UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2017/70

Exploring regional and gender disparities in Beninese primary school attendance

A multilevel approach

Kyle McNabb*

March 2017

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: This study combines household survey data from the Beninese Demographic and Health Survey with school supply statistics in order to investigate regional and gender disparities in primary school attendance rates in Benin. Despite almost unparalleled increases in enrolment since the 1990s, Benin remains virtually ignored in the literature surveying school attendance. Results of a logistic regression model highlight the important role played by factors such as household wealth and religion and show that, despite progress, gender disparities in education persist in Benin. The opportunity cost of attending school is also investigated and, in order to account for regional disparities in attendance, a multilevel model is estimated; results from a random slopes model highlight those communes where reductions in the cost of schooling could see the greatest improvements in attendance rates.

Keywords: economic development, demand for schooling, Africa, school attendance, multilevel model
JEL classification: I25, O12, I21

Acknowledgements: I acknowledge support from Heriot-Watt University, Edinburgh.

* UNU-WIDER, Helsinki, Finland, McNabb@wider.unu.edu

This study has been prepared within the UNU-WIDER project on 'Gender and development'.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-294-6

Typescript prepared by Joseph Laredo.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Denmark, Finland, Sweden, and the United Kingdom.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Benin has seen almost unparalleled improvements in primary school attendance since 1990, yet remains virtually ignored in the literature surveying education outcomes in developing countries. The Gross Enrolment Rate (GER)—the number of students attending school as a percentage of the school-age population—soared from around 50 per cent in 1990 to well over 100 per cent in 2012.[1] The Net Enrolment Rate (NER), defined as the percentage of school-age students attending school, stood at 95 per cent in 2012 (UIS 2105), whilst the average in Sub-Saharan Africa (SSA) was just 77 per cent. However, despite these impressive increases in enrolment, severe geographical and gender disparities remain in Benin's primary school attendance rates. According to the most recent Demographic and Health Survey (DHS) data, there are still communes where fewer than one-third of children attend.[2]

This study complements individual- and household-level data from the DHS with detailed commune-level schooling data from the Beninese Institute for Statistics (INSAE) in order to assess the importance of demand-side (i.e. on the part of the student, or their family) and supply-side factors (i.e. availability of school facilities) on school attendance rates. Results from a logistic regression model suggest that household wealth, religion, parental education, and the supply of schools all predict the likelihood of a child attending school. Results also indicate that as average distance to school increases, the likelihood that boys who work in the field will attend school decreases, to a greater extent than those who do not work. This echoes the findings of, for example, Huisman and Smits (2009) and Lincove (2012) and also presents evidence that distance to school acts as a useful proxy for the opportunity cost of attending, which is greater when time spent travelling to school replaces potentially income-generating work.

Whilst much of the existing research on primary school attendance (e.g. Huisman and Smits 2009) has acknowledged that factors at the community, district, or national level play an important role in explaining school attendance, few have sought to explicitly model this econometrically. As a result, estimations fail to account for cluster-level interdependence. This study also employs a multilevel logistic model, which accounts for unobserved heterogeneity between higher-level clusters (i.e. at the household and commune level), in order to assess the level at which most variation in school attendance rates is seen. A three-level *random intercepts* model shows that there are a large number of communes with a significantly lower than average primary attendance rate. However, only around 11 per cent of this variation is attributable to factors at the commune level. After controlling for the number of schools and average distance to school, less than 5 per cent of the variation in attendance is due to commune-level factors, suggesting that the majority of the between-commune variation in attendance rates is attributable to factors at the household level.

In a similar vein to Delprato and Sabates (2015), a *random slopes* model is also estimated, where the wealth coefficient is allowed to vary between communes. This identifies those communes where attendance is below average but the effect of household wealth on school attendance is above average. Such results could aid policymakers in identifying those areas where interventions that

[1] Impressive GERs, however, cannot be taken at face value. By definition, a value greater than 100 might point to a system playing catch-up, or a large number of students entering late or repeating grades. Both are likely to apply in the Beninese context.

[2] Benin is divided into 12 departments and 77 communes.

raised household wealth (or lowered the cost of schooling) might be most effective in increasing school attendance rates.

This study also investigates the problems inherent in the measurement of school enrolment or attendance; where public officials such as schoolteachers or principals have an incentive (due to top-down funding replacing school fees, as was the case in Benin 10 years ago) to report higher enrolment rates, then official statistics might inflate the true number of attendees. Indeed, evidence is presented that enrolment figures from the Beninese DHS are somewhat lower than those from the UNESCO Institute for Statistics (UIS) or INSAE.

The broad contributions to the related literature are therefore threefold: first, evidence is presented on the determinants of primary school attendance for Benin, a context that has not previously been considered in the empirical literature surrounding primary schooling. Second, the use of a multilevel model helps to provide insights into regional disparities, which many similar studies, from other countries, neglect to consider. Finally, the comparison between official statistics and household survey data provides further detailed evidence of the problems with using official statistics, noted in a number of other studies.

This paper proceeds as follows. Section 2 provides a discussion regarding the disparities in enrolment statistics from different sources, before considering the Beninese context in detail. Section 3 discusses the theoretical predictions and empirical results surrounding school attendance in developing countries, with an emphasis on not only the economic rationale, but also the sociocultural factors that might dictate whether or not a child is sent to school. In particular, results from similar studies in SSA are surveyed. Section 4 presents the variables and methodology chosen for this paper. Results for the single-level logistic model are presented in Section 5, whilst the multilevel strategy is outlined and presented in Section 6. Section 7 concludes.

2 Measuring attendance: the Beninese context

2.1 Data considerations

The various sources reporting enrolment or attendance statistics for the period in question in Benin appear to tell somewhat different stories. Table 1 illustrates that official enrolment statistics, from UIS or INSAE, are consistently higher than those from the DHS.

Table 1: Differences in enrolment statistics by source

Source	Indicator			
	NER (%) 2006	NER (%) 2012	GER (%) 2006	GER (%) 2012
DHS	57.01	71.05	86.16	96.41
UIS	82.58	94.86	98.79	122.77
INSAE	-	-	99.59	-

Sources: As stated.

However, there are a number of reasons why statistics from the DHS might not only differ from those supplied by the government (INSAE, UIS[3]) but also be a more reliable and useful indicator of school attendance. Over-reporting on the part of public officials can lead to upward bias in

[3] Whilst not identical, the UIS and INSAE statistics track each other very closely.

enrolment statistics (Glewwe and Kremer 2006; Sandefur and Glassman 2015). Sandefur and Glassman (2015) investigated such a bias toward over-reporting enrolment statistics in a panel of 46 surveys in 21 African countries, finding this bias to be prevalent in cases where low-level public servants (in this case, teachers or school principals) had incentives to misreport official statistics. This was particularly true in countries such as Kenya or Rwanda, where pupil fees had been replaced by top-down per pupil grants—exactly the case in Benin during the period in question here. These authors, and also fhi360 (2013), point out that part of the discrepancy between *enrolment*, which is reported in official statistics, and *attendance*, which is measured by the DHS, might also arise from children that enrol in school but rarely attend. Indeed, the fact that enrolment is now free in Benin might lead some to enrol with little intention of ever attending.

Thus, due to the potential question marks raised over the validity of official statistics, this study will rely on the DHS data where possible, as these are widely used, understood, and accepted as representative. The respondents are unlikely to answer strategically, unlike public officials, who might have incentives to over-report the number of children enrolled in order to maximize funding for their schools. Moreover, an indicator of *attendance* is perhaps a more useful indicator than *enrolment* in the sense that it captures the number of children who actually attend school, rather than those merely registered to attend. Unfortunately, however, the DHS does not investigate the frequency of attendance at school, asking only whether a child attended or not in a particular school year.

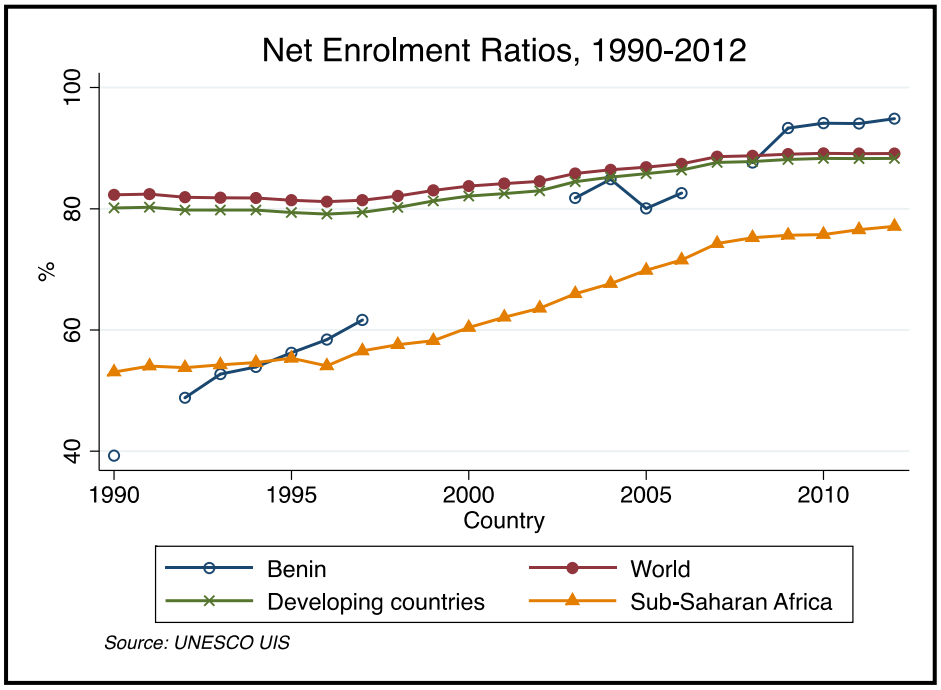
2.2 The Beninese context

Benin provides an intriguing case study in education and development. Along with an economic crisis that forced the closure of teacher training colleges and large-scale cuts to the civil service in the 1980s, the socialist regime's failed attempts at reform (see Allen 1989) left Benin ranking amongst the worst performing countries in the world with regard to GERs and gender parity; fewer than 50 per cent of all children and fewer than one in three girls were attending school in 1990; the ODI (2011: 4) described the education system at this time as 'deeply dysfunctional and inequitable'. However, the democratically elected government prioritized education in 1990 and, as a result of systemic reform, Benin has seen almost unparalleled (at least in SSA) progress in terms of enrolment rates.[4] Between 1990 and 2010, average adult years of education (Barro and Lee 2013) increased from 2.13 to 4.35 (only Mali saw a greater increase in the same period), gross and net enrolment soared, and the gender gap was virtually eliminated in many regions. In 2006, primary education was made free for all. Figure 1 plots Benin's progress in net enrolment rates against the average for SSA, developing countries, and the world. Despite missing data for many years, the trend is clear: NERs rose from around 40 per cent (52 per cent male, 27 per cent female) in 1990 to 94 per cent (99 per cent male, 88 per cent female) in 2012.[5]

[4] For a more detailed account of the backdrop to the 1990 reforms, see ODI (2011).

[5] I rely here on the UIS data, as it allows comparisons with world and regional averages.

