Leading education and social research
Institute of Education
University of London

Department of Quantitative Social Science

Teacher Quality in Sub-Saharan Africa: Pupil fixed effects estimates for twelve countries

Christopher F. Hein
Rebecca Allen

DoQSS Working Paper No. 13-08
May 2013

Disclaimer

Department of Quantitative Social Science, Institute of Education, University of London
20 Bedford Way, London WC1H 0AL, UK

Teacher Quality in Sub-Saharan Africa: Pupil-fixed effects estimates for twelve countries

Christopher F. Hein[1] and Rebecca Allen[2]

Abstract

This paper estimates the relationship between teacher characteristics and teacher quality by applying point-in-time pupil-fixed effects. It uses a large cross-sectional dataset of grade 6 teaching and learning in 12 sub-Saharan countries. The findings are generally in line with the existing literature that finds such observable characteristics to be weak predictors that significantly differ in their effects across countries. Teacher subject competency test scores, the only consistent predictor of teacher quality across African countries in other studies, are only significant in the Seychelles. Contrary to US studies, we do not find consistent returns to teacher experience. Our estimates suggest that teacher characteristics are complementary rather than mutually exclusive. The analysis presented here provides comparable estimates of within-school variation of teacher quality and unique lower-bound estimates of teaching ability.

JEL classification: I20, I21, I25, O15, O55, O57

Keywords: Teacher Quality, Teacher Characteristics, Africa, SACMEQ, Complementarity

---

[1] Department of Quantitative Social Science, Institute of Education, University of London (chein@ioe.ac.uk)

[2] Department of Quantitative Social Science, Institute of Education, University of London (r.allen@ioe.ac.uk)

The authors would like to thank Kirstine Hansen and Tarek Mostafa for their feedback.

Introduction

There is common agreement that teachers matter in respect to pupil learning outcomes, and that there is substantial variation in estimated teacher quality in many countries. However, estimating the impact of teacher characteristics on pupil attainment is never straightforward because teachers are not randomly allocated to schools and classrooms and many pupil characteristics that determine attainment are not observed. High quality datasets from developed countries such as the US, with multiple observations of both teachers and pupils, have produced convincing estimates that have made a substantial contribution to our understanding of teacher labour markets. However, few developing countries possess even basic teacher-pupil matched data and, given it is reasonable to assume teacher labour markets in developed and developing countries are very different, we are restricted in the extent to which we can use US findings to inform policy elsewhere. This paper contributes to this gap by estimating within school variation in teacher quality and its conditional correlation with teacher characteristics across 12 sub-Saharan countries. The methods we use require relatively few identification assumptions because most pupils are observed in maths and English lessons with a different teacher. So, we can hold constant pupil characteristics to relate differences in attainment across subjects with differences in teacher characteristics.

The typical economic approach to estimating teacher effectiveness is to apply an education production function that adapts Theory of the Firm to the educational context (cf. Hanushek, 1979). Evidence from the developed economy context is predominantly from the US, where there is access to rich longitudinal datasets that link teachers to pupils and annual measures of pupil scores, which purport to reveal a pupil's progress made while in a particular teacher's classroom. This body of literature focuses on estimating the total impact of teachers on pupil test scores within a school, by using teacher fixed effects to estimate the total variance of pupil test scores due to the respective teacher. These are expressed in standard deviations to allow comparability across studies. Clotfelter et al. (2007), for example, use administrative data from North Carolina consisting of 1.8 million pupils in grades three to five over a period from 1994/95 to 2003/04 to estimate that the contribution of teacher quality variation to variation in pupil achievement is 0.15 standard deviations on maths and 0.08 standard deviations on reading scores. Nye et al. (2004) use data from the Tennessee STAR randomized experiment consisting of an eight-year panel of 79 schools in 42 districts, estimating an effect of between 0.065 and 0.1 standard deviations on reading and an effect of between 0.104 and 0.135 standard deviations on maths achievement. Kane et al. (2006) use panel data from New York City from 1998/99 to 2004/05 following pupils in grades four to eight to find teacher fixed effects of 0.06 standard deviations on reading scores in middle school and 0.1 standard deviations in primary school; the effect being stronger on maths scores (0.13 standard deviations in elementary school and 0.08 standard deviations in middle school). Rockoff (2004) uses panel data following 10,000 pupils for five grades and 300 teachers from two districts in New Jersey over a time span of ten years from 1989/90 to 1999/2000, estimating teacher fixed effects ranging between 0.08 and 0.11 standard deviations. Rivkin et al. (2005) and Hanushek (2005) use data from the Texas Lone Star panel following pupils from grade three to eight from 1995/96 to 2000/01, but can only match pupils to the average maths teacher by grade. They estimate an effect of 0.15 standard deviations, which is similar to the findings of Aaronson et al (2007) who estimate maths teacher effectiveness ranging between 0.13 and 0.15 standard deviations, using High School data from Chicago over three school years. In the UK, Slater et al (2011) find teacher fixed effects of between 0.167 and 0.189 standard deviations in maths, science and English GCSE scores in a sample of 740 teachers in 33 schools in England matched to a total 7305 pupils from two points in time.

