Gender gap in current school enrolment: selection among "irregular" student?*

Michel Tenikue†
CEPS/INSTEAD, Luxembourg

September 2009

Abstract

Many developing countries face a pro-male gender gap in schooling, as boys are more likely to be enrolled at school than girls. This paper examines whether the current enrolment gap prevails at the same time among "regular" and "irregular" children. Regular children are children who complete primary education between the age of 12 and 15 years. Irregular children are the rest. We investigate the gender gap in schooling empirically using data provided by the 2001 Cameroon Household Survey. The empirical framework allows for a different gender effect among regular and irregular children. It also accounts for selection into the two groups. Results show no male-female difference among regular children. Among irregular children however, females are more likely to stay out of schools. We therefore suggest that, independently of the source of the gender gap, it seems to be at work mostly among irregular children.

*I began work on this paper while visiting the University of British Columbia. I am grateful to Siwan Anderson, Patrick François, Ashok Kotwal, Jean-Marie Baland for helpful comments. I would also like to thank Frederic Gaspart, Sylvie Lambert, Primila Krishnan, Vincenzo Verardi, Miguel Urquiola, Hanan Jacoby, David McKenzie and seminar participants at ECRU UCL.

†Email: michel.tenikue@ceps.lu

1 Introduction

It is widely recognized that, irrespective of the gender, investments in education have positive effects individual income and productivity. Unfortunately, in many developing countries, a pro-male gender gap is still observed in schooling (Orazem and King, 2008, World-Bank, 2008, Dar, Blunch, Kim, and Sasaki, 2002). Closing this gap is a priority for many governments[1] and is included in the Millennium Developments Goals.

The literature has suggested many sources of the gender gap in school enrolment: preference or parental discrimination (Kingdon, 2002), market incentives such as the male-female differences in the opportunity cost of time spent at school and the male-female difference in future earning prospects (Kingdon, 1998, Munshi and Rosenzweig, 2006), social norms about gender roles in familial relationships (Orazem and King, 2008, Rosenzweig and Schultz, 1982, Lahiri and Self, 2007). The driving forces of the gender gap cited generate a similar gender effect across different groups of children. Particularly a similar gender effect is expected among children of who started their educational process with a delay and those who did not, children who have repeated grades and those who have not. In this paper we investigate whether the gender gap in school enrolment is alike among on one hand, children who feature late enrolment or who have repeated grades and on the other hand, children with a regular course of study. In the sequel, we refer to the words "regular" and "irregular" to identify two groups of children. Regular children are those who complete primary education between the age of 12 and 15 years old. Irregular children are the rest, namely children who have not completed primary education by the age of 15. They have faced at least one of the following irregularities: delayed enrolment, repeating a grade, or interruption in the course of study.

We use data provided by the 2001 Cameroon Living Standards Measurement Survey (LSMS) and examine the gender gap in current school enrolment rates among children aged 12 to 19 years. We start at the age 12 for two reasons: (1) in general primary education is completed around the age of 12, (2) in our data, there is almost no male-female difference in school enrolment rates up to the age of 11[2]. Indeed, in our database 82% of children aged 6 to 11 years old are enrolled at school and this enrolment rate is split into 83% for boys and 81% for girls[3].

To refine the distinction between regular and irregular children, we take advantage of the fact that in Cameroon, a countrywide exam, called the Primary School Certificate (CEP), is organized at the end of primary school. This exam takes place the same days with the same set of questions for all candidates. Children undergo this exam when they are around 12 years old. A child is considered regular if he/she succeeds the exam by the age of 12 or before the age of 15 [4]. Using this criterion, we divide children aged 12 to 15 years into two groups. Regular

[1]The reduction of the gender gap in education is an important priority in the Poverty Reduction Strategy Papers of most countries.

[2]We document this in this paper. We typically suppose that, the current wave of children aged 6 to 11 properly represents the situation of children aged 12 or more when they were younger (stationarity assumption). We then show that there is no male female difference before the age of 11.

[3]Furthermore, the net enrolment rate in primary school is 78% and is split into 79% for boys and 77 % for girls (INS and Macro, 2005).

[4]We choose 15 years as upper limit to account for the fact that, around the age of 16 children attend another

children or children with CEP and irregular children or children with no CEP.

As a robustness check, we do a similar analysis with an alternative group of children. The reference exam is now the secondary school certificate (BEPC) organized at the end of the 10^{th} year of education. Children take this exam when they are around 16. Gender difference in school enrolment is investigated among children aged 16 to 19, taking into account the distinction between BEPC holders and those who do not.

Our empirical framework accounts explicitly for the fact that success completing primary education (success in CEP) is not a pure random realizations. It jointly estimates the probability of having the CEP and the probability of being currently enrolled in school. This approach has the advantage of controlling for the selection into groups of irregular and regular children. The main finding is that male-female differences in school enrolment is only observed among irregular children. This result suggests that female children are more penalized when they feature some irregularities in their course of study.

From the data, it is not possible to identify the exact economic mechanism driving our empirical results. However, given our static analytical framework, the results seem to highlight the role played by the schooling history of a given child. So, current schooling decision may depend on past schooling decisions and past schooling outcomes. For instance, if from the schooling history, parents realize that their daughter features some schooling irregularities (repeating grades), they may anticipate that, additional investment in her eduction would not be properly rewarded. While, in the case of a boy, despite the irregularities, they can still expect good returns. Such gender difference in perceived return to education can be driven by the fact, after the schooling period, females participate to housework (non income generating activities) while males barely do so.

The contribution of this paper is the light it sheds on the role played by the schooling history of a child on its probability to be currently enrolled at school. It suggests that, independently of the source of the gender difference, the pro-male gender gap in school enrolment seems to be important mainly among children who feature some schooling irregularities. Thus, reducing sources of schooling irregularities such as late enrolment and repeating grades should be added to the set of policies suggested by Glick (2008) to reduce gender schooling gap.

The paper is organized as follows. Section 2 explains the analytical framework and discusses the role played by schooling history rhythm. Section 3 describes the data, provides evidence on the absence of gender gap under the age of 11 and develops the empirical model. Section 4 provides the main estimation results. Section 5 discusses alternative mechanisms to interpret the results while section 5 gives some concluding comments.

exam which is considered separately in the paper. Further, more in general, in cameroon a child is not allowed to repeat 3 times the same grade in the same public school.

2 Analytical framework of gender gap in schooling

This section presents a model of school length. The household has two children: one son and one daughter. The life of a child is divided into two periods. The first is devoted entirely to education and the second is devoted to working. Parents decide on the amount of time their offspring spend at school. We denote the schooling time of child i by E_i and his/her working time T_i. The common retirement age R is defined by : $R = E_i + T_i$. The index i takes the value m for male and f for female.

We suppose that there is no within household competition over resources (Garg and Morduch, 1998) and that parents have perfect access to credit. Education is regarded as a risk-free investment (Kodde, 1986). Its cost, which is gender specific, is constant and denoted by c_i. The cost is composed of direct schooling costs (tuition fee, uniforms, books, transport...) and the opportunity cost of child's time devoted to education. Parents care about education because they are altruistic. Their schooling decisions reflect a tradeoff between schooling costs and future earnings (Baland and Robinson, 2000, Bedi and Marshall, 2002).

The utility of the parent depends on the wellbeing of his children and the cost of sending them at school. In a household with a boy and a girl, the utility is written as:

$$W_p = \alpha_m U_m(E_m, T_m, \lambda_m) + \alpha_f U_f(E_f, T_f, \lambda_f) - (c_m E_m + c_f E_f) \tag{1}$$

where U_i is the utility of child i. The parameters α_i represent the altruism coefficient of the parent towards the son and the daughter. The parameters λ_i summarize the schooling histories. They are known at the time of the current schooling decision and reflect irregularities faced (late enrolment, grades repeating, interruptions) by the child in the past. Children who started their educational process at the age of 6 and advance one grade every year have higher value of λ.

The utility of the child depends on λ, on the total amount of time spent at school and on the amount of time devoted to income generating activities or housework afterward. A boy spends his entire working time to income generating activities. His productivity is defined by $Q^m(E_m, \lambda_m)$. We assume that the marginal return to schooling is strictly positive and strictly decreasing ($Q^m_{E_m} > 0$, $Q^m_{E_m E_m} < 0$)[5]. We also suppose that return to regularity is positive and that return to education increases with regularity ($Q^m_{\lambda_m} > 0$, $Q^m_{E_m \lambda_m} > 0$). The utility of the boy is defined by:

$$U_m = Q^m(E_m, \lambda_m) T_m = Q^m(E_m, \lambda_m)(R - E_m) \tag{2}$$

Unlike a boy, a girl devotes part of her working time to housework. She chooses the amount of time, denoted t, she spends on household's activities such as child care[6], cooking and housekeeping[7]. The productivity of a girl is defined by $Q^f(E_f, \lambda_f)$ and her utility function is written

[5] This assumption is common in the literature, for example (Orazem and King, 2008, Baland and Robinson, 2000, Bommier and Lambert, 2000).