Figure 1: Net enrolment ratios, 1990–2012

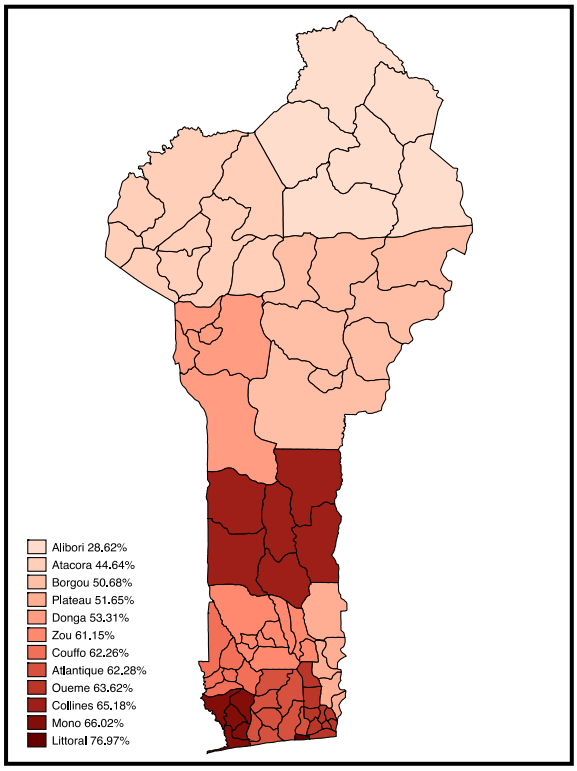


Gross enrolment rose from 51 per cent to 123 per cent in the same period. Progress in increasing the enrolment of girls in school has been particularly impressive: the Gender Parity Index (GPI) rose from 0.50 in 1990 to 0.89 in 2012; only Guinea matched this progress in SSA.[6]

Yet these figures do not tell the whole story in Benin, as large regional disparities persist at both the department and commune level. Figure 2 shows primary school attendance (male and female combined) by administrative department in 2006. The data in this case come from the 2006 wave of the DHS. Given the discussion above regarding statistics from the DHS, these numbers might best be defined as *net attendance rates*; i.e. the percentage of primary-school-age children whose parents reported that they were attending primary school in 2006.

[6] The GPI is calculated as female gross enrolment divided by male gross enrolment. Thus this statistic equals 1 when gender parity is achieved.

Figure 2: Net primary school attendance by department, 2006



Mean = 57%.

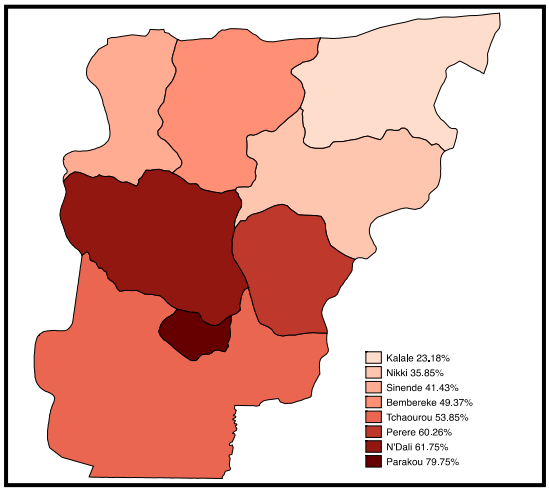
Source: Author's calculations from the 2006 DHS.

In 2006, some regions were still struggling with net attendance rates of as little as 29 per cent (department average), whilst others saw rates of over 75 per cent. The national average was 57 per cent.[7] It is also insightful to examine the differences at the communal level. The department of Borgou, for example, is depicted in Figure 3. Net attendance rates in 2006 ranged from just 23 per cent (Kalale) to almost 80 per cent (Parakou). Whilst the Beninese government has continued to prioritize access to education, regional disparities such as those outlined above persist. Indeed, by 2011/12 many communes had net attendance close to 90 per cent, but some still lagged behind in the 20–30 per cent range.[8]

[7] Note that this number differs somewhat from that reported in Figure 1, for the reasons outlined in Section 4.

[8] Unfortunately, INSAE has not made available its school supply statistics for 2011/12, so the main analysis here focuses on the 2005/06 round of the DHS, where complementary statistics from INSAE are available.

Figure 3: Net enrolment rates by commune, Borgou department, 2006



Source: Author's calculations from the 2006 DHS.

Whilst the NER (or in this case, the net *attendance* rate) represents a significantly more useful indicator than *gross* enrolment, it still does not tell the whole story; the official definition (UNESCO 2012) is: 'The number of children of official primary school age who are enrolled in primary education as a percentage of the total children of the official school age population'. It therefore pays no attention to what grade a child is in; an 11-year-old, who is, strictly speaking, of school age, having just entered the first grade is counted as enrolled—yet they have arrived in primary school some five years late. Whilst the empirical analysis here does not specifically consider on-time enrolment, a look at the age distribution of those attending primary school is nonetheless interesting. Figure 4 uses data from two waves of the DHS in order to highlight changes in the age distribution of primary enrolment in Benin between the 2005/06 and the 2011/12 school years. In 2011/12, the mode age of primary school attendees was 9 years old, at which over 75 per cent of children were attending primary school; this compares with 10 years old in 2005/06.

Indeed, the ages containing the highest percentage of children in school in 2011/12 were 6, 7, 8, 9, 10, and 11—pertaining to the six official years of school. In 2005/06, the ages with the highest percentages of children in school were 8, 9, 10, 11, 12, and 13. Given that official school age in Benin is 6–11 years old, this might reasonably be taken as an indication that, by 2011/12, more children were attending and completing school earlier, if not still strictly on time. Some of the biggest improvements come when looking at older children: the earlier survey showed that, for example, around 17 per cent of 16-year-olds were still in primary school. By 2011, that number had fallen to just 7 per cent.

Figure 4: Primary school attendance age distribution: 2005/06 vs. 2011/12

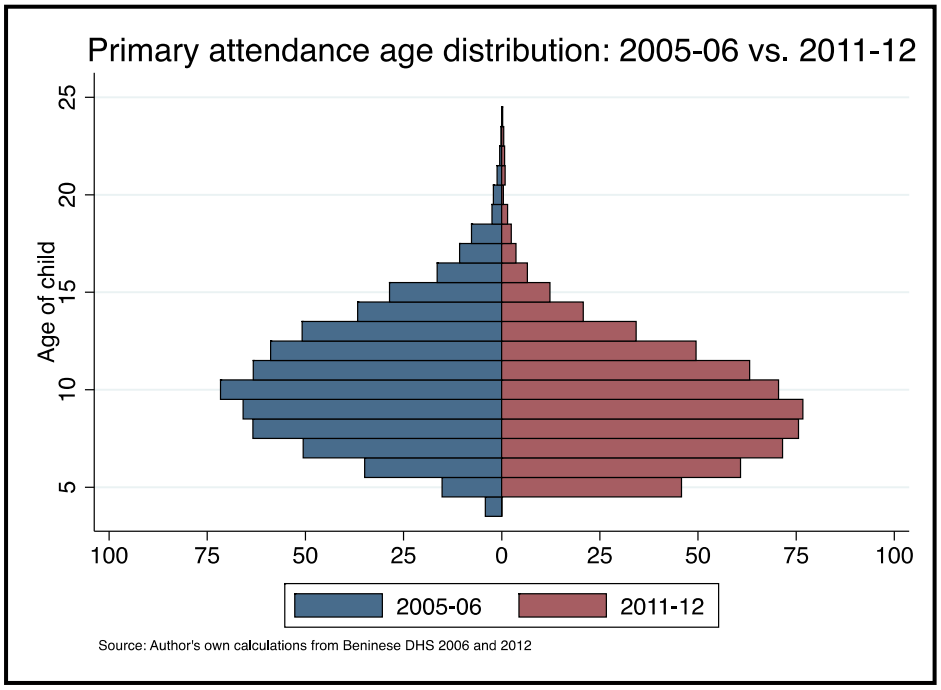
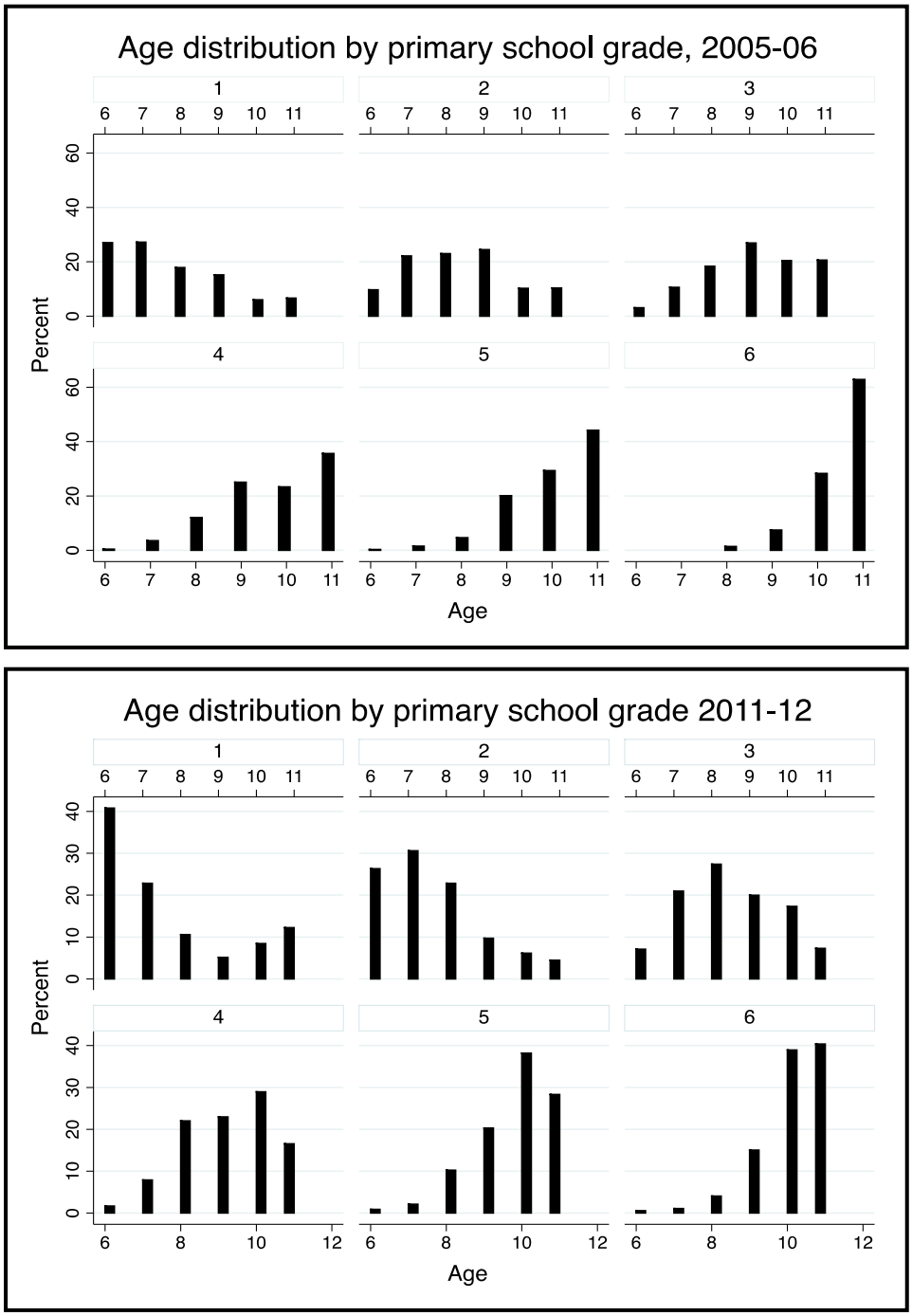


Figure 5 goes further, comparing the age distribution of school-age students in each primary grade, in 2005/06 and 2011/12. There is a clear trend toward earlier, or *on-time* enrolment of students: The 'official' ages for grades 1–6 are 6–11 years old. The 2005/06 data show that many students were either entering school late, or playing catch-up by repeating grades, with the result that the mode age for each grade was often higher than it should have been. By 2011/12, dramatic improvements had been seen in this regard: with the exception of the 4th grade, the most common age of children in each grade was as expected. Whilst Figure 5 restricts the sample to those of school age, the conclusions (with regard to the mode age in each grade) are unaltered when older students are accounted for, though the spread of the distribution is somewhat wider.[9]

[9] Not shown, but available upon request.

