An analysis of school dropout in Mozambique, 2014-15

Felix Mambo, Maren Basso, Vincenzo Salvucci, Ricardo Santos

[DRAFT – NOT FOR CITATION]
February 2019

Abstract

Educational attainment in Mozambique is one of the lowest in the world, only about 2,5 years on average. The primary school enrolment is 88%, but the survival rate is only 37,5%. Despite the high economic growth rates experienced until 2015, Mozambique's poverty rates remain high, when compared to its neighbours' and the relation between poverty and educational attainment seems to be strong but needs to be further explored. Using data from the Mozambican Household Budget Survey 2014/15, we first study the determinants of school dropout; then, we study in more detail the variables that are more associated with school dropout in the year of the survey, exploiting the panel structure of the 2014/15 Household Budget Survey. Finally, we also present the results of a propensity score matching analysis using the rural community data available in the same survey.

Regarding school dropout, we find that age, child labour, household head's gender and education, access to services are particularly associated with the probability of going to school and not abandoning it. We also find that if a girl gets married or becomes pregnant this highly increases the probability of dropout. In 2014/15 the result of teen girl pregnancy as one of the most important characteristics associated with dropout is confirmed. When the propensity score matching analysis is performed for rural communities, we find that building a new school or renovating an existing one, and building or improving a water facility, can have an effect of reducing the probability of dropping out. Also, setting up a social protection programme seems to influence school dropout both in the past and in the year of the survey. This is particularly interesting given that the Mozambican Government is discussing and re-designing its approach to social protection and its social protection programmes to include more communities and people.

1. Introduction

Despite the high economic growth rates experienced until 2015, Mozambique's poverty rates remain fairly high (46.1 percent of the population) and higher than those of most neighbouring countries (DEEF, 2016). Moreover, its population has experienced relatively high growth rates, so that the absolute number of poor people in 2014/15 is comparable to the one observed in 1996/97 – about 12 million people (DEEF, 2016). At the same time, inequality has also rapidly increased between 2008/09 and 2014/15, particularly in urban areas and in the more developed southern region (DEEF, 2016). A simple analysis of the determinants of poverty performed on the same data has shown that the relation between poverty and education seems to be strong and increasing in the levels of education achieved (Ibraimo and Salvucci, 2017). However, the educational attainment in Mozambique is one of the lowest in the world, being 2,5 years on average. The primary school enrolment is 88%, but the survival rate is only 37,5%.

In this paper we contribute to the literature on school participation and school dropout in Mozambique exploring first, the determinants of school dropout, using the most recent household budget survey available; second, we take advantage of the panel structure of this household budget survey to further explore the drivers of school dropout and establish a more precise causation link between dropout and time-varying child/household characteristics; finally, we study the relation between school dropout and the characteristics of the communities where school-age children live, using propensity

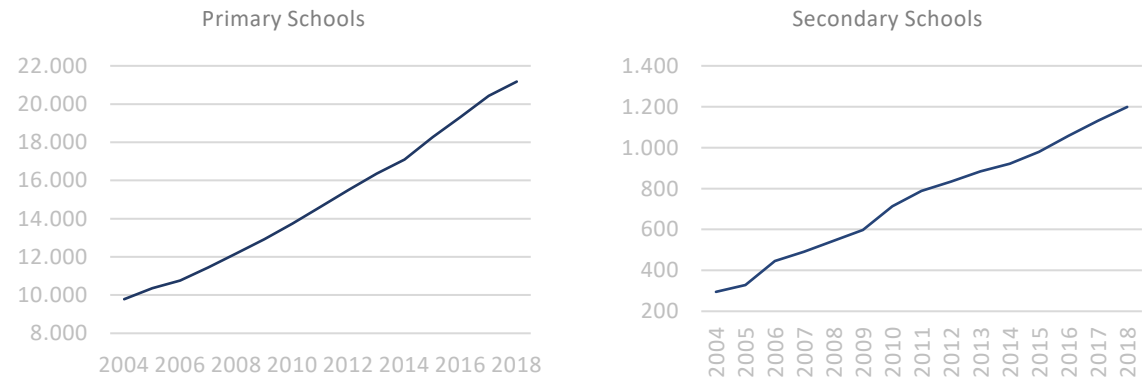
score matching methods to evaluate the relevance of changes occurred at the community-level like the building or improvement of new and existing infrastructures (such as education, water, electricity, roads or health infrastructures), or the introduction of social protection programmes.[1]
We find that age, child labour, household head's gender and education, access to services and distance to school are particularly associated with the probability of dropout. We also find that if a girl gets married or becomes pregnant this highly increases the probability of dropout. When analysing the dropout determinants for 2014/15 using the panel structure of the data, we observe that teen girl pregnancy is confirmed as one of the most important characteristics associated with dropout.
When the analysis is performed at the community level, we find that building a new school or renovating an existing one, building or improving a water facility, or setting up a social protection programme can have an effect on reducing the probability of dropping out. The paper develops as follows: in Section 2 we describe the context and motivation of the paper; Section 3 introduces the data and the variables used in the analysis; Section 4 presents the methodology, while Section 5 describes the main results; Section 6 concludes.

2. Context

With nearly half of its population estimated to be of school age, 48% with ages ranging from 5 to 25 years old according to the population projections of the Mozambican Statistics Bureau (INE), Mozambique must look to education as one of its most critical sectors.
A review of some of the country's statistics of primary education, suggests that there is a significant investment by the Government of Mozambique (GoM) and its development partners in improving the access to education, from its beginning.

This can be shown by the investment alone in school building by the GoM. In 15 years, from 2014 to 2018, the number of primary schools in the country more than doubled, with more than one thousand being built each year in the 2015-2017 period. Growing from a lower base, the number of secondary schools quadrupled during the same period.