[6] Child care is considered broadly to include maternity leave.

[7] Hersch and Stratton (1997) provides a detail list of housework activities.

as:

$$U_f = Q^f(E_f, T_f, \lambda_f) = Q^f(E_f, \lambda_f)(R - E_f - t) + u(t) \tag{3}$$

were t is the amount of time spent on housework. We assume that u is positive and concave with u(0)=0 and $u'(0) = +\infty$. We also assume that the features of the productivity function of a girl (Q^f) are similar to those of a boy (Q^m). The choice of t is optimal and satisfies the following first order condition:

$$-Q^f(E_f, \lambda_f) + u'(t^*) = 0 \tag{4}$$

This equation means that, the marginal cost of spending time on housework is $Q^f(E_f, \lambda_f)$ and the marginal gain is u'(t) for the daughter.

Parents decide on the amount of time the son and the daughter spent at school to maximize their utility. Substituting for U_m and U_f in equation (1), the problem is written as:

$$\begin{aligned} \underset{E_m, E_f}{Max} \quad & \alpha_m\Big(Q^m(E_m, \lambda_m)(R - E_m)\Big) \\ & + \alpha_f\Big(Q^f(E_f, \lambda_f)(R - E_f - t*) + u(t*)\Big) \\ & - (c_m E_m + c_f E_f) \end{aligned} \tag{5}$$

The first-order conditions for time spent at school by the boy and the girl are given by:

$$\alpha_m \frac{\partial Q^m(E_m, \lambda_m)}{\partial E_m}(R - E_m) - \alpha_m Q^m(E_m, \lambda_m) - c_m = 0 \tag{6}$$

$$\alpha_f \frac{\partial Q^f(E_f, \lambda_f)}{\partial E_f}(R - E_f - t^*) - \alpha_f Q^f(E_f, \lambda_f) - c_f = 0 \tag{7}$$

These equations imply that, parents keep children at school until the marginal benefit equals the marginal cost. The gain is related to an increase in productivity induced by an additional year spent at school. The marginal cost involves both direct cost c_i and the opportunity cost since additional time spent at school reduces the amount of time spent on income generating activities.

This framework is general and can generate, under plausible assumptions, many often cited sources of systematic gender gap in schooling[8]. For instance, holding other factors to be the same for both sexes, if the return to education is higher for the boy, that is $Q^f_E(E, \lambda) < Q^m_E(E, \lambda)$, then there will be a pro-male gender gap in schooling. Similarly, if the cost of schooling is higher for female ($c_f > c_m$) all else equal, then the daughter will receive less schooling. Finally if for some cultural factors or social norms, the empathy of parents is higher for the son than for the

[8]See Alderman and King (1998), Glick (2008) and others.

daughter ($\alpha_m > \alpha_f$)[9], then all else equal, the daughter will receive less education.

We now impose no gender difference in empathy ($\alpha_m = \alpha_f$), in the cost of education ($c_f = c_m$) and in productivity ($Q^f = Q^m$). So, the quality of the schooling history (λ) and the choice of time devoted to housework remain the only sources of the gender gap in schooling.

Proposition 1 *If the son and the daughter have identical schooling histories, the daughter will receive less education if she participates to housework.*

Proof. We rewrite the equations (6) and (7) as:

$$\alpha Q_E(E_m, \lambda)(R - E_m) - \alpha Q(E_m, \lambda) - c = 0 \tag{8}$$

$$\alpha Q_E(E_f, \lambda)(R - E_f - t^*) - \alpha Q(E_f, \lambda) - c = 0 \tag{9}$$

If $t^* = 0$, the equalities in equations (8) and (9) are satisfied only if $E_f = E_m$. So the son and the daughter will spend the same amount of time at school. If $t^* > 0$, equalities (8) and (9) hold simultaneously if and only if

$$\alpha Q_E(E_m, \lambda)(R - E_m) - \alpha Q(E_m, \lambda) = c$$

and

$$\alpha Q_E(E_f, \lambda)(R - E_f) - \alpha Q(E_f, \lambda) = c + \alpha Q_E(E_f, \lambda)t*$$

We have that $c + \alpha Q_E(E_f, \lambda)t^* \geq$ c. So $E_f \leq E_m$ because the function $Q_E(x, \lambda)(R - x) - \alpha Q(x, \lambda)$ in decreasing in x.

■

Proposition 2 *If the elasticity of the return to education with respect to the schooling history (λ) is smaller than the elasticity of time devoted to housework with respect to λ, the a marginal improvement in schooling history will have bigger effect of the education of the daughter.*

Analytically, this proposition is written as : if $\frac{\lambda}{Q_\lambda} Q_{E\lambda} < \frac{\lambda}{t^*}(-\frac{\partial t^*}{\partial \lambda})$ then $\frac{\partial E_f}{\partial \lambda} > \frac{\partial E_m}{\partial \lambda}$.

The consequence of this proposition is that, if parents realize that their daughter features some irregularities in her schooling process, they would anticipate that, after her schooling period, she would devote a larger part of her working time to housework. The additional amount of time devoted to education would then not be properly rewarded. If on the contrary, they realize that their daughter is doing well at school, they will anticipate that, after her schooling period, she will spent a larger amount of time on income generating activities and will keep investing in her education. This mechanism is not at work for a son because he barely participates in housework.

Proof. The proof is in appendix. ■

[9]The preferences of parents are biased. They have a "taste for discrimination" Becker (1971).

3 Data, econometric model and model specification

3.1 Data

We use data of the 2001 Living Standard Measurement Survey in Cameroon. The sample had 11,000 households. The data provides information on the current enrolment status of children, on their highest grade completed and on the certificates they have. From the data, we can identify the stage of education of each child in 2001. The cameroonian education system consists of four stages: primary, secondary, post secondary and university. At the end of primary school (grade 6), students have to take the primary schooling certificate (CEP). The CEP is a countrywide exam organized independently by the Ministry of Basic Eduction. It takes place during the same days with the same set of questions for all candidates. Children take this exam when they are around 12 years old. A similar exam, called the secondary school certificate (BEPC), is organized at the end of secondary school. Children with a regular school progression rhythm attend the BEPC exam when they are around 16 years old.

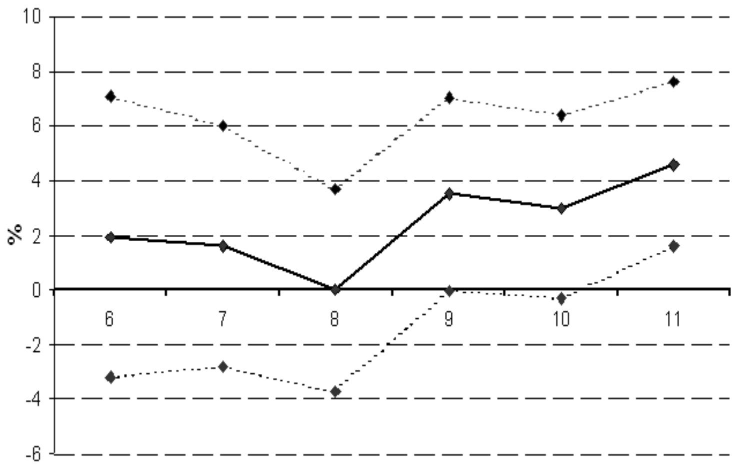


Figure 1: Difference between male and female current enrolment rates: 6-11 years old
The dashed lines indicate the the 95% confidence interval.

3.1.1 Choice of the sample

To get to the end of primary education, children have to be enrolled over 5 years and completed successfully the corresponding grades. Figure 1 shows the difference in male and female current enrolment rates by age and its 95% confidence interval. The lower bound of the confidence interval is negative up to the age of 10. Under a stationarity assumption, this figure suggests, in our data, the male-female difference in current enrolment rates is not significant before the age of 11.

We observed however significant difference in male and female enrolment rates from the age of 11. It exceeds 6% from the age of 12 onward (Table 1).