Figure 5: Age of children by grade, 2005/06 and 2011/12



Source: Author's own calculations from the 2006 and 2012 DHS.

3 Theoretical predictions and empirical evidence

3.1 Demand side

Becker (1975: 45) argues that 'The most important single determinant of the amount invested in human capital may well be its profitability or rate of return.' In other words, it makes sense to invest in human capital, or in this case send a child to school, if the expected benefits outweigh

the costs incurred.[10] Costs, of course, are measured in both direct and indirect terms—the former constituting items such as school fees, books, transport, and uniforms, and the latter the opportunity cost of attending (Becker 1975). The opportunity cost of attending school is often higher in developing countries than elsewhere: children are expected to work in order to contribute to total household income, particularly in rural areas. Often, children enrolled at the start of the school year are pulled out of school and required to help with the harvest in autumn (Colclough et al. 2000). Children might make a direct contribution to household income, by working on the farm or in the marketplace, or an indirect contribution, where their help around the home or family business frees up the time of adults to earn more money (Colclough et al. 2000). Empirical work often proxies cost of schooling with a measure of distance to school; time spent travelling to school could be used, for example, to work on the family farm, or help with household chores. Studies such as Lincove (2012) and Huisman and Smits (2009) found distance to school to be inversely related to the likelihood of attendance; Delprato and Sabates (2015), however, found no effects on the likelihood of late entry in Nigeria.[11]

Turning to the benefits of attending primary school, these might not be immediately clear to parents or students in a developing country context. If there is a lack of job market opportunities in an area, then there will be a limited expected return to education. Even if opportunities were to arise in the future, parents might not reasonably be able to foresee this happening. If a child is expected to, for example, work on a family plot of land, then numeracy and literacy skills might well be of limited value, at least relative to the physical strength that he or she could have been building, which may prove of more use for his or her future work. More generally, the majority of jobs in an area might not require a formal education, or the education offered at schools may be deemed inappropriate for the predominant type of employment in the area. Similarly, Colclough et al. (2000) highlight that in contexts where gender discrimination exists and the gender balance in labour markets is skewed in favour of males, the benefits of education will be lower for girls. Thus, even with equitable access to schooling, there may still be significant challenges to convincing parents of the benefits of sending their daughters to school.

Weighing up the costs and benefits of sending a child to school requires full knowledge of the potential future benefits. In developing countries, it is by no means guaranteed that parents will be able to accurately measure or estimate such benefits. If information on job market opportunities is unavailable, education is deemed unnecessary for rural farm labour, or if families live in a community where very few adults are educated, then parents (especially those who have not attended school themselves) are likely to undervalue the benefits. In a context where either the future benefits of education are unknown or parents display time-inconsistent preferences (i.e. are hyperbolic discounters who undervalue future benefits), the benefits of education will be undervalued and it is less likely that parents will send their children to school. The costs, whether direct or indirect, are more easily observable.

Costs and benefits must also be weighed in terms of household wealth, income, and expenditure. If the costs of education are small in relation to any of these measures, then it is more likely that parents can afford to send their children to school. It is commonplace for household surveys to stratify households into wealth quintiles (the DHS, for example, does this via a principal components analysis); empirical results often find that the likelihood of attending school increases

[10] This is in a household production function framework, where parents are deemed to make investment decisions on the part of all household members.

[11] Often it is not possible to have a precise measure of distance between house and school, so approximations based on the population and area of a state must be made (e.g. Huisman and Smits 2009). When even these data are missing, a simple rural or urban dummy might be included in estimations and, within reason, pick up some of the same effect.

from the lowest quintile to the highest (see, for example, Huisman and Smits 2009; Kazeem et al. 2010), although this is not always the case. Lincove (2012), for example, found no direct effect of household wealth on school attendance in Uganda, but did find that the effects of other explanatory variables varied in magnitude according to the wealth quintile. Delprato and Sabates' (2015) multilevel analysis of late entry to schooling in Nigeria highlighted that community- or state-level wealth effects were greater than those at the household level.

The economic rationale for sending a child to school is, however, just one side of the story: there are various sociocultural differences with respect to religion, caste, tradition, or tribe that might interlink with economic decisions to dictate the norms followed by parents with regard to education. Indeed, it is probable that these factors are driven more by norms at the community or state level than at the household level (Delprato and Sabates 2015). In societies where patriarchal norms persist, parents may place a higher value on the education of boys than girls. This often stems from the lack of a social security or pension system, meaning that male children are expected to provide for their parents in old age; concurrently, it is common that daughters join their husband's family at marriage (Colclough et al. 2000; Huisman and Smits 2009), and thus their own parents will realize no financial reward from their education.

However, the findings of Eloundou-Enyegye and Calvès (2006) provide a challenge to this traditional viewpoint: their study suggests that in some contexts, married women often remit to their own families (e.g. in Cameroon) and that their capacity to do so actually increases the more educated they are. Thus a potential paradox exists: parents may be unwilling to educate their daughters in the first place, but more educated daughters might actually remit more money. Eloundou-Enyegye and Calvès (2006) also note the significant control that women in West Africa, and Benin in particular, have over their own earnings—a finding corroborated by field surveys cited in LeMay-Boucher and Dagnelie (2014), which confirm the existence of disconnected financial spheres between husband and wife in Benin. A further disadvantage for girls is that, if the nearest school is quite far away, they might not be allowed to attend until they are slightly older, due to the perceived dangers of walking alone, or the physical effort of walking a long distance.[12]

Religion also plays a significant role in the likelihood of children—specifically girls—being sent to school. Csapo (1981) cites the distrust of Western education by Muslims and traditional Islamic views on the education of women, as outlined in the Qur'an, as potential barriers to education for girls in Nigeria. Lincove (2015) found that Muslim children in Nigeria were, on average, 23 per cent less likely to attend school—but the effect was more than double for girls (31 per cent) than for boys (15 per cent)—a finding echoed by Kazeem et al. (2010), who found the same result for Nigeria, although the order of magnitude is dramatically larger: their regressions suggest that Muslim children were five times less likely to attend school than Christians. Lincove (2012), however, found no effect of being a Muslim on school attendance in Uganda, and Buchmann's (2000) regression results found that Muslim children were no less likely to attend school than their Christian counterparts in Kenya. So, whilst theory predicts that different religions or traditions might place different importance on schooling, the empirical evidence for SSA is mixed. It may well be the case that in many countries, parents of all religions are increasingly willing to educate both their sons and daughters.

The education level of parents is another important consideration: it is highly probable that if parents have attended school themselves, and benefited from the education received, they will be more likely to send their own children to school. In terms of the economic rationale, this might

[12] See Colclough et al. (2000) for a thorough discussion of the barriers that face many girls in SSA.

allow the parent to better estimate and appreciate the benefits of education for their own child; this may be particularly true for girls if their mother has attained a certain level of education. Lincove (2015), for example, found that mothers' years of education was a significant determinant of school attendance for both girls and boys in Nigeria and Uganda. Huisman and Smits (2009) also uncovered positive effects of both parents' education level on school attendance; the effect of mothers having at least some primary education was stronger for girls' than boys' likelihood of being in school.[13]

Various household- or family-level factors such as household size, birth order, and whether or not a child is adopted might also be considered as explanatory factors with regard to a child's likelihood of attending school. In terms of economic rationale, a larger number of children increases competition for limited household resources (Lincove 2015). As a result, it might be that children with more siblings cannot attend school, as they are required to complete more tasks at home in order to contribute to household income; and the direct costs of sending many children to school are obviously higher. At the same time, however, a larger number of children might well increase the likelihood that they can attend school, as there are more hands to work and contribute to overall household income. Thus the direction of this effect, if it exists, is unclear. Colclough et al. (2000) presented evidence that, in Ethiopia, the average number of children in a household was higher for school attendees than for dropouts; Glick and Sahn (2006) found that, as the number of children in a household increased, there was no impact on the likelihood that a child would attend a public primary school in Madagascar, although it did decrease the likelihood of a child attending private school. Similarly, Lincove (2015) found no effect of the number of children in a household in Uganda and only marginally significant effects for girls in Nigeria.

Where the number of children is large, it might be that only older (or younger) children are allowed to attend school; as a result, a number of studies have considered the effects of birth order. Huisman and Smits' (2009) regressions for a sample of 30 developing countries indicate that later-born girls are more likely to attend school than first-born girls; and Kazeem et al. (2010) found evidence that having at least one older brother or more than two older sisters increases the odds of attending school in Nigeria. Chernichovsky's (1985) results suggested that an increase in the number of children aged 7–14 increased the likelihood of a child's attending school in Botswana.

3.2 Supply side

The above discussion has highlighted that numerous factors interlink to dictate whether or not a child will be sent to school by his or her parents. However, to consider only factors on the demand side ignores many of the considerations that a parent might take into account. Including estimates of the supply of schooling, and the quality of that supply, is crucial to avoid omitted variable bias in the estimates obtained in empirical analyses. Often, however, institutional data for the total number of schools, pupil–teacher ratios (PTRs), etc. is available only at the national or regional level, making consolidation with survey data at the community or village level very difficult. As a result, many studies fail to accurately account for the supply of schooling.[14] A measure of distance to school might be considered as a proxy for the supply of schooling. School quality is often measured by the PTR, although this variable is clearly endogenous in regressions where the dependent variable is school attendance rates. Huisman and Smits (2009) employ an estimate of

[13] Parental level of education is a common explanatory factor of the likelihood of being in school. Other studies finding positive effects of said include Buchmann (2000), Deininger (2003), Delprato and Sabates (2015), Glick and Sahn (2006), Kazeem et al. (2010), and Lavy (1996).

[14] Handa (2002) is a notable exception; his study considered a range of factors related to school enrolment in Mozambique on both the demand and the supply side.

the number of teachers per 1,000 children ('Teacher Child ratio'), finding a positive and significant effect on school enrolment. This avoids the endogeneity problem, as the denominator is the total number of children of school age in the population, rather than the total number of students.[15] Other variables that have been used to represent school quality include the percentage of teachers who are qualified (e.g. Huisman and Smits 2009; Lavy 1996), the percentage of teachers who are female (often considered key in explaining girls' participation in education, e.g. Glick 2008; Huisman and Smits 2009), the use of multi-grade teaching (Glick and Sahn 2006), and the condition of classrooms. However, the effect of these variables on enrolment or attendance is mixed and it is not unreasonable to argue that some of these measures might matter more for explaining school *achievement* than attendance.