Figure 1:Number of Schools

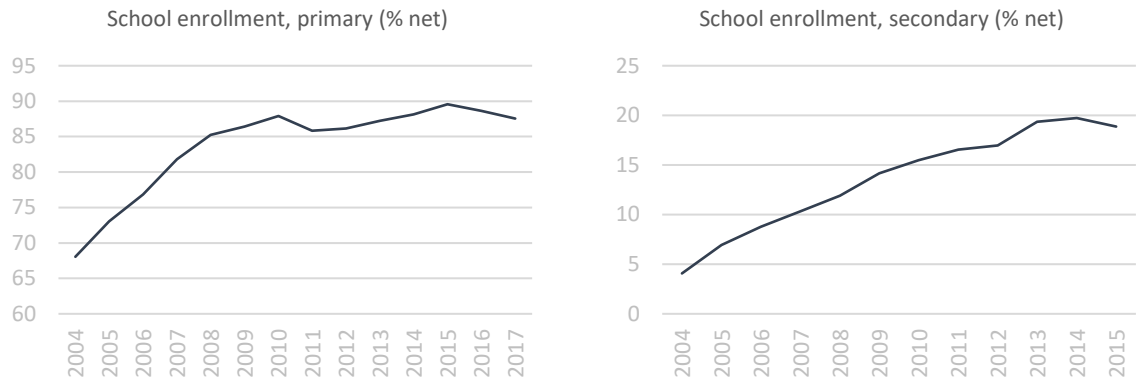


Source: GoM Ministry of Education (2018), Education Management Information System

Together with this investment, primary school enrolment has risen significantly, from 67.3% in 2004 to 86.9% in 2010. From then, improvement has come at a slower pace, still improving to 89.6% in 2015.

[1] The determinants of school participation are not analyzed here as they were already studied in a previous paper by Mambo (2017).

Figure 2:Net School Enrolment Rates, Primary and Secondary (%)



Source: World Development Indicators, accessed in 25-01-2019

UNICEF(2014) acknowledges that the investments put in place by the GoM led to this improvement. It highlights, in particular, measures directed to households and their demand for education, such as the abolition of school fees in 2003/04 or the provision of direct support to schools and free textbooks, along with measures directed to improving education supply, such as investments in classroom construction and teachers. The resulting increase in primary and secondary enrolment was also , according to their report, accompanied by a significant improvement in intake of children aged 6 years old, up to over 80%. Gender parity in primary and secondary enrolment has also reported to have improved.

However, there are still many challenges remaining in the provision of and access to basic education in Mozambique. UNESCO (2012) presents us estimates of approximately 1.2 million (or 23%) of out of school primary and secondary school aged children. Of these, they report 775,000 primary school aged children. They also report that 55% of these children are girls. De Walque & Valente (2016) highlight that the significant improvement in enrolment did not translate into completion of the very first level of education, primary school, by most Mozambican children. This has a repercussion on education achievement. De Walque & Valente (2016) report that young men age 18 in 2011 had achieved, on average, 6.3 years of education, while young women of the same age achieved even less, 5.5 years. In both cases, this is less than the seven years of full primary education in Mozambique. More precisely, only close to half of 18 years old men and 41% of 18 years old women had completed primary school, something a child entering at age 6 and not repeating would have achieved at age 13. The rate of completion of primary schooling is reported to be much lower among women, but also in rural and central and in the northern areas of the country. Only 34% for 18 years old men and 23% of 18 years old women had completed EP2 in rural areas. Fox et al. (2012) reported even lower official statistics of upper primary schooling (EP2) completion rate, especially in rural areas where even at age 19 it is only about 14% for males and 8% for females.

This is a worrying reality, acknowledged as such by the Minister of Education herself in a public speech in April 2018:

> *"From 2007 to 2016, students who complete seven years of schooling annually represent only 45.6% on average. And only 30% of the students who enter 1st grade annually complete the 7th grade in seven years. (...) Only last year about 550,000 primary school pupils stopped going to school as failure rates were also at alarming levels. In the 10th class, for example, the failure rates in the last five years were 46 percent on average".*
>
> *in APA-Maputo, Mozambique, April 28, 2018.*

These realities call out and highlight the need to uncover what can be the factors behind the high dropout rates. Among these we find the economic costs of education. As indicated above, studies such as Fox et al. (2012) suggest that the abolition of direct costs of education in the 2004/08 reform of the Mozambican Education System is behind the great increase in the school enrolment of children age 6 to 19. This despite an earlier finding by the World Bank (2005) suggesting that, while the direct costs of education have small to no influence on the likelihood of enrolment in primary school.

Other factors, such as distance to school, parents' education, particularly the mother's, gender of the head of household, household income (proxied by consumption) and children characteristics (age, gender, disabilities and other vulnerabilities, whether or not they live with their fathers) can significantly affect it (World Bank 2005, UNICEF 2010, UNESCO 2012). UNESCO (2012) highlights also other factors, related to poverty and socio-cultural norms that keep children from finishing school, such as early marriage and pregnancy, as well as factors related to quality of education such as lack of safe school spaces, overcrowded classrooms, and a lack of adequate number and quality teachers.

Aspects of quality of education (or lack thereof) are also highlighted by the respondents of the 2014/15 Household Survey. As per Table 1, below, households reported on problems of the schools their children attended. A majority reported lack of desks and chairs, one third reported infrastructures in very poor conditions and around one in ten reported lack of books, other school material or even teachers.

Table 1: Problems experienced in the school the child attends

	%		%
Lack of desk/chairs	50.46	*Lack of teachers*	8.86
Infrastructures in very poor conditions	34.37	*Other problems*	2.49
Lack of books	13.26	*Corruption*	1.20
Lack of school material	12.72	*No problems*	*42.35*

Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).

When asked why their children were not attending school, the respective families highlighted a set of reasons as per Table 2, below,

Table 2: Reasons why child does not attend school

	%		%
School is useless/no interest	49.85	*No available places*	3.55
Other problems	12.17	*Pregnancy*	3.14
School is very costly	9.13	*Failed*	1.72
School is very far	7.81	*Next grade doesn't exist*	1.20
Got married	5.93	*Reached the desired grade*	0.66
Works (at home or outside)	4.31	*Child is very young*	0.46

Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).

These factors can partly, or together, be behind the worrying levels of school dropout. This study seeks to contribute to validate these insights and to add further knowledge on the factors that lead to them.