Even if male and female children are equally likely to be enrolled in early years of primary

Table 1: Odds ratio and current enrolment rates by gender and age: 6-19 years

	Current enrolment rates				
Age	All	Female	Male	Diff (M-F)	Odds (F/M)
6	0.60	0.59	0.61	1.9	0.92
7	0.78	0.77	0.78	1.6	0.91
8	0.85	0.85	0.85	0.001	1.00
9	0.88	0.86	0.90	3.5	0.72
10	0.89	0.87	0.90	3	0.73
11	0.92	0.90	0.95	4.6*	0.51
12	0.87	0.83	0.90	6.7***	0.55
13	0.86	0.83	0.88	4.9**	0.67
14	0.83	0.79	0.86	7.3***	0.60
15	0.75	0.72	0.78	5.8**	0.73
16	0.72	0.66	0.78	11.9***	0.55
17	0.60	0.54	0.66	11.9***	0.61
18	0.50	0.43	0.58	15.1***	0.54
19	0.46	0.38	0.54	16***	0.52

The different is difference between male and female current enrolment rates. The odds ratio (Odds (M/F)) is the female over male ratio of the odds of enrolment rates. *** is significant at 1% and ** is significant at 5%.

Table 2: Percentage of children with no dalay: 6-19 years

	% of children with no delay				
Age	All	Female	Male	Diff (M-F)	Odds(F/M)
6					
7	99	99	99	0.10	1.00
8	87	87	88	1.00	1.00
9	84	83	85	2.40	1.00
10	79	79	79	0.40	1.00
11	74	74	74	0.60	1.00
12	64	64	65	1.00	1.00
13	61	63	58	-4.6*	1.00
14	51	54	49	-4.40	1.00
15	45	46	45	-0.80	1.00
16	45	42	47	4.40	1.00
17	36	35	37	1.60	1.00
18	27	24	30	6**	1.01
19	22	19	25	6.4**	1.01

The difference (Diff (M-F)) is the different in male and female percentages. ** is significant at 5% and * is significant at 10%. The odds ratio (Odds (M/F)) is the female over male ratio of the odds of percentages.

education, they might evolve at different rhythms. Table 2 shows the percentage of children with some form of schooling irregularities by age. Irregularities are summarized here with schooling delays. A delay can be due to late enrolment, repeating grades or drop out. It is defined with respect to the theoretical number of completed grades given the age of the child. The theoretical number of completed grades of child aged 7 is 1 given that he/she is supposed to have started primary education at 6. The theoretical number of completed grades of elder children is defined accordingly[10]. This table suggests that, on average, male and female children are equally likely to register a delay, and if anything, females are better off.

The percentage of children with no schooling delay is about 80% up to the age of 10, and drops sharply to about 65% at the age of 12.

Table 3 shows current enrolment rates of children who have witness a delay and those who

[10]It is actually the number of grades s/he would have completed at the time of the survey, had s/he entered at age 6 and advanced one grade each year.

have not in the course of their schooling process. Enrolment rates are lower among children with a schooling delay. There is no gender difference in school enrolment among children with no schooling delay. However, among children with a schooling delay, a significant (at 5%) gender difference appears at the age of 10. From this age onward, female children in this group seem to be less likely to be enroled at school.

Table 3: Odds ratio and current enrolment rates by gender and age: 6-19 years with and without delay

	Children with a delay				Children with no delay			
	Cur. enrolment rate			Odds	Cur. enrolment rate			Odds
AGE	Fem	Male	Diff(M-F)	Ratio	Fem	Male	Diff(M-F)	Ratio
6								
7	-	-	-	-	77	78	1.70	1.000
8	1.1	1.2	0.1	3.000	97	96	-1	1.000
9	27	39	12*	1.012	98	98	0.00	1.000
10	44	60	17***	1.006	99	98	-0.8	1.000
11	68	81	13***	1.002	98	99	1.5*	1.000
12	60	78	18***	1.004	97	97	0.07	1.000
13	62	78	16***	1.003	96	95	-0.4	1.000
14	63	77	14***	1.003	93	96	2.70	1.000
15	51	64	13***	1.004	98	96	-2.6	1.000
16	45	63	18***	1.006	95	95	0.05	1.000
17	32	49	17***	1.011	94	95	1.04	1.000
18	26	43	17***	1.015	93	91	-2.3	1.000
19	26	41	14***	1.014	88	94	6.07	1.001

The difference (Diff (M-F)) is the different in male and female percentages. *** is significant at 1% and * is significant at 10%. The odds ratio is the female over male ratio of the odds of enrolment rates. A child has a delay if his/her number of completed year of education is lower that the theoretical number a grades he/she must have completed given his/her age.

The definition of schooling delay used in tables (2) and (3) relies on schools specific requirements to move from one grade to the next one. The standards are potentially different from one school to another. To use a common standard, we rely national standardized exams. We consider for instance that a child has completed primary education if s/he has succeeded the national exam organized at the end of primary education (CEP). The quality of schooling history is then summarized by success or failure to these country-wide exams.

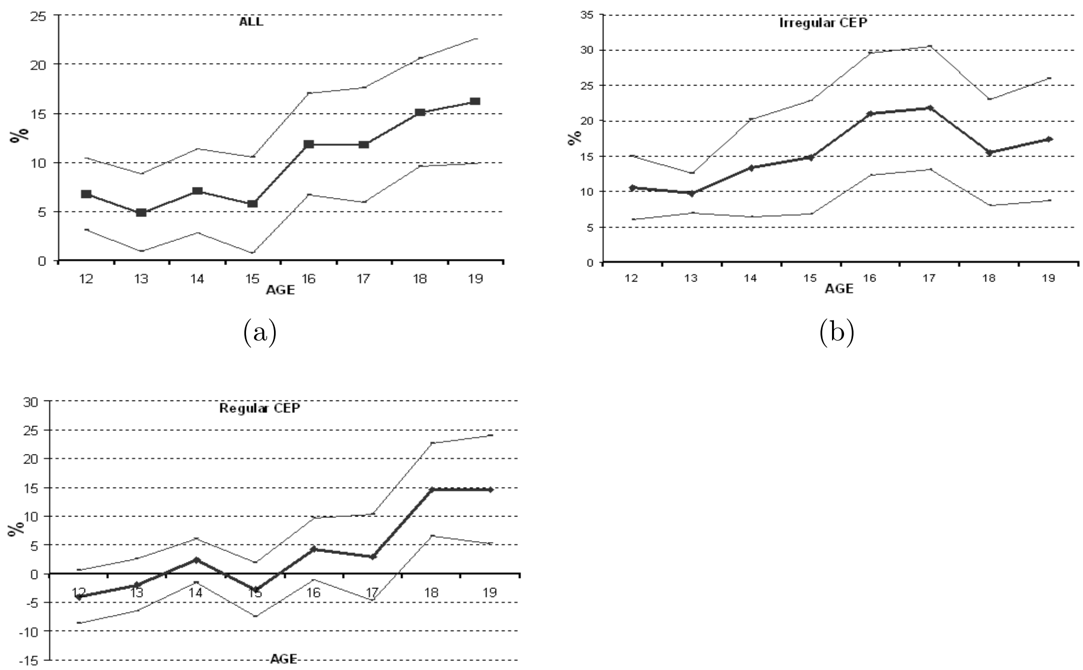
3.1.2 The sample

A child who starts primary school at the age of six and advances on grade every year will completed primary education by the age 12. Our sample therefore consists of children aged 12 to 19 years. It contains 9,585 children and is divided into two sub populations: children of 12 to 15 (53%) and children of 16 to 19 years old. Children aged 12 to 15 who have not completed primary education are called irregular children. They have got up to 3 extra years to succeed the CEP exam. Not completing primary education by this age indicates that have either repeated a grades, started schooling late[11] or drop out earlier. Similarly, children aged 16 to 19 years who have not succeeded the BEPC are called irregular children.

The total enrolment rate in our sample is 71% and is unequally split between males and

[11] Bommier and Lambert (2000) and Psacharopoulos and Patrinos (1997) study late enrolment or delayed enrolment.

females. Of the males, 76% are enrolled at school whereas 66% females are enrolled. Enrolment rate is 83% among children aged 12 to 15 years and is lower among children aged 16 to 19 years, about 57%. Figure 2) depicts the difference between male and female enrolment rate by age as well as the 95% confidence interval around it. Figure 2a is the difference for the sample as a whole. It indicates that male-female differences are significantly positive for all cohorts of age considered. As observed in African countries by Orazem and King (2008), the difference increases with age. Figure 2b plots the same difference for irregular children. It shows a similar positive difference that increases with age. However, the difference is not observed among children aged 12 to 17 who succeeded the CEP exam. Figure 2c shows that the difference is not significantly different from zero up to 17 years old and becomes positive only at the age of 18 and 19. The positive difference observed in this last part of the graph (age 18 and 19) is not very surprising. It corresponds to the situation 6 to 7 years after some children have got their CEP, long enough to give another signal on how far they can do at school, namely if they can succeed the BEPC or not.