4 Data and methodology

4.1 Data

The dependent variable of interest in the empirical analysis is a dummy variable equal to 1 if a child of primary school age was attending primary school in 2005/06. With this particular round of the DHS, the date at which the survey was carried out is important; the 2006 wave for Benin took place between 3 August and 18 November 2006. However, the Beninese government completed the elimination of primary school fees for all in 2006. Specifically, this announcement was made on 14 October 2006, which not only fell during the survey period, but also happened to be after the school year had started.[16] As a result, there may be children surveyed in August or September that were not attending primary school who might have done so in the absence of fees.[17] Unfortunately, it is not possible at this stage to pinpoint whether or not this policy was implemented immediately, or indeed was put in place for the following academic year. Nonetheless, in order to avoid any potential effects of this policy change, the dependent variable used is a dummy equal to 1 if a child was attending primary school during the previous school year (October 2005–July 2006).[18] Even then, there may be some confusion on the part of respondents: if they were surveyed in August (and perhaps even September), they may have understood the 'current year' to mean the one just past, as opposed to the one that was about to begin in October. Controls for the month in which the survey took place are included in order to attempt to correct for any misunderstanding on the part of the respondent in later survey months.

The DHS data are stratified as follows:

[15] There may be a greater degree of measurement error here, as population ratios are usually based on estimates.

[16] The school year runs from October to July in Benin.

[17] Information regarding primary school fees is unfortunately unavailable at either the national or the regional level.

[18] The DHS survey asked not only 'Did [the household member] attend school during the current year?' but also 'Did [the household member] attend school during the previous year?' Responses to the latter question are used here.

Benin (**1**)

↓

Departments (**12**)

↓

Communes (**77**)

↓

'Clusters' (Primary Sampling Unit [PSU]) (**750**)

↓

Households (**17,511**)

↓

Individuals (**90,650**)

↙ ↘

Not school age (**73,556**) School age (**17,094**)

Of the 90,650 individuals, 17,094 were of school age. Of these, 57 per cent were attending primary school. Control variables are included at the individual, household, and commune levels; Table 2 presents summary statistics of these. Just under half of the sample was female (48 per cent), 3.5 per cent were adopted, and 73 per cent of children of school age also worked.

Table 2: Summary statistics

Variable	Mean	Std. dev.	Min.	Max.
Net attendance 2005/06	0.570	0.495	0	1
Individual level				
Female	0.475	0.499	0	1
Adopted	0.035	0.183	0	1
Worked	0.729	0.444	0	1
Worked in the field	0.318	0.444	0	1
Domestic work	0.367	0.482	0	1
Other	0.044	0.206	0	1
Household level				
Household religion				
Catholic	0.254	0.435	0	1
Protestant	0.057	0.232	0	1
Other Christian	0.070	0.255	0	1
Celeste	0.051	0.221	0	1
Islam	0.256	0.437	0	1
Vodoun	0.212	0.408	0	1
Other traditional	0.033	0.179	0	1
Other religion	0.013	0.115	0	1
No religion	0.054	0.227	0	1
Household wealth				
Poorest	0.222	0.416	0	1
Poorer	0.220	0.414	0	1
Middle	0.205	0.404	0	1
Richer	0.193	0.395	0	1
Richest	0.160	0.366	0	1
Household head's education level				
Primary	0.214	0.410	0	1
Secondary	0.130	0.337	0	1
Higher	0.018	0.132	0	1
School considered essential?	0.790	0.407	0	1
Household size	8.174	4.104	2	36
Rural	0.639	.480	0	1
Commune level				
Distance to school	1.279	0.971	0.134	5.399
(log) Schools per 5–14-year-olds	-5.977	0.238	-6.577	-5.471
Observations: 17,094				

Sources: DHS; INSAE.

Turning to the household level, a set of dummy variables is included in order to control for the household head's religion, the DHS wealth index (a composite index constructed using principal components analysis, ranking households from 1 ['Poorest'] to 5 ['Richest']) and the education level of the household head (a set of dummy variables for none, primary, secondary, or tertiary).[19] One-quarter of households were Catholic, 26 per cent Islamic, and 21 per cent Vodoun. Around 65 per cent of household heads had no formal education whatsoever; of the remainder that had attended school, 21 per cent had a primary education, 13 per cent a secondary education, and the remaining 2 per cent a university education. Whilst ideally the child's mother's or father's education level would be included in the analysis here (as the head is not always the parent of the children included in the sample), it was not possible to identify the parent of each child from the 2006 Beninese DHS due to missing data. *School considered essential?* is a dummy variable taking the value equal to 1 if the respondent answered yes to the survey question 'Do you need to be able to send

[19] See DHS (2004) for a detailed report on the construction of the wealth index.

children to school?'[20] This variable might reasonably be expected to capture household stated preferences for education—some 79 per cent of households considered school to be essential. The average household size was just over 8 and this ranged from 2 to 36. Also included is a dummy variable equal to 1 if the child resided in a rural area (64 per cent did so).

At the commune level, controls are included for the average distance to school and the (log) number of schools per school-age children in each commune.[21] To construct a measure of average distance to school, I follow Huisman and Smits (2009) as follows:

$$Avg\ Distance = \frac{\sqrt{\frac{km^2}{No\ of\ Schools}}}{\pi} \tag{1}$$

One unavoidable weakness of this calculation is that it assumes children attend school in their home commune. Thus, it cannot account for situations where children live in one commune, but attend school in another. However, this is perhaps the best approximation available, given the data on hand.[22]

4.2 Methodology

Whilst geographical variation in attendance rates is explored in detail below in the form of a multilevel model, a simple logistic regression is presented first. The advantage of doing this lies in the ability to compute average marginal effects, which allow us to gain an understanding of the relative magnitudes of the covariates considered. The single-level logistic regression estimated takes the following form:

$$ln\left[\frac{p_i}{1-p_i}\right] = \beta_{0i} + \beta_1\boldsymbol{S}_i + \beta_2\boldsymbol{H}_i + \beta_2\boldsymbol{C}_i + e_i \tag{2}$$

where the dependent variable is a dummy equal to 1 if child *i* was of official school age and attended school during the 2005/06 school year. **S** is the vector of student-level characteristics; **H** the vector of household-level characteristics, and the vector **C** contains commune-level characteristics.

5 Results

Table 3 presents benchmark results for equation (2). In all following tables, the dependent variable is as outlined above. It can be considered a close approximation to the NER, although it is again perhaps best defined as a net *attendance* rate. Average marginal effects are shown.[23] Column 1

---

[20] The set of possible answers was 'No', 'Yes, essential', and 'Yes, more or less necessary'.

[21] Unfortunately, it was not possible to obtain population estimates for the age range 6–11, the official primary school age. Thus the number of schools per children aged 5–14 is shown.

[22] There are sometimes concerns surrounding endogeneity of distance-to-school variables: for example, the endogenous placement of schools, in areas where enrolment is low. However, this is unlikely to be a concern here: first, the data are cross-sectional and therefore only measure the number of schools at one point in time—any phenomenon where more schools were being built as a result of initially low attendance would not be captured. Second, the dependent variable is measured at the individual level and the distance variable at the commune level, so reverse causality is not likely to be an issue.

[23] See Bartus (2005) for an explanation of why it is preferable to present AMEs when a model contains a large number of dummy variables.

includes the entire sample, whilst columns 2 and 3 perform the same regression for boys and girls, respectively. Looking first at the individual-level factors, the likelihood of being enrolled in primary school is significantly lower for girls than it is for boys: this is in line with expectations, given the considerable gender disparities that exist(ed) in enrolment at primary school in Benin.

Table 3: Determinants of primary school attendance, logit estimation

	1	2	3	4	5	6	7	8	9
Sample	All	Boys	Girls	All	All	All	All	Boys	Girls
Female	-0.088***			-0.088***	-0.0879***	-0.088***	-0.0877***		
	(0.011)			(0.012)	(0.0114)	(0.011)	(0.0114)		
Age	0.481***	0.490***	0.469***	0.484***	0.481***	0.481***	0.481***	0.491***	0.471***
	(0.024)	(0.027)	(0.036)	(0.024)	(0.0237)	(0.023)	(0.0237)	(0.0262)	(0.0362)
Age^2	-0.025***	-0.025***	-0.025***	-0.025***	-0.0252***	-0.025***	-0.0252***	-0.0253***	-0.0251***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.00148)	(0.002)	(0.00148)	(0.00156)	(0.00228)
Adopted	-0.134***	-0.016	-0.203***	-0.134***	-0.135***	-0.136***	-0.136***	-0.0175	-0.204***
	(0.037)	(0.033)	(0.046)	(0.037)	(0.0367)	(0.036)	(0.0361)	(0.0329)	(0.0464)
Worked	-0.023	0.003	-0.048**	-0.024	-0.0229	-0.025	-0.0242	0.0836***	-0.0255
	(0.016)	(0.017)	(0.019)	(0.016)	(0.0158)	(0.016)	(0.0156)	(0.0243)	(0.0374)
Household size	-0.003**	-0.003**	-0.003	-0.003***	-0.00288**	-0.003**	-0.00319**	-0.00354***	-0.00332*
	(0.001)	(0.001)	(0.002)	(0.001)	(0.00125)	(0.001)	(0.00127)	(0.00133)	(0.00182)
Religion									
Islam	-0.073***	-0.089***	-0.057**	-0.078***	-0.0722***	-0.084***	-0.0846***	-0.0953***	-0.0606**
	(0.020)	(0.022)	(0.023)	(0.021)	(0.0197)	(0.019)	(0.0184)	(0.0227)	(0.0252)
Traditional / Other	-0.040***	-0.046***	-0.038**	-0.038***	-0.0412***	-0.037***	-0.0375***	-0.0468***	-0.0365**
	(0.011)	(0.013)	(0.015)	(0.012)	(0.0105)	(0.011)	(0.0112)	(0.0138)	(0.0164)
Household wealth level									
Poorest	-0.287***	-0.316***	-0.283***	-0.288***	-0.291***	-0.298***	-0.301***	-0.317***	-0.284***
	(0.019)	(0.023)	(0.025)	(0.019)	(0.0184)	(0.018)	(0.0182)	(0.0229)	(0.0251)
Poorer	-0.197***	-0.247***	-0.171***	-0.202***	-0.202***	-0.206***	-0.209***	-0.253***	-0.176***
	(0.017)	(0.022)	(0.020)	(0.018)	(0.0158)	(0.016)	(0.0161)	(0.0219)	(0.0207)
Middle	-0.107***	-0.169***	-0.067***	-0.111***	-0.111***	-0.115***	-0.117***	-0.172***	-0.0737***
	(0.015)	(0.022)	(0.019)	(0.016)	(0.0139)	(0.015)	(0.0143)	(0.0211)	(0.0202)
Richer	-0.031*	-0.092***	0.005	-0.034*	-0.0348**	-0.036**	-0.0380**	-0.0954***	-0.000232
	(0.017)	(0.022)	(0.019)	(0.018)	(0.0154)	(0.016)	(0.0157)	(0.0213)	(0.0208)
Household head's education									
Primary	0.073***	0.064***	0.080***	0.079***	0.0733***	0.074***	0.0739***	0.0682***	0.0860***
	(0.011)	(0.016)	(0.016)	(0.011)	(0.0114)	(0.011)	(0.0113)	(0.0152)	(0.0155)
Secondary	0.130***	0.125***	0.136***	0.134***	0.131***	0.131***	0.131***	0.128***	0.139***
	(0.013)	(0.017)	(0.018)	(0.013)	(0.0125)	(0.013)	(0.0127)	(0.0176)	(0.0175)
Tertiary	0.071*	0.170***	0.043	0.072*	0.0725*	0.073*	0.0735*	0.182***	0.0454
	(0.041)	(0.036)	(0.057)	(0.040)	(0.0414)	(0.041)	(0.0407)	(0.0351)	(0.0556)
School considered essential?	0.008	0.004	0.013	0.017	0.00697	0.010	0.0103	0.00935	0.0218
	(0.015)	(0.016)	(0.020)	(0.014)	(0.0148)	(0.015)	(0.0151)	(0.0148)	(0.0198)
Rural	-0.021	-0.030**	-0.014	-0.024*	-0.00245	-0.021	0.327	-0.0338**	-0.0167
	(0.013)	(0.014)	(0.016)	(0.014)	(0.0227)	(0.013)	(0.271)	(0.0150)	(0.0167)
Distance to school	-0.016	-0.032***	0.002	-0.034***	-0.00787			-0.00250	-0.00270
	(0.010)	(0.011)	(0.013)	(0.010)	(0.0131)			(0.0133)	(0.0170)