3. Data and variables used

The analysis is based on the Mozambican Household Budget Survey 2014/15 (IOF14). The data was collected by the National Statistics Institute (INE), throughout a one-year period, between August 2014 and August 2015. Households were interviewed three times in the first (mid-August to mid-November 2014), second (mid-November 2014 to mid-February 2015) and fourth quarter of that 12-month period (mid-May to mid-August 2015). The IOF14 dataset contains data from a representative sample of around 11,000 households (11,505 in the first quarter of the survey, 10,368 in the second and 11,315 in the fourth quarter). The sample is representative of the Mozambican population as a whole, but

also of rural and urban populations and of those in each of the eleven provinces of the country, including Maputo City. The main household questionnaire is accompanied by a community questionnaire for rural areas only.

The IOF14 provides information on a wide set of individual, household and community characteristics, including demographics; education; health; employment; daily, monthly and annual expenditures; durable goods, land and livestock; receipts and transfers. In our analysis we make use of child-specific variables like age, sex, birth order, mother/father alive, number of siblings, child labour, chronic deficiencies; household level variables like household head's gender, age, education level and occupation; socio-economic characteristics such as consumption level, access to water/sanitation/electricity, ownership of durable goods, distance to school; community characteristics like transportation access, presence of schools in the community, natural shocks; time and geographic controls.

The descriptive statistics for the variables used are presented in Table 3.

Table 3. Descriptive statistics for variables used in the analysis

Variable	*Description*	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Dropout anytime	*Abandoned school at any point in time*	44,983	0.13	0.34	0	1
Dropout 1415	*Abandoned school between 2014 and 2015*	44,983	0.05	0.22	0	1
Sex	*Sex*	44,983	0.51	0.50	0	1
Age	*Age*	44,983	11.17	3.26	6	17
Birth order	*Birth order*	44,983	2.12	1.25	1	16
Orphan	*Orphan (either father and/or mather died)*	44,983	0.16	0.37	0	1
Not in the hh	*Not in household (either father and/or mather do not live in the household)*	44,983	0.47	0.50	0	1
Child works	*The child works*	44,983	0.28	0.45	0	1
Disability	*The child has some form of permanent disability or sickness*	44,983	0.01	0.11	0	1
Late	*The child is attending a school grade that is more than two years lower than the one corresponding to his/her age*	44,791	0.68	0.47	0	1
Married	*The child is married*	44,983	0.11	0.31	0	1
Pregnant	*The child is pregnant*	44,983	0.00	0.07	0	1
Head age	*Household head age*	44,983	45.02	12.72	14	95
Head woman	*Household head is a woman*	44,711	0.27	0.44	0	1
Head no educ	*Household head has no education*	44,983	0.27	0.45	0	1
Head 5y educ	*Household head has completed 5 years of education*	44,983	0.40	0.49	0	1
Head 7y educ	*Household head has completed 7 years of education*	44,983	0.16	0.37	0	1
Head 10y educ	*Household head has completed 10 years of education*	44,983	0.10	0.29	0	1
Head 12y educ	*Household head has completed 12 years of education*	44,983	0.05	0.21	0	1
Head +12y educ	*Household head has completed more than 12 years of education*	44,983	0.03	0.16	0	1
Head in agriculture	*Household head works in agriculture*	44,750	0.60	0.49	0	1
Poor	*The household is poor (monetary poverty)*	44,983	0.47	0.50	0	1
Water source	*The household has access to a safe water source*	44,973	0.58	0.49	0	1
Sanitation	*The household has access to improved sanitation*	44,973	0.33	0.47	0	1
Roof	*The household lives in a house with good quality roof*	44,973	0.49	0.50	0	1
Electricity	*The household has access to electricity*	44,973	0.35	0.48	0	1
Durable goods	*The household has access to at least three durable goods out of a list of ten basic items of common use*	44,869	0.57	0.50	0	1

Access to transport	*The community has access to transportation means*	43,886	0.67	0.47	0	1
School	*There is at least one school in the community*	43,877	0.91	0.29	0	1
Flood	*The household lives close to the area affected by the 2015 flood*	44,983	0.29	0.45	0	1
Distance school 0-15	*Distance to school is less than or equal to 15 minutes*	44,983	0.48	0.50	0	1
Distance school 16-30	*Distance to school is between 16 and 30 minutes*	44,983	0.29	0.45	0	1
Distance school 31-60	*Distance to school is between 31 and 60 minutes*	44,983	0.17	0.38	0	1
Distance school 60+	*Distance to school is more than 60 minutes*	44,983	0.06	0.24	0	1

Note: the descriptive statistics presented consider the observations for each child in the various survey quarters as stacked. There are about 18,000 children included in the analysis, not all of them were observed in all three survey quarters. Statistics for temporal and geographic controls not shown.
Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).

4. Methodology

Using the data described in Section 2, we first study the determinants of school dropout. In this case, we use a very simple probit model to analyze those child, household and community characteristics that show a stronger association with the probability of having abandoned school at any point in time. The model is simply described by the following equation:

$$Y_i^* = \alpha_i + \beta_i X_i + \varepsilon_i, \quad (1)$$

where for each child *i*, *Y** is a binary variable indicating whether the child has dropped out at some point in time, and *X* represents a set of child, household and community characteristics.

In order to analyze in more detail those children that abandoned school between 2014 and 2015, we also study the characteristics that are more associated with school dropout over the survey period, exploiting the panel structure of the IOF14. Indeed, from the IOF14 we know whether a child is at school during August-November 2014 (last school quarter in the Mozambican school system) and not at school in the first and third quarter of the subsequent school year (corresponding to the second and fourth survey quarters). In this case, we use a fixed-effect linear panel regression model that allows us to look more specifically into the time-varying variables affecting school dropout in the survey year.

The linear probability model estimated is the following:

$$Y_{it} = \alpha_{it} + \beta_{it} X_{it} + \delta_i + \varepsilon_{it}, \quad (2)$$

where for each child *i* in quarter *t*, *Y* is a variable indicating whether the child has abandoned school over the survey year 2014/15, and *X* represents a set of child, household and community characteristics. The model is estimated using child fixed-effect.