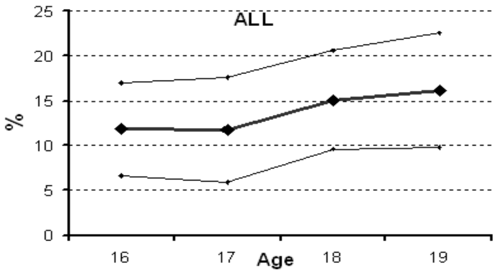


(a) (b) (c)

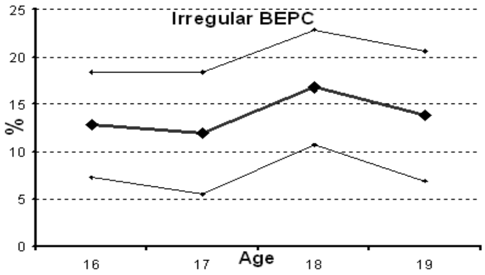
Figure 2: Male-Female differences in enrolment rates: children aged 12 to 19

Figure 3 shows the difference between male and female enrolment rates by age for children aged 16 to 19 as well as the 95% confidence interval around it. Children with and without BEPC are also considered separately. In this case as well, the gender effect is driven by the difference among irregular children.

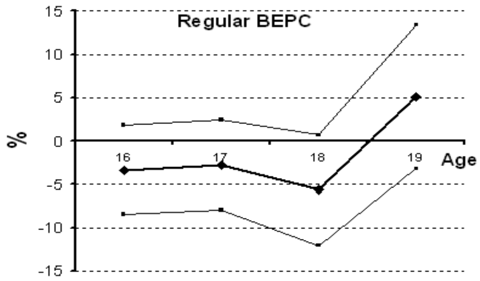
Table 4 presents enrolment rates by gender and age groups. It indicates that, on average, being a boy increases significantly the likelihood to be enrolled at school. If we divide the sample with respect to schooling regularity, we observe higher enrolment rates and no gender difference among regular children. In contrast, irregular children have lower enrolment rates and feature a



(a)



(b)



(c)

Figure 3: Male-Female differences in enrolment rates: children aged 16 to 19

The reference exam in this figure is the BEPC and the age group is 16-19.

significant gender differences in enrolment rates. So the gender gap observed seems to be driven mainly by the difference observed among irregular children.

Table 4: Enrolment rates by gender and age group

	Age 12-15			Age 16-19		
	All	Irregular	Regular	All	Irregular	Regular
Female	79.8	70.1	93.8	50.5	44.8	93.0
Male	85.8	81.9	92.2	64.5	59.3	92.2
Male-Female	6.0***	11.8***	-1.6	14.0***	14.5***	-0.8

Enrolment rate of the 12-15 is 83%. Enrolment rate of the 16-19 is 57%. ***indicates significant at 1%. Male-Female is the difference between Male and Female enrolment rates. All indicates all children of the age group considered. Irregular means have not succeeded the CEP exam for children aged 12 to 15 and the BEPC for children aged 16 to 19. Regular means have succeeded the CEP exam for children aged 12 to 15 and the BEPC for children aged 16 to 19. We do not exclude children between 16 and 19 years old who do not have the CEP in this table. If they were excluded, we would observed a change in the magnitude of the difference but but not change in the pattern.

Table 5 shows the allocation of the sample by gender. It suggests that pattern of the gender difference in enrolment rates is not driven by the gender structure of the sample. Actually, the results are reproduced in two sub-samples with different gender structures. In the first, children aged 12 to 15, males and females are equally represented and the proportion of irregular children is higher among males. In the second, children aged 16 to 19, they are slightly more females and the proportion of irregular children is higher among females.

Table 5: Allocation by gender

	Age 12-15		Age 16-19	
	% in sample	% Irregular	% sample	% Irregular
Female	50	57	51	84
Male	50	61	49	79
Total	100	41	100	18
Male-Female		4***		-5***

% in sample is the percentage male and female in the sample . % Irregular is the percentage of irregular children in the sub samples of males and females. Male-Female is the difference between Male and Female percentages. ***indicates significant at 1%.

3.2 Econometric model

We measure the quality of the schooling history with a dichotomous variable. In one group, there are children with irregular schooling histories. It the other, there are there children with a regular schooling history. The objective of this section is then to provide an econometric framework that allows to test whether the male-female difference in school enrolment exists within the two groups of children.

Regular and irregular children are in different stage of education. For instance, irregular children of 12 to 15 years have not completed primary education while regular children have. Regarding school enrolment, the former are to be enrolled in primary school while the latter are to be enrolled in secondary school[12]. Figure 4 illustrates the structure of the modeling strategy.

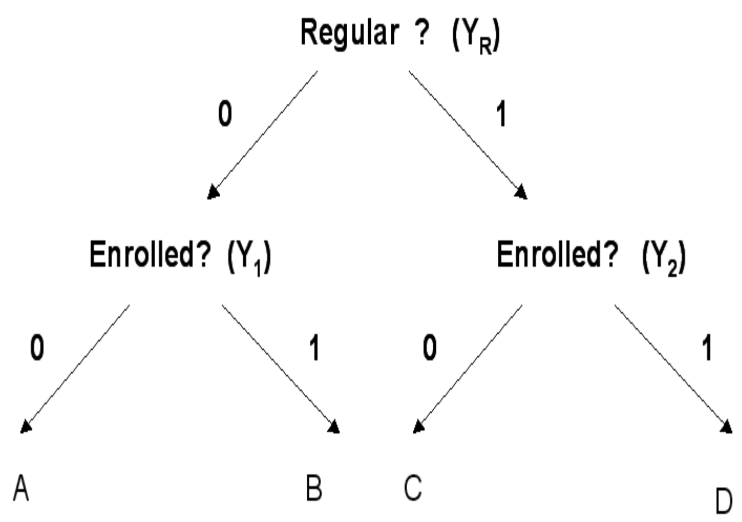


Figure 4: A model of school enrolment of the 12-15

A indicates irregular children who are not currently enrolled. B indicates irregular children who are currently enrolled. They are enrolled in primary education. C indicates regular children who are not currently enrolled. D indicates regular children who are currently enrolled. The are enrolled in secondary education.

If being regular was a pure random even, then the probability to be enrolled in school would be independent of the probability of being a regular child. The selection into the two regularity groups would be ignored and the econometric framework would consist of estimating binary

[12]Ten children in the data were in secondary school and did not have the CEP. They were excluded from the sample.

choice models. However, this possibility is unrealistic. Children from different families have different probabilities to attend or to succeed exams at given age (Cameron and Heckman, 2001, Mare, 1980, Dreze and Kingdon, 2001). The selection into the groups of children should be explicitly taken into account.

We use the structure provided by figure 4 to describe the econometric framework. Let Y_R be a binary variable indicating whether a child is regular or not. Let Y_1 and Y_2 be two binaries variables representing respectively school enrolment decisions of irregular and regular children. The variable Y_1 (respectively Y_2) is relevant only when the variable Y_R takes the value 0 (respectively 1). The probabilities to be enrolled are given by:

$$\begin{aligned} P(\text{B}) &= P(Y_1 = 1 \mid Y_R = 0) * P(Y_R = 0) \\ &= P(Y_1 = 1, Y_R = 0) \end{aligned}$$

$$\begin{aligned} P(\text{D}) &= P(Y_2 = 1 \mid Y_R) * P(Y_R = 1) \\ &= P(Y_2 = 1, Y_R = 1) \end{aligned}$$

Let us defined Y_j^* the index function associated to Y_j (j=R,1,2) by:

$$\begin{aligned} Y_{ji}^* &= Z'_{ji}\gamma_j + u_{ji} \\ &= \beta_j^0 + \beta_j^F * Female_i + X'_{ji}\beta_j + u_{ji} \end{aligned}$$

where i indicates individuals, $\gamma_j = (\beta_j^0, \beta_j^F, \beta_j)$ is the vector of parameters, $Z_j = (1, Female, X_j)$ is the set of regressors and u_{ji} is the unobserved error term. Female is a dummy variable that takes the value 1 for females. We assume that $u_{ji} \sim N(0, 1)$.

The outcomes being irregular and being enrolled are observed on the same child. They are related by a child specific unobserved heterogeneity (Lillard and Willis, 1994, Pal, 2004). Therefore, u_{R_i} and u_{1_i} are correlated. Similarly u_{R_i} and u_{2_i} are correlated. We denote the correlation coefficients by : $corr(u_{R_i}, u_{1_i}) = \rho_1$ and $corr(u_{R_i}, u_{2_i}) = \rho_2$.

