M. Chinen (2008) 'Educational Impact of a School Breakfast Programme in Rural Peru', International Journal of Educational Development, 28: 132-48

Davies, P. (2003) The School Class Size Debate: Does Size Matter? In Education Forum, Briefing Paper 6.
Deaton, A. S. (2010) 'Instruments, Randomization, and Learning about Development', Journal of Economic Literature, 48: 424–55.
Diagne, A. (2012) Le Sénégal face aux Défis de l'Education: Enjeux et Perspectives pour le XXIème Siècle. Karthala et CRES edn.
Felouzis, G. (1997) L'efficacité des Enseignants, Collection Pédagogie d'Aujourd'hui, Presse Universitaire de France, Paris.
Glass, G.V. and M. Smith (1979) 'Meta Analysis of Research on Class Size and Achievement', Educational Evaluation and Policy Analysis, 1: 2–16.
Heckman, J. (1976) 'The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables, and a Simple Estimator for Such Methods', Annals of Economic and Social Measurement, 5: 475–92.
Heckman, J. (1979) 'Sample Selection Bias as a Specification Error', Econometrica, 47: 153–611.
Heckman, J. J. (2010) 'Building Bridges Between Structural and Program Approaches to Evaluating Policy', Journal of Economic Literature, 48: 356–98.
Imbens, G. W. (2010) 'Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)', Journal of Economic Literature, 48: 399–423.
Imbens, G.W. and J.D. Angrist (1994) 'Identification and Estimation of Local Average Treatment Effects', Econometrica, 62: 467–75.
Jarousse, J. and A. Mingat (1989) Les Effets des Conditions d'Enseignement sur les Acquisitions des Elèves à l'Ecole Primaire: Le Cas du Togo. IREDU, Dijon.
Kvalsvig, J., R. Cooppan and K. Connolly (1991) 'The Effects of Parasite Infections on Cognitive Processes in Children', Annals of Tropical Medecine and Parasitology, 85: 551–68.
Ma, X. (2007) Gender Differences in Learning Outcomes, Document de Référence pour le Rapport Mondial de Suivi sur l'EPT.
Miguel, E. and M. Kremer (2004) 'Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities', Econometrica, 72: 159–217.
Piketty, T. (2004) L'impact de la Taille des Classes et de la Ségrégation Sociale sur la Réussite Scolaire dans les Ecoles Françaises: Une Estimation à Partir du Panel Primaire 1997, Document de travail EHESS, Paris.
Rivers, J.C. and W.L. Sanders (2002) Teacher Quality and Equity in Educational Opportunity: Findings and Policy Implications, In L. T. Izumi, W. M. Evers (Eds.). Stanford, CA: Hoover Press.
Roy, A. (1951) 'Some Thoughts on the Distribution of Earnings', Oxford Economic Papers, 3: 135–46.
Schwille, J., T. Eisemon, F. Ukobizoba, R. Houang, D. Kwon and R. Prouty (1991) Is Grade Repetition Always Wasteful? New Data and Unanswered Questions, Harvard Institute for International Development, Cambridge, MA.
UNESCO (2000) Forum Mondial sur l'Education. Education pour Tous: Tenir nos Engagements Collectifs, UNESCO Dakar.
Vermeersch, C. and M. Kremer (2004) School Meals, Educational Achievement and School Competition: Evidence from a Randomized Evaluation, Oxford University, Mimeo.