Finally, we also present the results of a propensity score matching analysis using the community data for rural areas to check whether some changes occurred at community level particularly influence school dropout. More specifically, we consider the impact of building or renovating a school, a water facility, an electricity facility, a road, a health center, and of the existence of a social protection programme. In this case, we use the nearest neighbor propensity score matching method with five "neighbors" being selected from the control group for each treated observation.[2]

5. Results

As discussed in Section 3, we first study the determinants of school dropout using the sample of kids who declare having enrolled in school at least once in the past. The marginal effects obtained from the probit model estimation used to analyze those child, household and community characteristics that

[2] Both the *psmatch2* and the *teffects psmatch* Stata commands were used in the analysis, showing qualitatively similar results.

show a stronger association with the probability of having abandoned school at any point in time are shown in the first column of Table 4.[3]

Regarding dropout determinants, we find that as age increases this is linked to a higher probability of abandoning school. As expected, if one of the parents does not live in the household this is associated with a higher probability of abandoning school, and the same is true if the child works or if he/she has a permanent illness/disability. If a child is older than two years compared to the grade he is supposed to be in, this is associated with a higher probability of abandoning school. Early marriage and especially pregnancy highly increase the chances of dropping out of school. Interestingly, if the household head is a woman, this is linked to a lower probability of abandoning school. The level of education of the household head is also strongly and consistently associated with lower probabilities of dropping out, and the same applies to access to some basic services such as safe water, quality sanitation, good quality roof or electricity. Greater distance to school also appears as being associated with higher probabilities of dropout. In the estimation presented in column 1 of Table 4 we make use of temporal (quarter) and geographic (province/rural-urban) controls. It emerges that living in some of the northern and central provinces (especially Niassa and Tete) is significantly associated with higher probabilities of dropping out compared to the capital Maputo.

Additional analyses are carried out in Table A.1 of the Appendix for girls and boys, for kids aged 6-13 and 14-17, and for kids in primary (up to the fifth and up to the seventh grade) and secondary school. Among the most noticeable differences, it can be seen that: being orphan seems to be linked to dropout only for boys, and that the coefficient associated with pregnancy almost doubles its size for the group of kids aged 14 to 17; also, if the household head is a woman, this is linked to a lower probability of abandoning school especially for younger kids (6 to 13 years old) and kids in primary school; when the distance to school is very high (above 60 minutes walking) this especially affects girls and secondary school kids.

In column 2 of Table 4, we also analyze in more detail the characteristics that are more associated with school dropout over the period 2014-15, exploiting the panel structure of the IOF14. As introduced in previous sections, the IOF14 provides information on whether a child is at school during August-November 2014 (last school quarter in the Mozambican school system) and not at school in the first and third quarter of the subsequent school year (corresponding to the second and fourth survey quarters). This estimation procedure allows us to provide a more precise indication about the causality mechanism at work, and the main results from the child fixed-effect linear panel regression model implemented[4] are the following: if either the father or the mother, or both, stop living in the household, this is associated with a 3.8 percentage points higher probability of dropping out; a similar effect is found if the child starts working; if a child gets married he/she has about 2.5 percentage point higher probability of abandoning school. Perhaps the most impressive result in this analysis is the one for when a girl gets pregnant: in this case, the probability of dropout increases by 29 percent.[5]

---

[3] We combine here all those kids who declare that they went to school but are not in school at the moment of the interview. Given that in our sample we also include teenager aged above the mandatory school age for Mozambique, we also check if the reply to the question on the reason for not going to school is: 'I quit because I reached the desired level of education'. In this case, we do not consider these individuals as having abandoned school because they quit school voluntarily.

A linear probability model was also tested without obtaining qualitatively different results. Therefore, only the marginal effects from the probit model estimation are shown here.

[4] Here we implement a linear probability model because the Stata command *xtreg* allows to use sample weights in the estimation, whereas the panel version of the probit command does not. Our choice was also reinforced by the fact that the predicted probabilities never exceed the range [0,1].

[5] In the Mozambican school system, when a girl gets pregnant she is obliged to switch to attend night school courses. However, in many areas of the country – due to availability of transportation, personal security, among other issues – this entails that most of the time the girl has no other choice but to abandon school entirely. One possible objection to this result could be that girls who become pregnant are very likely not to go to school in the immediate months after giving birth or during pregnancy, but that they could go back to school after some

In this case as well, we provide detailed additional estimations in Table A.2 of the Appendix for girls and boys, for kids aged 6-13 and 14-17, and for kids in primary (up to the fifth and up to the seventh grade) and secondary school. We can notice that in this case the coefficient for the variable gender of the household head is positive and significant for kids aged 6-13 and for kids in primary school. This apparently contradictory result can be interpreted in the following way: in general, if the household head is a woman this is associated with a lower probability of school dropout for the whole population of kids 0-17; however, if the household head gender changes from man to woman from one quarter to another, which is what is measured in the panel regression presented in column 2 of Table 2, this increases the chances for younger kids of abandoning school. This makes sense, as the change in the household head gender from man to woman is frequently associated to the death, migration or household abandonment of the (male) household head, which can lead to the decision to drop one or more kids from school. We can also observe that if either the father or the mother, or both, stop living in the household from one quarter to another because of death, migration or household abandonment of any of them, this mostly affects girls and younger kids (aged 6-13 and/or enrolled in primary school). Regarding pregnancy, it is interesting to observe that the probability of dropping out from school increases immensely if the girl is less than 13 years old (about 82%) and becomes 34% for girls in secondary school.

Table 4. Determinants of dropout at any point in time and dropout in the period 2014-15

VARIABLES	*(1)* *Dropout anytime*	*(2)* *Dropout 1415*
Sex	0.0106*	
	(0.00565)	
Age	0.0298***	
	(0.00127)	
Birth order	-0.00451*	
	(0.00247)	
Orphan	0.00848	-0.00221
	(0.00738)	(0.0108)
Not in the hh	0.0387***	0.0384**
	(0.00814)	(0.0155)
Child works	0.0576***	0.0323***
	(0.00517)	(0.00743)
Disability	0.0493***	
	(0.0188)	
Late	0.0265***	-0.0512***
	(0.00793)	(0.00824)
Married	0.0592***	0.0254***
	(0.00715)	(0.00594)
Pregnant	0.298***	0.291***
	(0.0319)	(0.0472)
Head age	-8.33e-05	
	(0.000247)	
Head woman	-0.0324***	0.176*
	(0.00803)	(0.0899)
Head 5y educ	-0.0154**	
	(0.00711)	
Head 7y educ	-0.0280***	
	(0.00993)	

time. This is possible in theory, but given our data we cannot observe whether this happens or not for the girls in our sample. However, we can provide some evidence using the data in our hands that it is quite unlikely that girls with young babies go back to school. Among the girls in school age with young babies (0-1 years old), the percentage of those going to school is about 12%, significantly lower than the rate for the entire population of girls in school age (about 84%).