This formulation is general and leaves open the possibility that different sets of regressors determine enrolment of irregular children and regular ones. It also leaves open the possibility that, a given regressor affects differently the likelihood to be enrolled of a child among the irregular group and a child among the regular group. We indeed expect the gender effect to be different for irregular and regular children. We will test the hypothesis that β_1^F is significant and negative while β_2^F is not significantly different from zero. The assumption that the gender gap in current school enrolment is independent of the schooling history corresponds to having both β_1^F and β_2^F significantly negative.

The predicted probability of the observed outcomes for any observation i is given by $\Phi_2(\mu_{ji}, \Omega_j)$ where Φ_2 is a bivariate standard normal cumulative distribution function with arguments $\mu_{ji} = (\kappa_{ji} Z'_{ji}\gamma_j, \kappa_{R_i} Z'_{R_i}\gamma_R)$ and Ω_j. The index j takes the values primary or second. The symbol κ_j denotes a "sign" variable and is defined by $\kappa_j = 2 * Y_j - 1$ for each observation. The 2 by 2 matrices

Ω_j have constituent elements Ω_{jtk} defined by $\Omega_{j11} = \Omega_{j22} = 1$ and $\Omega_{j12} = \Omega_{j21} = \kappa_j\kappa_1\rho_j$.

The log-likelihood function of observations is:

$$L = \Sigma_i^n 1_{iA}log(P_{iA}) + 1_{iB}log(P_{iB}) + 1_{iC}log(P_{iC}) + 1_{iD}log(P_{iD}) \tag{10}$$

where $1_{ix} = 1\{x\}$ is an indicator function with x =A,B,C,D.

We do not have a reasonable variable that would affect the probability to be irregular but that would not affect the probability to be enrolled at school. The identification of this model relies therefore on the nonlinearities of the functional form used[13] (Wooldridge, 2002).

If we assume that there is no child specific unobserved heterogeneity, we would have that $\rho_1 = \rho_2 = 0$. It implies that u_R, u_1 and u_2 are independent. The parameters γ_R, γ_1 and γ_2 could be estimated with three simple probit models.

3.3 Model specification

Section 2 indicates that, the probability to be enrolled in school is affected not only by the gender but also by other factors. To separate out the effects of these factors, it is necessary to specify a model with a suitable set of explanatory variables.

Participation to labor market: Child labor is seen as an opportunity cost of child's time devoted to schooling. As such child labor is endogenous to schooling. It is therefore inappropriate to use the indicator variable of child labor participation to explain school participation. However, demand factors of child labor can be used as instruments. The first set of instruments are related to demand of child labor from the household. In general the activities in which the household engages in, shape the demand for child labor within the household. As a result, we use the covariates self employed, agriculture and trader. These variables indicate respectively whether the head of the household or his spouse is self-employed, works in agriculture or runs a business.

Basu (1998) suggests that child labor and adult labor are substitutes. Thus the extend of local job opportunities for children can be provided by the local adult participation rate in the labor market. Adults local participation rates in labor market[14] and in the informal sector are used as regressor in the schooling equations. These participation rates also reflect demand factors of child labor originating from outside the household[15] and the extent of future job opportunities (Rosenzweig and Schultz, 1982). In addition, we expect that child labor occurs mainly in informal sector.

Supply side variables: One needs to control for school supply-side factors on enrolment decisions (Dreze and Kingdon, 2001, Bedi and Marshall, 2002). They reflect school quality, school proximity and cost (Behrman, Ross, and Sabot, 2008). Public primary school is free in Cameroon and tuition fees in public secondary school are identical across the country. Thus the relevant discriminating part of the cost of education are related to transportation cost and

[13]This is similar to what is commonly done with probit selection equation of nonresponse that lead to missing data.

[14]Pal (2004) uses this variable for a study on school attainment in Peru.

[15]Fafchamps and Wahba (2006) show that children living in or near cities are more likely to work.

Table 6: Descriptive statistics of the variables

	AGE 12-19		AGE 12-15		AGE 16-19	
	mean	sd	mean	sd	mean	sd
Age	15.34	2.28	13.46	1.13	17.44	1.09
Female	0.50		0.50		0.51	
Head has primary level	0.35		0.35		0.35	
Head has secondary level	0.33		0.32		0.35	
Head has university level	0.07		0.07		0.08	
Head sex (Male=1)	0.78		0.78		0.77	
Age of head	46.15	13.18	47.23	12.56	44.95	13.75
Muslim	0.39		0.38		0.41	
Christian	0.69		0.68		0.71	
Estimated expenditure per capita	3.14	3.13	2.91	3.09	3.39	3.15
Rural	0.28		0.31		0.25	
Head or spouse non wage worker	0.67		0.69		0.65	
Head or spouse (non wage) in agri	0.40		0.43		0.36	
Head or spouse (non wage) in trade	0.21		0.21		0.22	
Son/daughter of head	0.65		0.72		0.57	
Distance to private primary school	2.28	4.72	2.46	4.92	2.08	4.48
Distance to public primary school	1.06	1.88	1.13	2.08	0.99	1.61
Distance to private secondary school	4.07	6.06	4.37	6.27	3.74	5.80
Distance to public secondary school	3.64	4.98	3.87	5.17	3.38	4.74
Local participation to labor market	0.71	0.17	0.72	0.17	0.70	0.17
Local participation rate to informal sector	0.43	0.21	0.45	0.22	0.42	0.21
# of 0-5 years	1.07	1.26	1.13	1.28	1.00	1.23
# of male of 6-11 years	0.73	0.96	0.81	1.01	0.63	0.89
# of female of 6-11 years	0.69	0.94	0.77	0.98	0.60	0.90
# of male of 12-15 years	0.71	0.82	0.93	0.84	0.48	0.72
# of female of 12-15 years	0.72	0.84	0.93	0.86	0.50	0.75
# of male of 16-19 years	0.66	0.85	0.45	0.74	0.90	0.91
# of female of 16-19 years	0.62	0.76	0.42	0.67	0.85	0.79
# of male of 20-35 years	0.63	0.96	0.54	0.92	0.73	1.00
# of female of 20-35 years	0.70	0.87	0.73	0.86	0.67	0.88
# of over 35 years	1.56	1.03	1.63	1.00	1.49	1.06
N	9585		5067		4518	

The mean of a dummy variable is proportion.

other schooling direct costs (uniform, books, PTA fees,...). The data provides only the distance to the nearest primary or secondary school but not on other school quality characteristics. If we assume that transportation cost are related to distance, we can proxy transportation cost by the distance to the nearest school. We consider separately distance to a public and a private schools to reflect differences in these two types of schools. We further distinguish distance to a primary and a secondary school to reflect differences in the supply of the two levels of education. In our sample, the average distance to a public primary school for a given child is 1 kilometer. S/He needs to walk on average one more kilometer to reach the nearest private primary school. In general secondary schools are located farther than primary ones. A child needs to walk on average 2 to 3 additional kilometers to reach a secondary school.

The residential location (urban or rural) of a child in a developing country may reflect the quality of the local public services to which he has access. Well trained teachers prefer working in cities than in villages, distorting therefore the quality of services offered by schools.

Household resource constraints: In a country with imperfect credit market, household resource constraints play an important role on school investments (Baland and Robinson, 2000, Basu, 1998). We expect children from credit constrained households, actually poor households, to have a lower probability to attend school. The data does not provide information on credit constraints. They however contain household expenditure which is considered to be an indicator for household standard of living. But household expenditures are potentially related to schooling decisions in a household model. Consequently, we use predicted values[16] instead of actual values, of household expenditure per capita.

Household preferences and structure: Section 2 noted the importance of parental preferences on schooling decisions. These preferences are difficult to quantify. The education level of parents is used to capture their attitude towards schooling decisions.

Finally, we control for the structure of the household in terms of the number household member per age groups(Psacharopoulos and Patrinos, 1997). We consider separately the number of females and males in the schooling age groups. This allows to capture differential effect related to their presence in the household (Garg and Morduch, 1998). Further the number of adults in the households controls for the number of economically active family members (Manacorda, 2006).

Summary statistics of selected regressors classified by age groups are shown by Table 11 in appendix.

[16]The identifying instruments were household head characteristics, household structure, durable goods owned, geographical dummies, usable land owned. The adjusted R-squared of the OLS regression is 0.3.

4 Results

The model described by equation 10 is estimated on children aged 12-15[17]. Estimated coefficients are shown in table 7. The coefficients ρ_{21} and ρ_{31} are significantly different from zero. This suggest that the child specific heterogeneity is significant. We then conclude that ignoring the correlation structure of the data would give rise to inconsistent estimates.