ln Schools per 5–14-year-olds	0.159***	0.152***	0.165***			0.182***	0.142***		
	(0.044)	(0.041)	(0.056)			(0.039)	(0.0449)		
Rural*Distance					-0.0131				
					(0.0135)				
Rural*ln Schools per 5–14-year-olds							0.0580		
							(0.0452)		
Worked*Distance								-0.0640***	-0.0184
								(0.0129)	(0.0209)
Observations	17,094	8,969	8,125	17,094	17,094	17,094	17,094	8,969	8,125

Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01; Average Marginal Effects shown. Controls for the survey month included but not shown.

Source: Author's estimations.

Specifically, the results suggest that girls aged 6–11 were around 9 per cent less likely to attend primary school than boys. Age and its square suggest a non-linear relationship with the dependent variable, reflecting the pattern outlined in Figure 4 above. The results also suggest that adopted children are less likely to attend school than those related to the family; in a context where income might be low it is understandable that households might give preference to the education of biological children. Interestingly, however, column 3 highlights that the significance of this variable comes entirely from adopted girls, who are around 20 per cent less likely to be enrolled in school than a biological daughter; adopted boys appear to face no such disadvantage compared with biological sons. Further inquiry showed that of four types of child (biological son, biological daughter, adopted son, and adopted daughter), adopted daughters were the least likely to attend school, even when compared with adopted sons.[24] An alternative way to frame this inquiry is to include an interaction term (*female*adopted*) in the full sample. When tested, this yielded an identical result, with a similar marginal effect (0.19). Lincove (2012) found a similar result in Uganda, but her results showed a larger impact for fostered boys than fostered girls—the opposite of what is presented here—whilst Huisman and Smits (2009) found an overall negative effect on the likelihood of foster children attending school in their panel of 30 developing countries. *Worked* is a dummy variable equal to 1 if children in the sample carried out any kind of work alongside studying. Again, girls appear to be at a disadvantage compared with boys, the likelihood of attending school being 5 per cent lower for girls who work than for those who do not.[25] This is explored in more detail below.

Turning to the household's religion, the results suggest that children of Islamic households were around 7 per cent less likely to be sent to school than those from Christian homes (the reference category).[26] Those of parents following a *traditional/other* religion were about 4 per cent less likely to attend. The DHS also inquired as to the ethnicity of individuals, but this was often closely correlated to religion and did not provide any additional insights when tested in the model (not shown). It is clear that, as household wealth increases, so too does the likelihood that children attend primary school. For instance, those in the lowest wealth quintile were some 29 per cent less likely to attend school than those in the richest (the reference category). This echoes results in studies such as Huisman and Smits (2009) and Delprato and Sabates (2015) (Nigeria), which also found an increasing likelihood of school attendance as household wealth level increased, although Lincove (2012) found no effects of household wealth on school attendance in Uganda. The effects of household wealth again suggest some difference by gender: girls from the poorest households face a slightly lower disadvantage relative to boys than those in the upper quintile—perhaps a reflection of the higher value placed on boys' work. This is explored in more detail below. Similarly, household heads that had attended primary or secondary school were more likely to send their own children to primary school than those with no education (the reference category); there is little difference here by gender. Obviously, it might be the case that richer families are often more educated, or it might be the case that Christian families are more likely to be wealthy (reside in wealthier regions, etc.). In order to investigate, Tables 4a and 4b display the predictive margins of being in a Christian home by wealth level and the predictive margins of having a more educated household head by wealth level, respectively.

---

[24] Results not shown, but available upon request.

[25] The results of a logit model where *female* and *worked* were interacted revealed a similar result (not shown), with the AME around -5%.

[26] The 'Christian' category includes 'Protestant', 'Catholic', 'Celsete', and 'Other Christian'.

Table 4a: Predictive margins of religion, by household wealth level

Wealth level	Christian	Predictive margin
Poorest	No	0.394
Poorest	Yes	0.455
Poorer	No	0.489
Poorer	Yes	0.551
Middle	No	0.585
Middle	Yes	0.644
Richer	No	0.655
Richer	Yes	0.710
Richest	No	0.675
Richest	Yes	0.728

Table 4b: Predictive margins of household head's education, by household wealth level

Wealth level	Head's education	Predictive margin
Poorest	None	0.383
Poorest	Primary	0.464
Poorest	Secondary	0.528
Poorest	Tertiary	0.446
Poorer	None	0.478
Poorer	Primary	0.561
Poorer	Secondary	0.623
Poorer	Tertiary	0.543
Middle	None	0.575
Middle	Primary	0.655
Middle	Secondary	0.710
Middle	Tertiary	0.637
Richer	None	0.647
Richer	Primary	0.720
Richer	Secondary	0.769
Richer	Tertiary	0.704
Richest	None	0.667
Richest	Primary	0.738
Richest	Secondary	0.785
Richest	Tertiary	0.723

Source: Author's estimations.

It is clear that, holding wealth level constant, household religion and household head's education still have an impact on the likelihood of a child's attending school. In particular, children of Christian parents are consistently around 5–6 per cent more likely to attend school, regardless of household wealth level. Turning to Table 4b, again it is clear that holding wealth constant, the likelihood that a child is sent to school is greater if the household head has primary education than no education, and greater for those with secondary than primary education. Interestingly, holding wealth level fixed, the household head having a tertiary education does not increase the probability that a child will attend school compared with a secondary education.

Looking at the commune-averaged variables in Table 3, male children in rural areas are less likely to be enrolled than males in urban areas; it also seems that distance to school matters only for boys. This result emerges in direct conflict to the thinking that parents will be less likely to send their female children to school because of the long distance to walk, fear of attack, etc. The supply of schools is, however, important—the variable *Schools per 5–14-year-olds* reflects the number of schools in a given commune, divided by the population aged 5 to 14 years.[27] The positive and significant coefficient on this variable for children of both genders suggests that, having controlled for demand-side factors, school availability also matters for attendance rates. However, these results warrant closer inspection; due to the relatively high correlation between average distance to school and number of schools (-0.57), the effects of these variables are examined in isolation. The results in columns 4 and 6 suggest that both distance and the number of schools affect the likelihood that a child attends. However, when interacted with the rural dummy (columns 5 and 7), it appears that neither distance to school nor the number of schools has a stronger effect in rural areas, although the signs are in the expected direction.

As a further check on the role of distance to school, columns 8 and 9 investigate the extent to which this variable can shed light on the opportunity cost of attendance by subdividing the sample by gender and examining the interaction of distance to school and whether or not a child worked.[28] There is no significant effect of distance on girls' attendance if they worked (column 9), but the result for boys presents evidence that distance to school is an important concern for those that worked alongside studying (column 2). The negative and significant AME on the interaction term suggests that boys who worked alongside studying were even less likely to attend a school that was far away than those who did not work. This lends support to the hypothesis that the opportunity cost of schooling is taken into account when deciding whether or not to send a child to school. Indeed, it suggests that the opportunity cost of sending boys to school is higher than for girls.

A number of explanations for the differences observed between the genders might be considered: first, boys' labour, or the nature of their labour, may contribute more to family income than girls' and so the opportunity costs of travel time are more heavily felt. Indeed, of those children carrying out 'work in the field', 68 per cent were boys; conversely, only 37 per cent of children that carried out 'domestic work' were boys. Second, the climate in Benin at that time—of promoting enrolment for all—might have seen parents under pressure to (be seen to) send their daughters to school, so that even those required to work were encouraged to attend to a greater extent than were sons. This result echoes the findings of, for example, Colclough et al. (2000), who found that boys in both Guinea and Ethiopia that had dropped out of school did so primarily to earn money, although Buchmann (2000) found that from a sample of 146 children who had dropped out of school in Kenya, only one did so for employment and one to help in the household.

5.1 Testing an alternative dependent variable

Whilst the dependent variable used above adheres closely to a measure of net enrolment in Benin, it is not a perfect barometer of primary school attendance. Figure 4 showed that many children in primary school in 2005/06 fell outside the official age category of 6–11. Furthermore, related studies in this field have imposed different criteria for classifying school attendance status: Huisman and Smits (2009) consider only those aged 8–11 in a panel of 30 countries; Lincove

[27] This age range was chosen because INSAE produces population estimates for ages 5–9 and 10–14. It was deemed preferable to use these figures as published, rather than construct an age range of 6–11, which would require the use of arbitrary estimates of population growth.

[28] No significant differences were uncovered here when differentiating between rural and urban, or when including the *Number of schools* variable.

(2015), those aged 6–13 in Nigeria and Uganda (although crucially, they omit any children that had already progressed to secondary school); Buchmann (2000), those aged 13–18 in Kenya; Lincove (2012), those aged 6–12 years old in Uganda. To the best of my knowledge, only Deininger (2003) provides estimates for a number of alternative age ranges (Uganda).

Whilst some of this disparity in the dependent variable used naturally results from different official starting ages and lengths of school cycles in each country, it can hamper the comparability of results across studies and across countries. An appealing approach is outlined in fhi360 (2013), which argues in favour of classing the school age as between 7 and 14 years old and abandoning the arbitrary definitions of 'primary' or 'secondary'. These bounds are influenced by the fact that 'in all countries, compulsory education begins by age 7 or earlier' and 'the ILO minimum age convention establishes age 15 as the minimum legal age for entering any form of employment' (fhi360 2013: 48). As a result, any child falling within this age bracket should be expected to be in school. To follow this definition for the Beninese sample would lead to the loss of those children aged 5 and 6 who are attending school. However, by extending the upper age range of the sample to 14, the analysis can account for many more children attending, or who have completed, primary school. In the robustness checks included in Appendix A, the dependent variable is thus a dummy equal to 1 if (i) a child is aged between 5 and 14 and (ii) (s)he is either in or has completed primary school.[29] All of the previously reported results hold using this alternative dependent variable; the only differences arising are small changes in the magnitude of certain independent variables.