Head 10y educ	-0.0510***	
	(0.0103)	
Head 12y educ	-0.0729***	
	(0.0132)	
Head +12y educ	-0.0729***	
	(0.0162)	
Head in agriculture	-0.00547	
	(0.00753)	
Poor	-0.00462	0.00600
	(0.00580)	(0.00677)
Water source	-0.0187**	
	(0.00856)	
Sanitation	-0.0266***	
	(0.00885)	
Roof	-0.0194**	
	(0.00918)	
Electricity	-0.0316***	
	(0.00978)	
Durable goods	-0.0125	
	(0.00780)	
Access to transport	0.00490	
	(0.0101)	
School	-0.0128	
	(0.0129)	
Flood	-0.00675	0.0105
	(0.0105)	(0.0138)
Distance school 16-30	0.0172**	
	(0.00770)	
Distance school 31-60	0.0351***	
	(0.0105)	
Distance school 60+	0.0252*	
	(0.0133)	
Constant		0.0765***
		(0.0264)
Temporal controls	Yes	Yes
Geographic controls	Yes	No
Observations	43,102	44,520
Number of kids		18,056

Standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

Note: There are about 18,000 children included in the analyses, not all of them were observed in all three survey quarters.
Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).

5.1 Results at rural community level using propensity score matching

Here we present the results from the analyses of the relation between school participation/dropout and some specific characteristics of the communities where school-age children live, using propensity score matching methods to evaluate the relevance of changes occurred at the community level, like the building of new schools and/or improvement of existing ones, the construction of new water/electricity/road/health facilities or the introduction of social protection programmes.

The IOF14 includes a community questionnaire module, only for rural areas, in which a series of relevant questions with respect to the community are asked to key informants on topics such as demographics, economics and infrastructures, education and health, social protection and agriculture. In this module, one of the most interesting questions for our purpose is the one about improvements in infrastructures occurred in the last two years. It is particularly important in the case of education infrastructures to assess if this has an effect on school dropout, and to assess whether other investments in infrastructures or other programmes show a positive statistical association as well.

Until now we used data on all children and tried to control reasonably comprehensively across children based on a series of time invariant factors or based on the panel structure of the survey. In this

subsection we limit the analysis to rural communities and compare reasonably similar kids based on their observables using propensity score matching methods. Finally, we compute the average treatment effect on the treated (ATT) (Rosenbaun and Rubin, 1983). In our case, the 'treatment' is given by the building of new schools and/or improvement of existing ones, the construction of new water/electricity/road/health facilities or the introduction of social protection programmes. We designate the treatment group as the group of school-age children surveyed in rural communities where these changes occurred. However, given that the population was not actually 'treated' in the sense of a formal experiment, we prefer to define this group of observations as 'exposed' group, rather than 'treated'. The control group is thus the group of school-age children surveyed in rural communities where these changes did not take place. To compute the ATT, we assume that all relevant differences between school-age children in the exposed and non-exposed communities are captured by their observables and seek to select a control group from the non-exposed pool of observations for which the distribution of the observables is as similar as possible to the distribution of the observables in the treated group. Various matching methods exist. Here, we implement a nearest neighbor matching (Leuven and Sianesi, 2003). Robustness checks (not shown) are also performed using one-to-one matching, without obtaining qualitatively different results. In this case, we use the nearest neighbor propensity score matching method with five 'neighbors' being selected from the control group for each treated observation.[6] The observable characteristics selected as controls are the same used in the analysis of school participation and school dropout presented in Section 3.

The tests performed (not shown) suggest that matching reduces differences in sample means between the control and treated group with respect to the observable characteristics. Following Leuven and Sianesi (2003), we do not use sample weights to implement the matching procedure.

The descriptive statistics for the variables introduced in this section are in Table 5, whereas the results – ATT and statistical significance – are presented in Table 6. We find that building or renovating a school can lower probability of dropping out by about 2 percentage points (panel a). The effect is slightly bigger (4.6 pp) for building or improving a water facility in the community (panel b), and slightly smaller (1.9 pp), but only significant at 10%, for building or improving an electricity network facility (panel c). Building or improving a health facility in the community seems to have contributed to lower the probability of dropout in 2014/15, by about 1.2 pp. A slightly bigger effect is found for the case in which in the community a social protection programme has been put in place (3 percentage points lower probability of dropout, panel d); in this case the effect is also significant for the case of dropout in 2014/15 (about 2 percentage points). No effect is found in the case of building or improving a road (panel e).

Table 5. Descriptive statistics for the additional variables used in the propensity score matching estimations

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
School built/renovated	17,935	0.311108	0.46296	0	1
Water facility built/renovated	17,935	0.269464	0.443694	0	1
Electricity facility built/renovated	17,935	0.092646	0.289944	0	1
Social protection programme	17,935	0.216347	0.411765	0	1
Road facility built/renovated	17,935	0.117457	0.321973	0	1
Health facility built/renovated	17,736	0.48191	0.499687	0	1

[6] Both the *psmatch2* and the *teffects psmatch* Stata commands were used in the analysis, showing qualitatively similar results, but only the results obtained using the *teffects psmatch* are shown in what follows. The main difference comes from the fact that *teffects psmatch* implements a method to estimate the standard errors of the estimator that matches on estimated treatment probabilities as derived by Abadie and Imbens (2012). The reason for using nearest the neighbor matching method is motivated by the fact that we have a pool of controls that is relatively big compared to the pool of treated observations and could thus obtain better matches compared to the one-to-one matching. Nonetheless, results obtained using the one-to-one matching with replacement are qualitatively not dissimilar from the ones presented.