The coefficient on gender is significantly negative for irregular children. This means that an irregular female aged 12 to 15 is less likelihood to be enrolled at school. In contrast, the coefficient on gender is positive and non significant for regular children. So, the effect of gender on the likelihood to be enrolled in school is not significantly different from zero among regular children aged 12 to 15. If any effect, it would be to increase the likelihood of female to be enrolled. The coefficient on gender in the column labeled regular in table 7 is not significant. It suggests that the gender plays no role in the selection into the groups of regular and irregular children.

As a robustness check, we estimate the model on children aged 16 to 19. Estimated coefficients are presented in table 9 in appendix. Here as well, the gender effect on school enrolment is significant only for irregular children.

Table 7: Estimated coefficients of the joint model on children aged 12 to 15.
Regularity if based on CEP exam.

	Regular	**Enroll irregular**	**Enroll regular**
Age	0.368	-0.198	0.178
	(19.79)***	(4.15)***	(2.09)**
Female	0.090	-0.411	0.128
	(1.47)	(4.74)***	(1.32)
Head has primary level	0.457	0.717	0.281
	(7.15)***	(9.98)***	(2.02)**
Head has secondary level	0.868	1.090	0.732
	(11.51)***	(10.24)***	(4.75)***
Head has university level	1.052	1.140	0.873
	(9.19)***	(5.18)***	(3.90)***
Head sex (Male=1)	-0.232	-0.449	-0.047
	(4.33)***	(5.61)***	(0.45)
Age of head	0.012	0.010	0.013
	(5.20)***	(3.46)***	(3.75)***
Muslim	0.004	-0.146	0.070
	(0.09)	(1.80)*	(0.90)
Christian	0.375	0.833	0.376

Continued on Next Page…

[17]The results of the model estimated on children age 16 to 19 is in Table 9 in appendix. We also estimated the model on the pool sample of all children. We keep the definition of irregular children in the process and consider enrolment more generally. The results (not shown here) on gender effect were qualitatively similar to those presented in this paper.

Table 7 – Continued

	Regular	**Enroll irregular**	**Enroll regular**
	(6.57)***	(10.73)***	(3.96)***
Estimated expenditure per head	0.070	0.053	0.079
	(6.36)***	(2.74)***	(4.56)***
Rural	-0.040	-0.033	0.047
	(0.56)	(0.36)	(0.39)
Head or spouse non wage worker	-0.057	0.001	-0.245
	(0.84)	(0.01)	(2.08)**
Head or spouse (non wage) in agri	-0.108	-0.067	0.092
	(1.50)	(0.71)	(0.70)
Head or spouse (non wage) in trade	-0.106	-0.080	-0.031
	(1.60)	(0.92)	(0.30)
Son/daughter of head	0.064	0.381	0.343
	(1.37)	(5.81)***	(3.29)***
Distance to private primary school	-0.037	-0.011	
	(5.31)***	(1.87)*	
Distance to public primary school	-0.017	-0.021	
	(0.96)	(1.96)**	
Distance to private secondary school	-0.005		0.003
	(0.88)		(0.26)
Distance to public secondary school	-0.034		-0.080
	(4.43)***		(5.54)***
Local participation to labor market	-0.168	0.074	-0.711
	(0.96)	(0.31)	(2.24)**
Local participation rate to informal sector	-0.044	-0.054	-0.053
	(0.33)	(0.36)	(0.24)
# of 0-5 years	-0.035	0.003	-0.000
	(1.75)*	(0.13)	(0.00)
# of male of 6-11 years	0.016	0.018	-0.032
	(0.70)	(0.67)	(0.80)
# of female of 6-11 years	0.039	0.004	0.067
	(1.66)*	(0.15)	(1.67)*
# of male of 12-15 years	0.066	0.063	0.063
	(2.05)**	(1.44)	(1.20)
# of female of 12-15 years	0.043	-0.004	0.067
	(1.32)	(0.08)	(1.30)
# of male of 16-19 years	0.033	0.012	0.075
	(1.13)	(0.28)	(1.55)
# of female of 16-19 years	0.093	0.044	0.065
	(2.92)***	(0.94)	(1.25)
# of male of 20-35 years	-0.006	0.051	-0.005
	(0.27)	(1.39)	(0.15)
# of female of 20-35 years	0.035	0.029	-0.024
	(1.29)	(0.79)	(0.50)
# of over 35 years	-0.022	0.100	-0.037
	(0.79)	(2.73)***	(0.82)

Continued on Next Page...

Table 7 – Continued

	Regular	**Enroll irregular**	**Enroll regular**
Constant	-6.345	2.205	-3.261
	(19.74)***	(3.43)***	(2.00)**
Rho1		0.546	
		(2.28)**	
Rho2		0.860	
		(6.38)***	
Observations		5067	
Test: $\beta^F_{primary} = \beta^F_{second}$		$\chi^2(1) = 18$	p=0.000

The dependent variables Regular, Enroll irregular and Enroll regular are all dummy variables. Regular takes value 1 if the child is considered regular. Enroll irregular takes value 1 if the Irregular child is enrolled in school. Enroll regular takes value 1 if the regular child is enrolled in school. Absolute value of z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

5 Discussion

Section 2 suggests one plausible mechanism to explain why the gender gap on current enrolment is observed only among irregular children. If from the schooling history parents realize that their daughter features some irregularities in her school progression rhythm, they would anticipate that, after her schooling period, she would devote a larger part of her working time to housework. Further investment in her education would therefore not be properly rewarded. Parents would then have lower incentive to keep investing in her education. Table 8[18] provides indirect evidence of this mechanism. It show the average number of hours devoted to income generating activities over month. It suggests that, on average, male and female children who were regular at school indeed spent the same amount of time in income generating activities after their schooling period. Female who were irregular spent on average a lower amount of time on income generating activities. However, figures in this table are not exempt of bias due to determinants of occupational choice (Banerjee and Newman, 1993). We can therefore think of two other possible mechanisms.

Table 8: # of hours devoted to income generating activities over a month

		20 - 24 years	
		Female	Male
CEP	Weak	35	46
	Able	39	48
	Regular-Irregular	4*	2
BEPC	Weak	36	47
	Able	42	45
	Regular-Irregular	6*	2

First, on might think that parents under invest in female's education from the age of 6 onward

18

and that what we observe at the age of 12 is simply the continuation of what started many years earlier. This scenario correspond to having female education costs higher than male education costs. In this case, we would then have a pro male gender gap in school enrolment at every age. But tables 1 and 3 suggest that, under stationarity assumption, there is not gender gap in current enrolment rate before the age of 10. Furthermore, when the gender difference appears, it is observed only among children featuring some schooling irregularities. By the same token, the possibility that parents invest more in female education in early age is not very plausible.

Second, the (opportunity) cost of schooling might be the same for male and female around the age of 6 but becomes higher for female around the age of 12. If this were the case, all female children will face the same cost around the age of 12 and they would all be less likely to enrolled at school. We would then observe a gender gender gab both among irregular and regular children. However, we have observed a gender gap in current enrolment rates only among irregular children(see figure 2). It thus suggests that, difference in cost is not driving force the our results.

We have assumed in the empirical model (section 3.2) that there is a child specific heterogeneity that affect both schooling regularity and enrolment. If this hypothesis is relaxed, then u_1 and u_R are independent. The error terms u_R and u_2 are independent as well. Estimating model 10 is reduced to estimating two independent probit models. Table 10 in appendix presents the corresponding estimation results. They show some differences compared to the initial results reported on table 7. For instance, on the role played by the education of household head or the indicator of family wealth[19]. However the message on gender gap is not altered.

One limitation of our data is that they do not provide the exact year at which children succeeded exams. It is not possible to know when a child, say of 15 years old, succeeded the CEP exam. Moreover, we do not know how an irregular child, say of 13 years old, would fare the next year. Thus with our definition of regularity, a child aged 15 who succeeded the CEP when he was 14 years old is classified as regular while a child aged 13 who did not have the CEP is classified as irregular. We check that our result is robust to the definition of regularity by estimating the model only on children aged 14 and only on children aged 17. The results, not shown, remain similar.

6 Conclusion

The paper examines the gender difference in the likelihood to be enrolled at school. Current literature suggests that in many African countries, male children are more likely to be enrolled at school than females. It treats identically children who evolve differently at school. We introduce a new approach that consists of allowing a different gender effect on the probability to be enrolled in school for children with different schooling histories. We study current enrolment among regular and irregular children. Regular children are those who complete primary education between the age 12 and 15.

[19]They have minor effect on the probability to enrolled in secondary school (Mare, 1980).