6 A multilevel approach

The results presented in Section 5 have highlighted that factors on both the demand and supply side were important determinants of school attendance in Benin in the 2005/06 school year. As shown in Section 2, stark regional disparities exist in Benin with regard to attendance rates. Whilst a number of studies acknowledge that factors at the household, community, or district level might impact on school attendance in SSA (for example, Handa 2002; Huisman and Smits 2009; Lavy 1996), only Delprato and Sabates (2015) have explicitly modelled this econometrically by taking account of unobserved heterogeneity between higher-level clusters. The consequence for the econometric analysis of ignoring this is a violation of the assumption that observations are independent from one another; unobserved heterogeneity at higher levels leads to cluster-level interdependence between units (Skrondal and Rabe-Hesketh 2004). The traditional approach to dealing with such data is to turn to a multilevel (or hierarchical) linear model (MLM) (HLM), the simplest of which is the random intercepts or *variance components* model, which estimates a random intercept for every higher-level unit, such as commune or department.

6.1 Random intercepts model

After testing various multilevel structures, it turned out that, whilst all performed better than the single-level model above, a three-level model, as shown below, was preferred to any two-level model.[30]

[29] Thus, children currently in secondary school and those who have completed primary school but not gone on to secondary education are also included.

[30] This judgement was made on the basis of LR-test statistics.

The basic variance components model takes the form

$$ln\left[\frac{p_{ijkl}}{1-p_{ijkl}}\right] = \beta_0 + \mu_k + \varphi_{jk} + \varepsilon_{ijk} \tag{3}$$

where

$$\mu_k \sim N(0, \sigma_\mu^2)$$

$$\varphi_{jk} \sim N(0, \sigma_\varphi^2)$$

$$\varepsilon_{ijk} \sim N(0, \sigma_\varepsilon^2) \tag{4}$$

$\ln\left[\frac{p_{ijk}}{1-p_{ijk}}\right]$ is the log-odds that child *i* in household *j* in commune *k* is attending school. β_0 is the intercept shared by all individuals, households, and communes. μ_k is the effect of commune *k*, φ_{jk} is the effect of household *j*, and ε_{ijk} is the child-level residual error.

The models here are computed using second-order penalized quasi-likelihood (PQL2) in *MLwiN* (Rabash et al. 2015) via the Stata module *runmlwin* (Leckie and Charlton 2013). When selecting the appropriate means by which to estimate multilevel equations, it is necessary to choose a method that is unbiased, but also computationally feasible. Simulations in Rodriguez and Goldman (2001) show that, out of the choice of 1st and 2nd order marginal quasi-likelihood (MQL) and 1st and 2nd order PQL, 2nd order PQL estimation provides the closest approximation to maximum likelihood estimation (MLE). Whilst, ideally, MLE would be used to obtain all the estimates, this is computationally very intensive for models beyond the null; Stata's commands such as *xtmelogit* often take many hours or days to converge, if they do at all. Appendix B displays estimates of the variance of the random effects in equation (3) using MQL1, MQL2, PQL1, PQL2, and MLE; whilst there is still a downward bias in the estimates of σ_μ^2 and σ_φ^2 compared with MLE, PQL2 performs substantially better than the other quasi-likelihood estimators. Given that MLE estimates for models beyond the null are computationally very difficult to obtain, PQL2 is the preferred method here.

Column 1 of Table 5 displays results from the random intercepts model, as per equation (3). Odds ratios are displayed, as it is not possible to compute marginal effects for multilevel models. Thus, it is difficult to compare the magnitude of the covariates with the single-level logistic regressions. The estimate of β_0 suggests that the log-odds of a child of school age attending primary school in an 'average' household/commune are 0.32.

In order to examine and interpret each of the variance components outlined above in (4), a variance partition coefficient (VPC) can be calculated, which, for the unconditional model in column 1 of Table 5, 'report[s] the proportion of the observed variance at each level of the model hierarchy' (Leckie 2013: 21). Thus, the VPC provides an indication of those levels at which the most residual variation in the likelihood of attending school exists. The discussion in Section 2 has already highlighted the significant inter-commune disparity in school attendance rates, so the expectation is that a significant amount of variation will exist at this level. The VPCs are calculated for the commune and household, respectively, as follows:

$$VPC_\mu = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\varphi^2 + \sigma_\varepsilon^2} \tag{5}$$

$$VPC_{\varphi} = \frac{\sigma_{\varphi}^2}{\sigma_{\mu}^2+\sigma_{\varphi}^2+\sigma_{\varepsilon}^2} \quad (6)$$

Table 5: Random intercepts and random slopes model

	Random intercepts 1	2	3	4	Random slopes 5
Individual-level characteristics					
Female		-0.538***	-0.582***	-0.582***	-0.578***
		(0.042)	(0.043)	(0.043)	(0.043)
Age		3.171***	3.183***	3.182***	3.163***
		(0.139)	(0.140)	(0.140)	(0.139)
Age^2		-0.167***	-0.167***	-0.167***	-0.166
		(0.008)	(0.008)	(0.008)	(0.008)
Adopted		-0.658***	-0.924***	-0.917***	-0.929***
		(0.117)	(0.118)	(0.118)	(0.117)
Worked		-0.297***	-0.167***	-0.156***	-0.137**
		(0.055)	(0.055)	(0.054)	(0.054)
Household-level characteristics					
Household size			-0.019***	-0.015**	-0.015***
			(0.007)	(0.007)	(0.007)
Religion					
Islam			-0.386***	-0.311***	-0.327***
			(0.078)	(0.079)	(0.079)
Traditional / Other			-0.401***	-0.414***	-0.413***
			(0.063)	(0.063)	(0.063)
Grand mean-centred wealth level			0.507***	0.497***	0.498***
			(0.024)	(0.024)	(0.033)
Household head's education					
Primary			0.460***	0.447***	0.329***
			(0.063)	(0.063)	(0.063)
Secondary			0.714***	0.701***	0.696***
			(0.089)	(0.089)	(0.089)
Tertiary			0.418**	0.368*	0.371*
			(0.207)	(0.206)	(0.206)
School considered essential?			0.154**	0.134**	0.125***
			(0.064)	(0.064)	(0.064)
Rural			-0.248***		-0.236***
			(0.062)		(0.062)
Commune-level characteristics					
Distance to school				-0.165**	-0.0210
				(0.079)	(0.077)
(ln) Schools per 5–14-year-olds				1.073***	1.128
				(0.303)	(0.289)
Intercept (β_0)	0.322***	-13.528***	-13.544***	-6.646***	-6.422***
	(0.089)	(0.577)	(0.603)	(1.842)	(1.764)
Random effects					
Level 2: Household					
Intercept variance	1.223	1.800	1.418	1.422	1.373
VPC_{φ}:	*0.240*	*0.310*	*0.279*	*0.287*	-
ICC:	*0.353*	*0.430*	*0.352*	*0.336*	-
Level 3: Commune					
Intercept variance	0.573	0.700	0.368	0.241	0.214
VPC_{μ}:	*0.113*	*0.120*	*0.073*	*0.049*	-
ICC:	*0.113*	*0.120*	*0.073*	*0.049*	-
Wealth slope variance	-	-	-	-	0.037
		-	-	-	-0.043
Intercept—Wealth slope covariance	-				
Observations:	17,094	17,094	17,094	17,094	17,094

Odds ratios shown. Standard errors in parentheses; *p<0.1; **p<0.05;***p<0.01.

Source: Author's estimations.

In the case of a binary outcome, σ_ε^2 is fixed at $\frac{\pi^2}{3} \approx 3.29$, the variance of the standard logistic distribution. For the null model, the VPCs are 0.240 and 0.113 for the household and commune, respectively. Thus, for the null model considered in column 1, 24 per cent of the variation in school attendance rates is between households, and only 11.3 per cent between communes. An alternative means by which to interpret variance components is the intra-class correlation coefficient (ICC), which measures correlation or similarity between observed responses within a given higher-level cluster unit. The ICCs for the commune level, ρ_μ, and the household level, ρ_φ, are calculated as follows:

$$\rho_\mu = \frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma_\varphi^2+\sigma_\varepsilon^2} \tag{7}$$

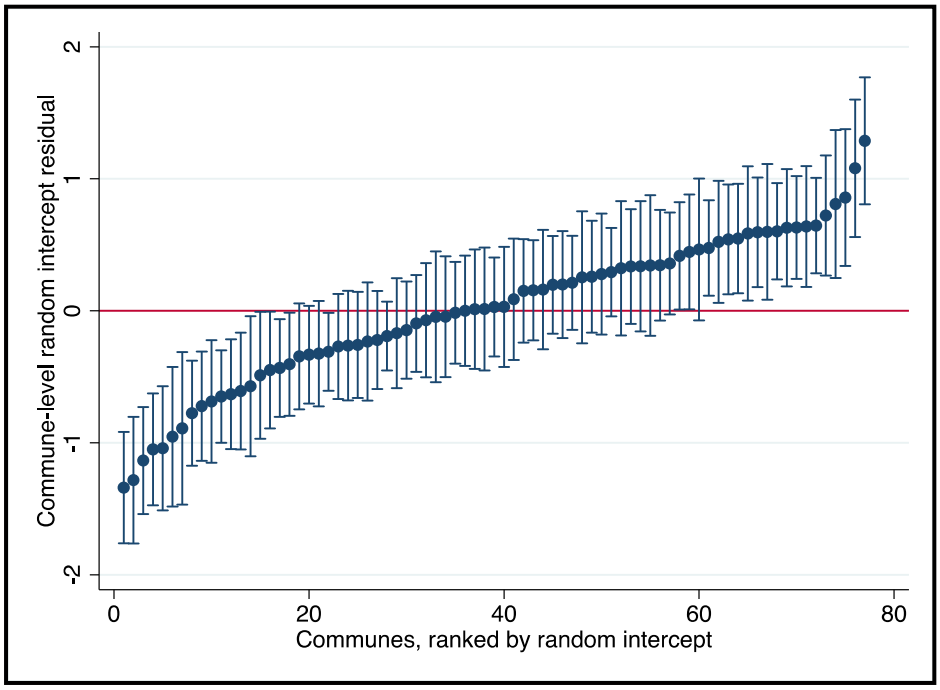
$$\rho_\varphi = \frac{\sigma_\mu^2+\sigma_\varphi^2}{\sigma_\mu^2+\sigma_\varphi^2+\sigma_\varepsilon^2} \tag{8}$$

Thus, an ICC of 0.113 for the commune level represents the between-commune correlation in the odds that a child is attending school. An ICC of 0.353 shows that the between-household within-commune correlation is much higher, suggesting a higher correlation between the odds that any two children from different households in the same commune are attending school and the odds that any two children from different households in different communes are attending school.