Table 6. Average treatment on the treated (ATT)

		Dropout anytime	*Dropout 1415*
a)	*School built/renovated*	-0.023***	-0.004
		(0.006)	(0.004)
b)	*Water facility built/renovated*	-0.046***	-0.005
		(0.015)	(0.004)
c)	*Electricity facility built/renovated*	-0.019*	-0.004
		(0.010)	(0.005)
d)	*Social protection programme*	-0.032***	-0.019***
		(0.006)	(0.004)
e)	*Road facility built/renovated*	-0.004	0.001
		(0.007)	(0.005)
f)	*Health facility built/renovated*	-0.009	-0.012**
		(0.009)	(0.006)

Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).

6. Conclusions

In this study we analysed the determinants of school dropout occurred at any point in time and in 2014/15 in Mozambique using the 2014/15 Household Budget Survey data, the most recent household budget survey that is available. We also presented results for a propensity score matching analysis that used the rural community data available in the same survey.

As discussed in the introduction, education is one of Mozambique's most critical sectors. The sector receives significant investments by Government and development partners; the number of primary schools in the country more than doubled in recent years and the number of secondary schools quadrupled during the same period. At the same time, and following these investments, primary school enrolment has risen significantly, also thanks to the abolition of school fees and the provision of direct support to schools and free textbooks. However, school dropout remains a big challenge for the Mozambican education system, as highlighted for example by De Walque and Valente (2016), Fox et al. (2012) and by the Mozambican Minister of Education (2018). Aspects of quality of education services provided, such as lack of desks and chairs, infrastructures in very poor conditions and lack of books, other school material or even teachers, are often highlighted by the respondents of the 2014/15 Household Budget Survey and these factors can partly be behind the worrying levels of school dropout. However, it should also be stressed that when asked about the reasons for not attending school, a very high percentage of children or family members replied that school is useless/there is no interest. As discussed, these factors can partly be behind the worrying levels of school dropout, and this study has tried to contribute to add further knowledge on the factors that lead to the outcomes described.

In the first set of the analyses presented in this paper, we studied the child, household and community characteristics that show a stronger association with the probability of having abandoned school at any point in time using the sample of kids who declare having enrolled in school at least once in the past. We find that age, child labour, household head's education, access to services are particularly associated with the probability of going to school and not abandoning it. Moreover, our results suggest that early marriage and especially pregnancy highly increase the chances of dropping out of school, whereas if the household head is a woman, this seems to be linked to a lower probability of abandoning school. As expected, greater distance to school also appears as being associated with higher probabilities of dropout.

Analysing in more detail the characteristics that are more associated with school dropout over the period 2014-15, we find that if either the father or the mother, or both, stop living in the household, this is associated with a higher probability of dropping out and a similar effect is found if the child starts working. Teen marriage also shows a relation with a higher probability of abandoning school, but the most impressive result is the one for early pregnancy, which seems to be associated with an increase in the probability of dropout of about 29 percentage points.

When the propensity score matching analysis is performed for rural communities, we find that building a new school or renovating an existing one, and building or improving a water facility, can have an

effect of reducing the probability of dropping out. A similar effect is found for the case in which in the community a social protection programme has been put in place and in this case the effect is also significant for the case of dropout in the period 2014/15.

The analyses presented highlight the importance to even more deeply study the challenges faced by the education sector, and of school dropout in particular, uncovering the pull and push factors behind the high dropout rates still present in the country. At the same time, the results presented add some further knowledge on the factors that lead to the outcomes described, providing some evidence to policy makers and development practitioners on some of the most worrying characteristics that are associated with school dropout and on some of the public and private investments that show an association with a decrease in the dropout rates in the country.

References

Abadie, A., and Imbens, G.W. 2012. Matching on the estimated propensity score. Harvard University and National Bureau of Economic Research. http://www.hks.harvard.edu/fs/aabadie/pscore.pdf.

APA-Maputo, 2018. "School dropout figure worries Mozambique". APA-Maputo, Mozambique, April 28, 2018. http://mobile.apanews.net/en/news/mozambique-alarmed-by-school-dropout-figure

De Walque, D. and Valente, C. 2016. Preventing excess female school drop out in Mozambique. Conditional transfers and the respective role of parent and child in schooling decisions. IGC Working paper.

DEEF. 2016. Pobreza e Bem-Estar em Moçambique: Quarta Avaliação Nacional. Maputo: Ministry of Economics and Finance/DEEF.

Fox, L., Santibáñez, L., Vy, N. and André, P. 2012. Education reform in Mozambique, lessons and challenges, World Bank, Washington DC.

GoM Ministry of Education. 2018. Education Management Information System. Maputo: Ministry of Education.

Ibraimo, M., and Salvucci, V. 2017. Os Determinantes da Pobreza em Moçambique, 2014/15. Inclusive Growth in Mozambique Working Paper.

Leuven, E., Sianesi, B. 2003. PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Available at: http://ideas.repec.org/c/boc/bocode/s432001.html. Version 4.0.11, 22 October 2014.

Mambo, F. 2017. Procura por Educação Primária em Moçambique. Inclusive Growth in Mozambique Working Paper.

Rosenbaum, P.R., Rubin, D.B. 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70(1), 41–55.

UNESCO, 2012.

UNICEF, 2010. Child poverty and disparities in Mozambique 2010, UNICEF Mozambique, Maputo.

UNICEF, 2014. Situation Analysis of Children in Mozambique 2014, UNICEF Mozambique, Maputo.

World Bank, 2005.