Our econometric framework accounts explicitly for selection of children into the groups of regular and irregular children. It jointly estimates the probability of being a Regular child and being currently enrolled. Results of the estimated model show no male-female difference in the likelihood to be enrolled in school among regular children. However, among irregular children, females are less likely to be enrolled in school than males.

While enrolment has become less problematic for very young children, dropping out of school remain important for children aged 12 or more, especially female. Our results suggest that one way to reduce the gender gap is to give more schooling incentives to children in order to reduce grade repetition. This could be implemented by giving bursaries to children with a regular rhythm of school progression or by introducing a performance-based conditional cash transfer programme.

References

Alderman, H., and E. M. King (1998): "Gender differences in parental investment in education," *Structural Change and Economic Dynamics*, 9(4), 453–468.

Baland, J.-M., and J. A. Robinson (2000): "Is Child Labor Inefficient?," *Journal of Political Economy*, 108(4).

Banerjee, A. V., and A. F. Newman (1993): "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101(2), 274–98.

Basu, K. & Van, P. H. (1998): "The Economics of child labor," *The American Economic Review*, 88 (3), 412–427.

Becker, G. S. (1971): *The Economics of Discrimination*. The University of Chicago Press, Chicago.

Bedi, A. S., and J. H. Marshall (2002): "Primary school attendance in Honduras," *Journal of Development Economics*, 69(1), 129–153.

Behrman, J. R., D. Ross, and R. Sabot (2008): "Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural Pakistan," *Journal of Development Economics*, 85(1-2), 94–104.

Bommier, A., and S. Lambert (2000): "Education Demand and Age at School Enrollment in Tanzania," *The Journal of Human Resources*, 35(1), 177–203.

Cameron, S. V., and J. J. Heckman (1998): "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," *Journal of Political Economy*, 106(2), 262–333.

——— (2001): "The Dynamics of Educational Attainment for Blacks, Hispanics, and Whites," *Journal of Political Economy*, 109.

Dar, A., N.-H. Blunch, B. Kim, and M. Sasaki (2002): "Participation of Children in Schooling and Labor Activities: A Review of Empirical Studies," Social protection discusion paper 0221, World Bank.

Dreze, J., and G. G. Kingdon (2001): "School Participation in Rural India," *Review of Development Economics*, 5(1), 1–24.

Fafchamps, M., and J. Wahba (2006): "Child labour, urban proximity, and household composition," *Journal of Development Economics*, 79 (2), 374–397.

Garg, A., and J. Morduch (1998): "Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana," *Journal of Population Economics*, 11, 471–493.

Glick, P. (2008): "What Policies Will Reduce Gender Schooling Gaps in Developing Countries: Evidence and Interpretation," *World Development*, 36(9), 1623–1646.

Hersch, J., and L. S. Stratton (1997): "Housework, Fixed Effects, and Wages of Married Workers," *The Journal of Human Resources,*, 32(2), 285–307.

INS, and O. Macro (2005): "Enqute Démographique et de santé, Cameroun 2004," Rapport, Institut National de la Statistique.

Kingdon, G. G. (1998): "Does the Labour Market Explain Lower Female Schooling in India?," *The Journal of Development Studies*, 35(12), 39–65.

——— (2002): "The gender gap in educational attainment in India: how much can be explained?," *The Journal of Development Studies*, 39(2), 25–53.

Kodde, D. A. (1986): "Uncertainty and the Demand for Education," *The Review of Economics and Statistics*, 68(3), 460–67.

Lahiri, S., and S. Self (2007): "Gender Bias in Education: the Role of Inter-household Externality, Dowry and other Social Institutions," *Review of Development Economics*, 11(4), 591–606.

Lillard, L. A., and R. J. Willis (1994): "Intergenerational Educational Mobility: Effects and State in Malaysia," *Journal of Human Resource*, 29(4), 1126–1166.

Lina, A., A. Simon, and K. John (2006): "Why Do Girls in Rural China Have Lower School Enrollment?," *World Development*, 34(9), 1639–1653.

Manacorda, M. (2006): "Child Labor and the Labor Supply of Other Household Members: Evidence from 1920 America," *American Economic Review*, 96(5), 1788–1801.

Mare, R. D. (1980): "Social background ans school continuation decisions," *Journal of the American Statistical Association*, 75(370), 295–305.

Munshi, K., and M. Rosenzweig (2006): "Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy," *American Economic Review*, 96(4), 1225–1252.

Orazem, P., and E. M. King (2008): "Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy," *in T. P. Schultz and John Strauss, Handbook of Development Economics*, vol 4, Elsevier Science, Amsterdam.

Pal, S. (2004): "Child Schooling in Peru: Evidence From A Sequential Analysis of School Progression," *Journal of Population Economics*, 17(4), 657–680.

Psacharopoulos, G., and H. A. Patrinos (1997): "Family size, schooling and child labor in Peru - An empirical analysis," *Journal of Population Economics*, 10(4), 387–405.

ROSENZWEIG, M. R., AND T. P. SCHULTZ (1982): "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India," *American Economic Review*, 72(4), 803–15.

WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, Massachusetts.

WORLD-BANK (2008): "World Development indicators," Wdi online, The World Bank Group.

7 Appendix

Proof of proposition 2.

Proof.

From equations (6) and (7), we have the following static comparative equation for the son and the daughter.

$$\frac{\partial E_m}{\lambda} = -\frac{Q_{E\lambda}(R - E-) - Q_\lambda}{Q_{EE}(R - E - t^*) - Q_E(1 - t^*_E) - Q_E} \tag{11}$$

$$\frac{\partial E_f}{\lambda} = -\frac{Q_{E\lambda}(R - E - t^*) - Q_E t^*_\lambda - Q_\lambda}{Q_{EE}(R - E - t^*) - Q_E(1 - t^*_E) - Q_E} \tag{12}$$

We suppose that Q is sufficiently concave so that

$$Q_{EE}(R - E - t^*) - Q_E(1 - t^*_E) - Q_E < 0$$

. Given that $Q_E t^*_E > 0$ the denominator of equation (11) is smaller that the denominator of equation (12). A sufficient condition to have $\frac{\partial E_m}{\lambda} < \frac{\partial E_f}{\lambda}$ is that

$$Q_{E\lambda} t^* - Q_E t^*_\lambda > 0$$

The is equivalent to $\frac{Q_{E\lambda}}{Q_E} < -\frac{t^*_\lambda}{t^*}$ or $Q_{E\lambda}\frac{\lambda}{Q_E} < t^*_\lambda \frac{\lambda}{t^*}$.

■

Table 9: Estimated coefficients of the joint model on children aged 16 to 19.
Regularity if based on BEPC exam.

	Regular?	**Enroll Irregular**	**Enroll Regular**
Age	0.299	-0.268	-0.200
	(12.56)***	(7.67)***	(0.86)
Female	-0.375	-0.498	-0.032
	(4.91)***	(7.25)***	(0.11)
Head has primary level	0.246	0.272	-0.470
	(2.68)***	(4.36)***	(1.07)
Head has secondary level	0.642	0.755	-0.001
	(6.45)***	(10.05)***	(0.00)
Head has university level	0.860	1.007	0.030
	(6.62)***	(7.76)***	(0.05)
Head sex (Male=1)	-0.235	-0.291	-0.152
	(3.80)***	(5.27)***	(0.77)
Estimated expenditure per head	0.110	0.071	0.062
	(9.08)***	(5.15)***	(1.07)
Age of head	0.009	0.006	0.008
	(3.12)***	(2.42)**	(0.80)
Rural	-0.402	0.053	-0.706
	(3.67)***	(0.69)	(2.22)**
Head or spouse non wage worker	-0.177	-0.170	-0.359
	(2.22)**	(2.30)**	(1.68)*
Head or spouse (non wage) in agri	-0.006	0.112	0.420
	(0.07)	(1.47)	(1.42)
Head or spouse (non wage) in trade	0.090	0.073	0.079
	(1.10)	(1.04)	(0.35)
Son/daughter of head	0.372	0.535	0.618
	(6.32)***	(10.09)***	(3.26)***
Distance to private primary school	-0.015	-0.005	
	(1.07)	(0.78)	
Distance to public primary school	-0.064	-0.014	
	(1.96)**	(0.80)	
Distance to private secondary school	-0.015	0.002	-0.015
	(1.79)*	(0.32)	(0.64)
Distance to public secondary school	-0.016	-0.046	-0.034
	(1.20)	(6.51)***	(0.84)
Local participation to labor market	-0.232	-0.128	0.920
	(1.10)	(0.68)	(1.37)
Local participation rate to informal sector	0.039	-0.264	-0.847
	(0.21)	(1.97)**	(1.43)
# of 0-5 years	-0.038	-0.077	-0.098
	(1.39)	(3.57)***	(1.23)
# of male of 6-11 years	0.008	-0.008	0.033
	(0.24)	(0.31)	(0.32)