A caterpillar plot of the commune-level residuals (Figure 6) shows a significant number of communes where the 95 per cent confidence interval does not overlap with zero, suggesting that attendance rates are significantly higher or lower than average in these communes. The significant *between-commune-within-department* variance, discussed in Section 2, is also confirmed here; Figure 7 illustrates that whilst some departments, such as Alibori and Plateau, contain only communes with a negative random intercept residual, others, such as Atacora, Borgou, and Zou, contain communes where the random effect is both above and below the average.[31]

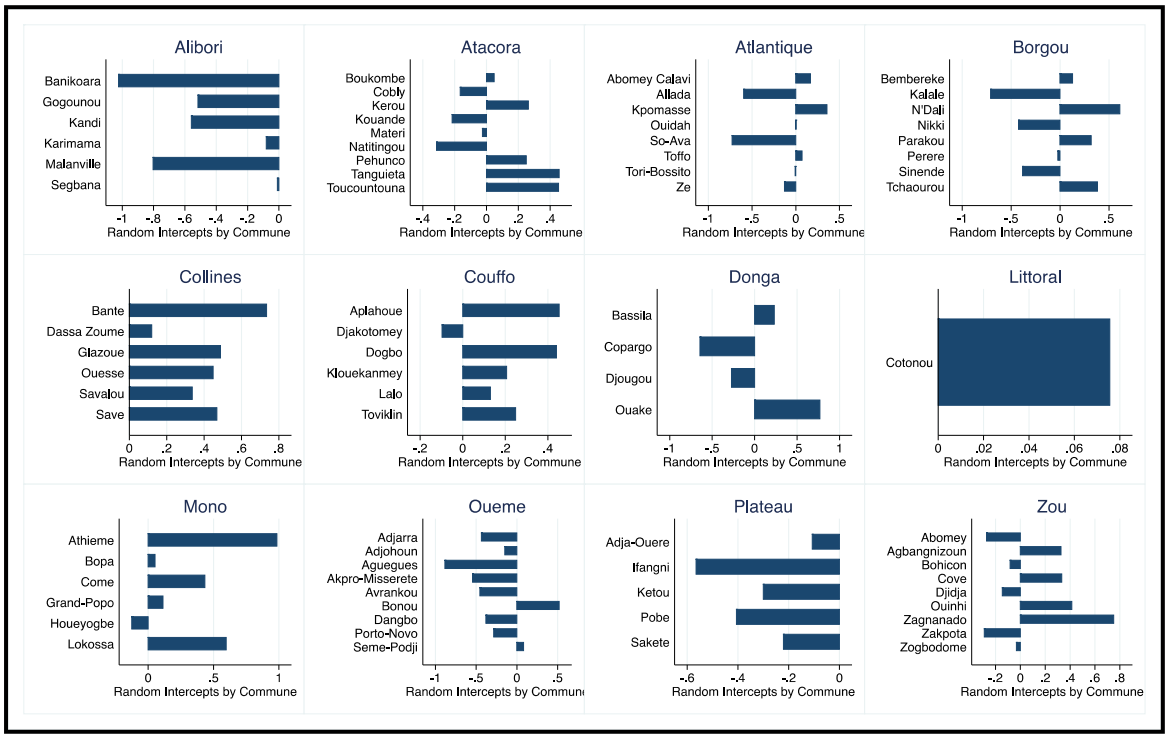
[31] Note that initial diagnostics deemed that the computational complexity of adding a department level to the model in Table 5 led to little gain in model performance. Therefore, it was not included, but the graphical insight here is nonetheless useful.

Figure 6: Random intercept residuals by commune



Source: Computed from random intercepts model, Table 5.

Figure 7: Between-commune-within-department variation in random intercepts



Source: Computed from random intercepts model, Table 5.

In column 2 of Table 5, covariates are included at the individual level. The VPC in columns 2–4 takes on a slightly different interpretation; in conditional models, it represents the degree of *unexplained* variance that exists at each higher level. Having controlled for individual-level factors, we see that 12 per cent of residual variance in school attendance exists between communes and some 31 per cent between households.

When household-level variables are included in column 3 of Table 5, the variance component attributable to the household level falls from 1.800 to 1.418, representing a reduction in between-household variance of 21 per cent.[32] Thus, the household covariates included (household wealth, religion, size, and stated preferences for education) explain around one-fifth of the residual variation in attendance rates. The remaining household-level variation is attributable to some unobserved factors not accounted for here. The VPCs for the household and commune fall to 27.9 per cent and 7.3 per cent, respectively.

In column 4, commune-level variables are added to the model. Immediately clear is that the commune-level variance component falls from 0.355 to 0.241; the VPC for communes falls from 7 per cent to 4.9 per cent. Thus, the inclusion of *distance to school* and *number of schools* in the model helps to explain around one-third of the commune-level variation in primary school attendance. This result suggests that the regional differences observed in Benin are due to other unobserved factors at the commune level, such as regional differences in labour markets, culture, or traditions. The model has also shown that, overall, relatively little of the regional variation in attendance rates displayed in Benin is attributable to commune-level factors: only 4.9 per cent of the total remaining variation is at the commune level, whilst some 30 per cent is due to factors at the household level. Thus, the greatest improvements in attendance rates might be realized by focusing on raising household income, or changing attitudes toward educating daughters.

6.2 Random slopes model

The random intercepts model of Section 6.1 assumed that the effects of each of the independent variables was fixed across communes, and across households within communes. In order to test the validity of this assumption, it is possible to estimate a *random slopes model*, which allows both the intercept and the coefficient (slope) of explanatory variables to vary randomly across higher-level units. The particular focus here is on commune-level effects of household wealth on school attendance. The model takes the form

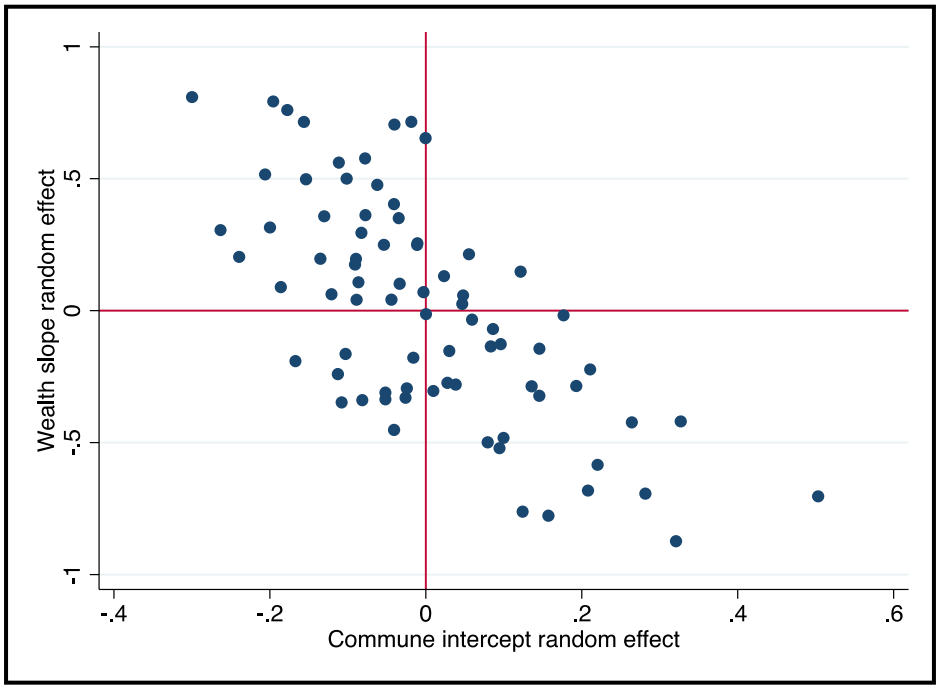
$$\ln\left[\frac{p_{ijk}}{1-p_{ijk}}\right] = \beta_0 + \beta_1 x_{1ijk} \dots\dots + \mu_{0k} + \mu_{1k} x_{1_k} + \varphi_{jk} + \varepsilon_{ijk} \qquad (9)$$

Column 5 of Table 5 reports the results of equation (9), which allows the slope of mean-centred wealth to vary across communes. The effect of wealth on the log-odds of attending school in commune *k* (i.e. the average effect of wealth) is given by $\widehat{\beta_1} + \widehat{\mu_{1k}}$, which in this case is estimated to be 0.498 + $\widehat{\mu_{1k}}$. The between-commune variance in the effects of wealth, $\widehat{\mu_{1k}}$, is estimated to be 0.037.

The between-commune variance in attendance, μ_k, falls from 0.241 to 0.214, which suggests that the distribution of wealth does indeed vary across communes (otherwise, $\mu_k$would have remained unchanged). The estimated *commune intercept–wealth slope* covariance is negative (-0.043), which shows that those communes with below average primary school attendance rates (i.e. where μ_{0k} < 0) also tend to have above average effects of wealth (i.e. where μ_{1k} > 0). Figure 8 plots μ_{1k} against μ_{0k}.

[32] In the multilevel models here, mean-centred wealth is included rather than the individual wealth quintiles, allowing estimation of the random slopes model, below.

Figure 8: Commune slopes vs. commune intercepts



Source: Computed from random slopes mode, Table 5.

In terms of potential policies to increase school attendance, this exposition is useful. Communes in the upper left quadrant represent those where the effect of wealth on attendance is above average, whilst school attendance itself is below average. Therefore, these communes represent those areas where the greatest improvements in school attendance rates could be realized through policy interventions that either raise income or lower the cost of schooling. The analysis of Section 5 suggests that lowering average distance to school (by building more schools, improving the road network, etc.) might free up time for boys to work (and thereby contribute to family income) alongside studying, thus lowering the opportunity cost of attending school. This provides a clear example of how supply- and demand-side considerations work hand in hand to determine whether a child is sent to school.

Conversely, communes lying in the lower left quadrant are those where attendance rates are below average, but the effect of wealth is also limited; increasing wealth levels here might have a limited impact on school attendance rates.

7 Conclusion

This study has sought to shed light on the determinants of primary school attendance rates in Benin, a country that despite seeing almost unparalleled improvements in school attendance, gender parity, and education completion over the last two decades, has been practically ignored in the literature. The results presented here have generated numerous insights.

First, the analysis of Beninese statistics has echoed the findings of, for example, Sandefur and Glassman (2015), who found that administrative statistics (in this case from either INSAE or UIS) overstated school enrolment compared with household surveys (DHS). For the school year 2005/06, DHS estimates of enrolment (attendance) were around 15 percentage points lower than those from UIS or INSAE.

Second, large regional disparities existed during the year of study (2005/06) and, indeed, still do in the most recent DHS data. The empirical analysis has presented evidence that factors on both the demand and supply side are predictors of whether or not a child will attend school. Richer households, those following a Christian religion, and those with more educated household heads were more likely to send their children to school. Despite the narrowing gender gap in Benin at the time, girls still had a lower likelihood of attending school than boys; adopted girls seemed at the greatest disadvantage. Focusing on the role of child labour, the empirical results highlight that boys who worked alongside studying faced a higher opportunity cost of travel time to school, but girls did not face a similar disadvantage.

Whilst much of the literature investigating the determinants of school attendance or enrolment has acknowledged that factors at the community or state level might play a part in determining enrolment trends, many studies continue to ignore higher-level clustering in the data. In attempting to shed light on the large regional disparities that exist in Beninese primary school attendance, the present work has explicitly modelled higher-level variation in school attendance by using a multilevel modelling strategy. A three-level random intercepts model, estimated at the individual, household, and commune levels, highlighted that there were a number of communes where primary school attendance was significantly lower than average. After controlling for individual-, household-, and commune-level factors, the model is able to explain a large portion of the between-commune variance in attendance rates. However, it also suggests that much of the residual variance in attendance between communes is actually due to factors at the household level.