Appendix A

Table A.1. Determinants of dropout at any point in time, various subpopulations

VARIABLES	*(1) Dropout anytime*	*(2) Dropout anytime*	*(3) Dropout anytime*	*(4) Dropout anytime*	*(5) Dropout anytime*	*(6) Dropout anytime*	*(7) Dropout anytime*	*(8) Dropout anytime*
	Entire population of kids aged 6-17	*Boys*	*Girls*	*Kids aged 6-13*	*Kids aged 14-17*	*Kids in primary school (up to 5th grade)*	*Kids in primary school (up to 7th grade)*	*Kids in secondary school*
Sex	0.0106*	-	-	0.0138***	-0.0170	0.00195	0.00574	0.0229
	(0.00565)			(0.00504)	(0.0142)	(0.00627)	(0.00600)	(0.0150)
Age	0.0298***	0.0275***	0.0322***	0.0123***	0.0659***	0.0300***	0.0320***	0.0476***
	(0.00127)	(0.00178)	(0.00158)	(0.00153)	(0.00594)	(0.00156)	(0.00142)	(0.00906)
Birth order	-0.00451*	-0.00308	-0.00401	-0.00114	-0.00907*	-0.00453	-0.00575**	0.00628
	(0.00247)	(0.00369)	(0.00297)	(0.00302)	(0.00533)	(0.00297)	(0.00277)	(0.00599)
Orphan	0.00848	0.0214**	-0.00487	0.0135*	0.00423	0.0129	0.00985	0.00639
	(0.00738)	(0.00995)	(0.00950)	(0.00799)	(0.0161)	(0.00894)	(0.00802)	(0.0162)
Not in the hh	0.0387***	0.0341***	0.0409***	0.0270***	0.0393**	0.0298***	0.0321***	0.0599***
	(0.00814)	(0.0115)	(0.00935)	(0.00735)	(0.0180)	(0.00907)	(0.00881)	(0.0202)
Child works	0.0576***	0.0488***	0.0652***	0.0318***	0.113***	0.0456***	0.0511***	0.0679***
	(0.00517)	(0.00697)	(0.00658)	(0.00483)	(0.0116)	(0.00565)	(0.00543)	(0.0125)
Disability	0.0493***	0.0308	0.0611***	0.0313*	0.0853*	0.0530***	0.0487**	-0.0350
	(0.0188)	(0.0280)	(0.0235)	(0.0174)	(0.0471)	(0.0204)	(0.0199)	(0.0656)
Late	0.0265***	0.0113	0.0472***	0.0178***	0.124***	-0.00444	0.00282	-0.0385**
	(0.00793)	(0.0109)	(0.0107)	(0.00594)	(0.0250)	(0.00907)	(0.00887)	(0.0184)
Married	0.0592***	-	0.0676***	-0.00200	0.273***	0.0243***	0.0430***	0.239***
	(0.00715)		(0.00738)	(0.00549)	(0.0328)	(0.00748)	(0.00755)	(0.0304)
Pregnant	0.298***	-	0.249***	0.235***	0.445***	0.445***	0.292***	0.262***
	(0.0319)		(0.0297)	(0.0500)	(0.0595)	(0.0562)	(0.0445)	(0.0370)
Head age	-8.33e-05	0.000200	-0.000296	-6.43e-05	1.53e-05	-6.12e-05	-1.27e-05	-7.38e-05
	(0.000247)	(0.000333)	(0.000313)	(0.000208)	(0.000560)	(0.000282)	(0.000268)	(0.000555)
Head woman	-0.0324***	-0.0307***	-0.0317***	-0.0274***	-0.0260	-0.0295***	-0.0304***	-0.0263
	(0.00803)	(0.0115)	(0.00971)	(0.00741)	(0.0182)	(0.00915)	(0.00880)	(0.0165)
Head 5y educ	-0.0154**	-0.0102	-0.0199**	-0.0132**	-0.0260	-0.0121	-0.0145*	-0.0181
	(0.00711)	(0.00872)	(0.00824)	(0.00621)	(0.0167)	(0.00739)	(0.00742)	(0.0207)
Head 7y educ	-0.0280***	-0.0194	-0.0359***	-0.0196**	-0.0553**	-0.0253**	-0.0290***	-0.0183

	(0.00993)	(0.0135)	(0.0109)	(0.00888)	(0.0216)	(0.0111)	(0.0108)	(0.0227)
Head 10y educ	-0.0510***	-0.0689***	-0.0405***	-0.0362***	-0.0859***	-0.0436***	-0.0448***	-0.0427
	(0.0103)	(0.0159)	(0.0135)	(0.00830)	(0.0249)	(0.0117)	(0.0113)	(0.0260)
Head 12y educ	-0.0729***	-0.0746***	-0.0850***	-0.0415***	-0.131***	-0.0387**	-0.0615***	-0.0651**
	(0.0132)	(0.0222)	(0.0282)	(0.0128)	(0.0321)	(0.0181)	(0.0151)	(0.0272)
Head +12y educ	-0.0729***	-0.0723**	-0.0916***	-0.0253	-0.150***	-0.00504	-0.0393*	-0.0952***
	(0.0162)	(0.0340)	(0.0274)	(0.0168)	(0.0376)	(0.0253)	(0.0214)	(0.0264)
Head in agriculture	-0.00547	-0.0106	0.000264	-0.00575	-0.00783	-0.0107	-0.00670	0.0265*
	(0.00753)	(0.0100)	(0.00971)	(0.00662)	(0.0160)	(0.00775)	(0.00768)	(0.0157)
Poor	-0.00462	-0.00851	0.000580	-0.0108**	0.0150	-0.00662	-0.00572	0.0147
	(0.00580)	(0.00795)	(0.00671)	(0.00521)	(0.0131)	(0.00624)	(0.00593)	(0.0153)
Water source	-0.0187**	-0.0234**	-0.0149	-0.00499	-0.0548***	-0.0132	-0.0191**	-0.0104
	(0.00856)	(0.0111)	(0.0101)	(0.00674)	(0.0192)	(0.00915)	(0.00908)	(0.0235)
Sanitation	-0.0266***	-0.0338***	-0.0171	-0.0118	-0.0501**	-0.0261***	-0.0233**	-0.0128
	(0.00885)	(0.0116)	(0.0115)	(0.00770)	(0.0204)	(0.0101)	(0.00939)	(0.0166)
Roof	-0.0194**	-0.0136	-0.0265**	-0.0139*	-0.0282	-0.0207**	-0.0172*	-0.0275
	(0.00918)	(0.0120)	(0.0109)	(0.00816)	(0.0199)	(0.00992)	(0.00965)	(0.0253)
Electricity	-0.0316***	-0.0305**	-0.0323***	-0.0190**	-0.0514**	-0.0188*	-0.0262***	-0.0311
	(0.00978)	(0.0139)	(0.0124)	(0.00935)	(0.0232)	(0.0104)	(0.00989)	(0.0210)
Durable goods	-0.0125	-0.0180	-0.00794	-0.0130*	-0.0152	-0.00931	-0.00852	-0.0179
	(0.00780)	(0.0113)	(0.00964)	(0.00692)	(0.0182)	(0.00830)	(0.00823)	(0.0215)
Access to transport	0.00490	0.00553	0.00672	0.0105	-0.0137	0.00244	-0.00107	0.0263
	(0.0101)	(0.0126)	(0.0118)	(0.00829)	(0.0225)	(0.0105)	(0.0103)	(0.0345)
School	-0.0128	-0.0222	0.000354	-0.0164	0.0128	-0.0206*	-0.0120	-0.0113
	(0.0129)	(0.0155)	(0.0170)	(0.0100)	(0.0318)	(0.0119)	(0.0128)	(0.0457)
Flood	-0.00675	-0.0173	0.00302	0.00350	-0.0365	-0.00478	-0.00637	-0.0208
	(0.0105)	(0.0140)	(0.0118)	(0.00879)	(0.0243)	(0.0107)	(0.0106)	(0.0311)
Distance school 16-30	0.0172**	0.0214*	0.0111	0.00958	0.0326*	0.0127	0.0172**	0.00228
	(0.00770)	(0.0109)	(0.00915)	(0.00673)	(0.0170)	(0.00847)	(0.00810)	(0.0165)
Distance school 31-60	0.0351***	0.0481***	0.0141	0.0273***	0.0498**	0.0350***	0.0344***	0.00676
	(0.0105)	(0.0134)	(0.0113)	(0.00945)	(0.0230)	(0.0116)	(0.0110)	(0.0261)
Distance school 60+	0.0252*	0.0106	0.0385**	0.00709	0.0772**	0.0189	0.0161	0.217**
	(0.0133)	(0.0188)	(0.0157)	(0.0109)	(0.0339)	(0.0139)	(0.0142)	(0.107)
Temporal controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	43,102	21,672	21,430	29,757	13,345	29,788	36,612	4,374