Continued on Next Page…

Table 9 – Continued

	Regular?	**Enroll Irregular**	**Enroll Regular**
# of female of 6-11 years	0.024	0.060	0.103
	(0.77)	(2.25)**	(0.99)
# of male of 12-15 years	0.040	0.101	-0.153
	(1.05)	(3.11)***	(1.31)
# of female of 12-15 years	0.101	0.047	0.091
	(2.77)***	(1.44)	(0.75)
# of male of 16-19 years	-0.004	0.102	-0.006
	(0.11)	(3.05)***	(0.06)
# of female of 16-19 years	0.143	0.243	0.067
	(3.56)***	(6.26)***	(0.48)
# of male of 20-35 years	0.012	-0.058	-0.100
	(0.48)	(2.35)**	(1.54)
# of female of 20-35 years	0.057	0.078	0.216
	(1.80)*	(2.77)***	(2.14)**
# of over 35 years	-0.000	0.065	0.144
	(0.01)	(2.17)**	(1.24)
Constant	-7.127	4.097	3.426
	(14.78)***	(6.58)***	(0.61)
Rho1		0.664	
		(3.60)***	
Rho2		0.415	
		(0.69)	
Observations		4518	
Test: $\beta_e^F nrolled = \beta_{enrolled_{post}}^F$		$\chi^2(1) = 2.31$	p=0.12

The dependent variables Regular, Enrolled irregular and Enroll regular all dummy variables. Regular takes value 1 if the individual have succeeded the BEPC exam. Enrolled irregular takes value 1 if the individual with no BEPC is enrolled at school. Enrolled regular takes value 1 if the individual with BEPC is enrolled at school. Absolute value of z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Estimated coefficients of 3 probit models: Case of the CEP.

	Regular	**Enroll irregular**	**Enroll regular**
	CEP	Enroll primary	Enroll secondary
Age	0.366	-0.286	-0.116
	(19.64)***	(11.14)***	(2.58)***
Female	0.108	-0.454	0.100
	(1.76)*	(5.33)***	(0.71)
Head has primary level	0.446	0.673	-0.090
	(6.96)***	(8.99)***	(0.58)
Head has secondary level	0.851	0.957	0.205
	(11.29)***	(8.75)***	(1.13)
Head has university level	1.040	0.929	0.220
	(9.08)***	(4.12)***	(0.82)
Head sex (Male=1)	-0.234	-0.423	0.169
	(4.35)***	(5.02)***	(1.49)
Age of head	0.011	0.008	0.008
	(5.06)***	(2.77)***	(1.56)
Muslim	0.004	-0.159	0.103
	(0.08)	(1.86)*	(0.97)
Christian	0.375	0.800	0.178
	(6.54)***	(9.87)***	(1.34)
Estimated expenditure per head	0.071	0.035	0.054
	(6.41)***	(1.82)*	(2.16)**
Rural	-0.028	0.008	0.151
	(0.38)	(0.09)	(0.92)
Head or spouse non wage worker	-0.065	0.029	-0.335
	(0.95)	(0.27)	(2.25)**
Head or spouse (non wage) in agri	-0.093	-0.057	0.282
	(1.30)	(0.58)	(1.76)*
Head or spouse (non wage) in trade	-0.098	-0.064	0.050
	(1.47)	(0.71)	(0.35)
Son/daughter of head	0.066	0.388	0.478
	(1.41)	(5.76)***	(4.72)***
Distance to private primary school	-0.036	-0.005	
	(5.00)***	(0.84)	
Distance to public primary school	-0.015	-0.017	
	(0.79)	(1.60)	
Distance to private secondary school	-0.004		0.017
	(0.65)		(1.27)
Distance to public secondary school	-0.039		-0.083
	(5.49)***		(5.03)***
Local participation to labor market	-0.130	0.114	-0.919
	(0.74)	(0.46)	(2.27)**
Local participation rate to informal sector	-0.065	-0.004	-0.133
	(0.48)	(0.02)	(0.41)
# of 0-5 years	-0.031	0.004	0.039

Continued on Next Page...

Table 10 – Continued

	Regular	Enroll irregular	Enroll regular
	(1.54)	(0.15)	(0.83)
# of male of 6-11 years	0.011	0.016	-0.064
	(0.47)	(0.55)	(1.18)
# of female of 6-11 years	0.035	-0.003	0.074
	(1.49)	(0.11)	(1.28)
# of male of 12-15 years	0.068	0.056	0.020
	(2.08)**	(1.22)	(0.26)
# of female of 12-15 years	0.042	-0.013	0.059
	(1.31)	(0.28)	(0.79)
# of male of 16-19 years	0.035	0.003	0.078
	(1.20)	(0.06)	(1.13)
# of female of 16-19 years	0.097	0.026	0.007
	(3.03)***	(0.54)	(0.09)
# of male of 20-35 years	-0.009	0.054	-0.000
	(0.40)	(1.42)	(0.01)
# of female of 20-35 years	0.035	0.020	-0.074
	(1.29)	(0.52)	(1.21)
# of over 35 years	-0.018	0.105	-0.026
	(0.64)	(2.79)***	(0.39)
Constant	-6.320	3.255	2.740
	(19.55)***	(7.50)***	(3.61)***
Observations	5067	2991	2076

The dependent variable CEP, Primary and secondary are all dummy variables. CEP takes value 1 if the individual have succeeded the CEP. Primary takes value 1 if the individual with no CEP is enrolled at school. Secondary takes value 1 if the individual with CEP is enrolled at school. Absolute value of z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 11: Descriptive statistics of the variables: other sub samples

	AGE 12-15				Age 16-19			
	NO CEP		With CEP		NO BEPC		WITH BEPC	
	mean	sd	mean	sd	mean	sd	mean	sd
Age	13.21	1.10	13.82	1.06	17.36	1.09	17.84	1.02
Female	0.48		0.52		0.52		0.44	
Head has primary level	0.38		0.30		0.38		0.23	
Head has secondary level	0.22		0.46		0.32		0.51	
Head has university level	0.03		0.12		0.05		0.19	
Head sex (Male=1)	0.80		0.75		0.78		0.74	
Age of head	48.00	12.65	46.12	12.34	44.73	14.10	45.94	11.99
Muslim	0.33		0.46		0.38		0.52	
Christian	0.59		0.81		0.68		0.84	
Estimated expenditure per head	2.06	2.71	4.15	3.18	2.92	2.89	5.54	3.36
Rural	0.42		0.15		0.30		0.05	
Head or spouse non wage worker	0.79		0.55		0.69		0.46	
Head or spouse (non wage) in agri	0.54		0.27		0.40		0.15	
Head or spouse (non wage) in trade	0.21		0.20		0.22		0.22	
Son/daughter of head	0.73		0.70		0.55		0.64	
Distance to private primary school	3.43	5.86	1.07	2.51	2.38	4.84	0.72	1.59
Distance to public primary school	1.35	2.60	0.82	0.82	1.05	1.74	0.71	0.71
Distance to private secondary school	5.66	7.10	2.51	4.20	4.19	6.17	1.70	3.00
Distance to public secondary school	5.09	6.08	2.10	2.62	3.77	5.10	1.63	1.63
Local participation to labor market	0.76	0.18	0.68	0.16	0.72	0.17	0.63	0.14
Local participation rate to informal sector of adults	0.49	0.23	0.39	0.19	0.43	0.21	0.33	0.16
# of 0-5 years	1.28	1.38	0.91	1.10	1.06	1.27	0.72	0.99
# of male of 6-11 years	0.87	1.08	0.73	0.89	0.65	0.91	0.54	0.80
# of female of 6-11 years	0.82	1.01	0.70	0.93	0.62	0.92	0.54	0.80
# of male of 12-15 years	0.96	0.85	0.88	0.84	0.48	0.73	0.44	0.67
# of female of 12-15 years	0.90	0.87	0.97	0.84	0.49	0.74	0.54	0.78
# of male of 16-19 years	0.42	0.72	0.49	0.77	0.88	0.92	0.97	0.84
# of female of 16-19 years	0.38	0.64	0.48	0.71	0.85	0.77	0.86	0.85
# of male of 20-35 years	0.51	0.87	0.59	0.99	0.71	0.95	0.82	1.19
# of female of 20-35 years	0.73	0.89	0.74	0.82	0.66	0.87	0.72	0.90
# of over 35 years	1.68	1.03	1.56	0.96	1.47	1.08	1.56	0.96
	2991		2076		3700		818	

The mean of a dummy variable is proportion.