A random slopes model suggests a number of communes where average school attendance is below average and the effect of household wealth is above average. Such regions might stand to benefit most, in terms of school attendance rates, from government policy that raised household incomes, or reduced the costs of schooling.

References

Allen, C. (1989). 'Benin'. In C. Allen, M.S. Radu, and K. Somerville (eds), *Benin, the Congo, Burkina Faso. Politics, Economics and Society*. New York: Pinter.

Barro, R., and J.-W. Lee (2013). 'A New Data Set of Educational Attainment in the World, 1950–2010'. *Journal of Development Economics*, 104: 184–98.

Bartus, T. (2005). 'Estimation of Marginal Effects Using Margeff'. *Stata Journal*, 5(3): 309–29.

Becker, G.S. (1975). *Human Capital, A Theoretical and Empirical Analysis, with Special Reference to Education*, 2nd Ed. New York: National Bureau of Economic Research, Columbia University Press.

Buchmann, C. (2000). 'Family Structure, Parental Perceptions, and Child Labor in Kenya: What Factors Determine Who Is Enrolled in School?' *Social Forces*, 78(4): 1349–79.

Chernichovsky, D. (1985). 'Socioeconomic and Demographic Aspects of School Enrollment and Attendance in Rural Botswana'. *Economic Development and Cultural Change*, 33(2): 319–32.

Colclough, C., P. Rose, and M. Tembon (2000). 'Gender Inequalities in Primary Schooling: The Roles of Poverty and Adverse Cultural Practice'. *International Journal of Educational Development*, 20(1): 5–27.

Csapo, M. (1981). 'Religious, Social and Economic Factors Hindering the Education of Girls in Northern Nigeria'. *Comparative Education*, 17(3): 311–19.

Deininger, K. (2003). 'Does Cost of Schooling Affect Enrolment by the Poor? Universal Primary Education in Uganda'. *Economics of Education Review*, 22: 291–305.

Delprato, M., and R. Sabates (2015). 'A Multilevel Analysis of Late Entry in Nigeria'. *International Journal of Research & Method in Education*, 38(1): 83–99.

DHS (2004). 'The DHS Wealth Index'. DHS Comparative Reports, 6. Available at: http://www.dhsprogram.com/pubs/pdf/CR6/CR6.pdf (accessed 1 March 2015).

Eloundou-Enyegye, P.M., and A.E. Calvès (2006). 'Till Marriage do us Part: Education and Remittances from Married Women in Africa'. *Comparative Education Review*, 50(1): 1–20.

fhi360 (2013). *Out of School Children. Data Challenges in Measuring Access to Education*. Education Policy and Data Center, fhi360, Available at: http://www.fhi360.org/sites/default/files/media/documents/Out_of_school_children_data_challenges.pdf (accessed 25 June 2015).

Glewwe, P., and M. Kremer (2006). 'Schools, Teachers, and Education Outcomes in Developing Countries'. *Handbook of the Economics of Education*, 2: 945–1017.

Glick, P. (2008). 'What Policies will Reduce Gender Schooling Gaps in Developing Countries: Evidence and Interpretation'. *World Development*, 36(9): 1623–46.

Glick, P., and D. Sahn (2006). 'The Demand for Primary Schooling in Madagascar: Price, Quality, and the Choice between Public and Private Providers'. *Journal of Development Economics*, 79: 118–45.

Handa, S. (2002). 'Raising Primary School Enrolment in Developing Countries: The Relative Importance of Supply and Demand'. *Journal of development Economics*, 69(1): 103–28.

Huisman, J., and J. Smits (2009). 'Effects of Household- and District-level Factors on Primary School Enrollment in 30 Developing Countries'. *World Development*, 37(1): 179–93.

Kazeem, A., L. Jensen, and C.S. Stokes (2010). 'School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban–Rural Residence, and Socioeconomic Status'. *Comparative Education Review*, 54(2): 295–319.

Lavy, V. (1996). 'School Supply Constraints and Children's Educational Outcomes in rural Ghana'. *Journal of Development Economics*, 51: 291–314.

Leckie, G. (2013). 'Module 11: Three-level Multilevel Models, Stata Practical'. Centre for Multilevel Modelling. Available at: https://www.cmm.bris.ac.uk/lemma (accessed 10 April 2015).

Leckie, G., and C. Charlton (2013). 'runmlwin – A Program to Run the MLwiN Multilevel Modelling Software from within Stata'. *Journal of Statistical Software*, 52(11): 1–40.

LeMay-Boucher, P., and O. Dagnelie (2014). 'The Divorced Financial Spheres of Beninese Spouses'. *Journal of International Development*, 26(1): 46–58.

Lincove, J.A. (2012). 'The Influence of Price on School Enrolment under Uganda's Policy of Free Primary Education'. *Economics of Education Review*, 31: 799–811.

Lincove, J.A. (2015). 'Improving Identification of Demand-side Obstacles to Schooling: Findings from Revealed and Stated Preference Models in Two SSA Countries'. *World Development*, 66: 69–83.

ODI (2011). 'Benin's Progress in Education: Expanding Access and Closing the Gender Gap'. Overseas Development Institute.

Rasbash, J., C. Charlton, W.J. Browne, M. Healy, and B. Cameron (2015). *MLwiN Version 2.35*. Centre for Multilevel Modelling, University of Bristol.

Rodriguez, G., and N. Goldman (2001). 'An Assessment of Estimation Procedures for Multilevel Models with Binary Response'. *Journal of the Royal Statistical Society*, Series A, 158(1): 73–89.

Sandefur, J., and A. Glassman (2015). 'The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics'. *The Journal of Development Studies*, 51(2): 116–32.

Skrondal, A., and S. Rabe-Hesketh (2004). *Generalized Latent Variable Modelling: Multilevel, Longitudinal, and Structural Equation Models*. Boca Raton, FL: Chapman & Hall/CRC.

UIS (2015). 'Enrolment Ratios'. UNESCO Institute for Statistics. Available at: http://data.uis.unesco.org (accessed 10 September 2015).

UNESCO (2012). 'UIS Frequently Asked Questions. Education Statistics'. Available at: http://www.uis.unesco.org/Education/Documents/FAQ-education-statistics-rev5-en.pdf (accessed 20 October 2015).

Appendix A: Replication of Table 3 using an alternative dependent variable

	1 All	2 Boys	3 Girls	4 All	5 All	6 All	7 All	8 Boys	9 Girls
Female	-0.093***			-0.093***	-0.093***	-0.093***	-0.093***		
	(0.010)			(0.010)	(0.010)	(0.010)	(0.010)		
Age	0.346***	0.352***	0.336***	0.347***	0.347***	0.347***	0.346***	0.353***	0.337***
	(0.008)	(0.011)	(0.011)	(0.009)	(0.009)	(0.009)	(0.008)	(0.011)	(0.011)
Age^2	-0.016***	-0.016***	-0.016***	-0.016***	-0.016***	-0.016***	-0.016***	-0.016***	-0.016***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Adopted	-0.158***	-0.033	-0.235***	-0.158***	-0.159***	-0.160***	-0.160***	-0.035	-0.236***
	(0.041)	(0.023)	(0.052)	(0.041)	(0.04q)	(0.040)	(0.040)	(0.023)	(0.053)
Worked	-0.035**	-0.016	-0.053***	-0.036**	-0.036**	-0.0362**	-0.036**	0.060**	-0.029
	(0.016)	(0.015)	(0.019)	(0.016)	(0.016)	(0.016)	(0.016)	(0.023)	(0.044)
Household size	-0.002**	-0.002**	-0.002	-0.003**	-0.003**	-0.003**	-0.002**	-0.003**	-0.003*
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Religion									
Islam	-0.069***	-0.083***	-0.055***	-0.073***	-0.073***	-0.079***	-0.079***	-0.088***	-0.058***
	(0.017)	(0.019)	(0.018)	(0.019)	(0.019)	(0.018)	(0.017)	(0.020)	(0.0210)
Other	-0.034***	-0.037***	-0.031**	-0.032***	-0.033***	-0.031***	-0.032***	-0.038***	-0.031*
	(0.011)	(0.012)	(0.015)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.016)
Household wealth level									
Poorest	-0.281***	-0.291***	-0.291***	-0.282***	-0.287***	-0.291***	-0.294***	-0.293***	-0.291***
	(0.017)	(0.020)	(0.023)	(0.017)	(0.017)	(0.016)	(0.016)	(0.021)	(0.023)
Poorer	-0.193***	-0.221***	-0.181***	-0.197***	-0.202***	-0.200***	-0.204***	-0.228***	-0.185***
	(0.014)	(0.018)	(0.017)	(0.014)	(0.015)	(0.014)	(0.014)	(0.019)	(0.017)
Middle	-0.114***	-0.155***	-0.090***	-0.118***	-0.123***	-0.121***	-0.124***	-0.160***	-0.096***
	(0.013)	(0.018)	(0.017)	(0.014)	(0.014)	(0.013)	(0.013)	(0.018)	(0.018)
Richer	-0.056***	-0.095***	-0.035***	-0.058***	-0.062***	-0.060***	-0.062***	-0.100***	-0.038***
	(0.012)	(0.018)	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.018)	(0.014)
Household head's education level									
Primary	0.064***	0.056***	0.072***	0.069***	0.069***	0.065***	0.065***	0.061***	0.077***
	(0.010)	(0.012)	(0.014)	(0.020)	(0.010)	(0.010)	(0.010)	(0.012)	(0.013)
Secondary	0.148***	0.143***	0.158***	0.152***	0.153***	0.149***	0.149***	0.146***	0.160***
	(0.011)	(0.014)	(0.015)	(0.012)	(0.012)	(0.011)	(0.011)	(0.015)	(0.015)
Tertiary	0.102***	0.238***	0.061	0.104***	0.106***	0.104***	0.105***	0.245***	0.063
	(0.028)	(0.025)	(0.042)	(0.027)	(0.027)	(0.028)	(0.028)	(0.025)	(0.041)
School considered essential?	0.008	0.004	0.013	0.016	0.015	0.009	0.010	0.010	0.020
	(0.013)	(0.013)	(0.017)	(0.012)	(0.011)	(0.013)	(0.013)	(0.012)	(0.016)

Rural	-0.018	-0.015	-0.023	-0.021	0.002	-0.018	0.448**	-0.020	-0.025*
	(0.012)	(0.013)	(0.014)	(0.013)	(0.024)	(0.012)	(0.217)	(0.013)	(0.015)
Distance to school	-0.014	-0.027**	0.002	-0.030***	-0.020			0.001	0.002
	(0.011)	(0.011)	(0.013)	(0.010)	(0.012)			(0.012)	(0.018)
(ln) Schools per 5–14-year-olds	0.145***	0.155***	0.133**			0.165***	0.111***		
	(0.043)	(0.042)	(0.052)			(0.038)	(0.041)		
Rural*Distance					-0.017		0.078**		
					(0.013)		(0.036)		
Distance*Worked								-0.062***	-0.020
								(0.012)	(0.025)
Observations	26,673	14,003	12,670	26,673	26,673	26,673	26,673	14,003	12,670

Source: Author's estimations.

Appendix B: Estimates of random effects for null model, by estimation method

Intercept variance estimate	Estimation method used				
	MQL1	MQL2	PQL1	PQL2	MLE
Commune	0.323	0.326	0.390	0.573	0.688
Household	0.387	0.399	0.701	1.223	2.020