Standard errors in parentheses

**** p<0.01, ** p<0.05, * p<0.1*

Note: There are about 18,000 children included in the analyses, not all of them were observed in all three survey quarters.

Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).

Table A.2. Determinants of dropout in the period 2014-15, various subpopulations

VARIABLES	*(1) Dropout 1415*	*(2) Dropout 1415*	*(3) Dropout 1415*	*(4) Dropout 1415*	*(5) Dropout 1415*	*(6) Dropout 1415*	*(7) Dropout 1415*	*(8) Dropout 1415*
	Entire population of kids aged 6-17	*Boys*	*Girls*	*Kids aged 6-13*	*Kids aged 14-17*	*Kids in primary school (up to 5th grade)*	*Kids in primary school (up to 7th grade)*	*Kids in secondary school*
Orphan	-0.00221	0.00681	-0.0136	-0.00470	0.00610	-0.00584	-0.00296	0.00516
	(0.0108)	(0.0147)	(0.0156)	(0.0118)	(0.0203)	(0.0119)	(0.0115)	(0.0245)
Not in the hh	0.0384**	0.0279	0.0476**	0.0488***	5.48e-05	0.0413**	0.0406**	0.0140
	(0.0155)	(0.0208)	(0.0208)	(0.0171)	(0.0324)	(0.0172)	(0.0159)	(0.0474)
Child works	0.0323***	0.0417***	0.0210**	0.0202**	0.0542***	0.0300***	0.0317***	0.0364
	(0.00743)	(0.00955)	(0.0102)	(0.00813)	(0.0128)	(0.00819)	(0.00766)	(0.0252)
Late	-0.0512***	-0.0561***	-0.0493***	-0.0612***	-0.0269*	-0.0520***	-0.0519***	-0.0432**
	(0.00824)	(0.0117)	(0.0105)	(0.00902)	(0.0157)	(0.00865)	(0.00876)	(0.0183)
Married	0.0254***	-	0.0277***	0.0147**	-0.0439***	0.0223***	0.0261***	0.00730
	(0.00594)		(0.00790)	(0.00649)	(0.00899)	(0.00656)	(0.00617)	(0.0550)
Pregnant	0.291***	-	0.291***	0.822***	0.258***	0.137**	0.270***	0.341***
	(0.0472)		(0.0473)	(0.136)	(0.0482)	(0.0605)	(0.0577)	(0.0919)
Head woman	0.176*	0.271	-0.0108	0.251**	-0.0570	0.212**	0.189**	0.0288
	(0.0899)	(0.179)	(0.0426)	(0.107)	(0.132)	(0.104)	(0.0957)	(0.0479)
Poor	0.00600	0.00629	0.00488	0.00218	0.0158	0.00553	0.00717	-0.0198
	(0.00677)	(0.00891)	(0.00876)	(0.00733)	(0.0128)	(0.00740)	(0.00700)	(0.0174)
Flood	0.0105	0.0160	0.00483	0.0126	0.00245	0.0128	0.0112	-0.00371
	(0.0138)	(0.0162)	(0.0149)	(0.0147)	(0.0192)	(0.0153)	(0.0145)	(0.0208)
Constant	0.0765***	0.0590	0.121***	0.0102	0.273***	0.0680**	0.0745***	0.104***
	(0.0264)	(0.0502)	(0.0187)	(0.0301)	(0.0429)	(0.0293)	(0.0279)	(0.0211)
Temporal controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic controls	No	No	No	No	No	No	No	No
Observations	44,520	22,414	22,115	30,811	13,718	32,848	39,809	4,443
Number of kids	18,056	9,057	9,037	12,584	5,515	13,527	16,153	1,666

Standard errors in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

Note: There are about 18,000 children included in the analyses, not all of them were observed in all three survey quarters.
Source: Authors' calculations using the 2014/15 Mozambican Household Budget Survey (IOF14).