So while the literature described above consistently shows that variation in teacher quality is substantial, identifying observable characteristics of teachers that consistently predict teacher quality is more difficult. In part this is because teacher characteristics are often poorly measured, but more seriously teachers are not distributed randomly across schools in respect to teacher quality, so that strong and poorly performing teachers cluster in schools (cf. Clotfelter, 2006). If well performing teachers are paired with good pupils, too much of the pupils' achievement would be attributed to the teacher instead of the unobserved of the pupil (Clotfelter et al, 2007), i.e. the teacher variables are then subject to endogeneity.

To illustrate the problem in teacher-pupil matched data, suppose the following equation is estimated:

$$A_{ijkt} = A_{ijkt-1} + \beta H_i + \gamma H_{it} + \tau T_j + \rho S_k + \varepsilon_{ijkt} \qquad (1)$$

where achievement A of pupil i in year or grade t, in school k and with teacher j is the sum of prior achievement, a vector of pupil time-variant and time-invariant characteristics H, a vector of teacher observable characteristics T and a vector of school level inputs S.

A number of approaches have been proposed to gaining unbiased estimates of teacher characteristics, τ. These include directly modelling the non-random matching of pupils and teachers that results from parental school choice (Vegas and De Laat, 2003) or propensity scores as a control variable (Bourdon et al., 2005). 'Value-added' estimation approaches include a measure of prior achievement, thus claiming to simultaneously control for ability as well as prior motivation of the pupil, family involvement and prior knowledge. If prior attainment is omitted, the teacher effect will be estimated for pupil i's education up to the point of measurement. Thus a fixed effect of τ will be biased upward, due to other teachers having contributed to the education of pupil i in the past.

Conversely, estimates of teacher observable characteristics $\tau_j \in T_j$ will be measured with error as they include the effect of other teachers. The nature of this measurement error may also be systematic; if teachers of similar quality group together in schools, the more similar their effectiveness, the harder it is to isolate the impact of one particular teacher.

As an alternative to the estimation of equation (1), longitudinal data allows construction of a 'differences' model with prior attainment replaced by lagged achievement or by achievement in another subject.

$$\Delta A_{ijkt} = \Delta \beta H_i + \Delta \gamma H_{it} + \Delta \tau T_j + \Delta \rho S_k + \Delta \varepsilon_{ijkt} \qquad (2)$$

Following the basic random-effects approach, set out in (1) and (2), there are two branches of estimation. Rockoff (2004), Clotfelter et al (2007) and Aaronson et al (2007) replace the vector of pupil characteristics with a pupil-fixed effect, exploiting variation in experience within the same pupil. This method though implicitly assumes that τ is constant for all T_j . Metzeler and Woessmann (2012) relax this assumption applying a correlated random effects model, via seemingly unrelated regressions. In contrast, Rivkin et al (2005), Hanushek (2005) and Nye et al (2004) apply teacher fixed effects, which exploit teachers' exposure to a variation in pupil and school characteristics. The latter has the advantages of increasing sample size and allowing for pupil confounders. Other approaches are modifications or hybrids of these two approaches (e.g. Slater et al., 2011; Aaronson et al., 2007).

Existing empirical estimates of the impact of teacher characteristics using the models described above are not consistent across countries and contexts. In developing countries, teachers' subject-specific knowledge is often proxied through tests administered to the teachers. Metzeler and Woessmann (2012) for example, find that in Peru teacher test scores account for 6 to 8.7 percent of a standard deviation in pupil maths achievement, but are not significant for pupil reading achievement. Such test are not administered in the US, thus teachers' subject-specific knowledge is proxied, for example, by teaching qualifications or passing scores on licensure exams for both of which Hanushek (2005) find no significant effect. Aaronson et al. (2005) also find no significant impact of academic and professional qualifications apart from teachers holding a certificate in bilingual education, which is significant at the ten percent level and reduces the teacher fixed effect by 0.08[3] standard deviations on pupil maths achievement. In respect to teaching experience, Rivkin et al (2005) and Hanushek (2005) argue that gains to experience occur at the beginning of the

3 Computed from data provided in Aaronson et al (2005)

teaching career. In particular, Rivkin et al (2005) find no further experience gains after five years. Rockoff (2004) finds an impact of 0.15 to 0.18 standard deviations for ten years of experience on pupil reading comprehension tests, but unlike Rivkin et al (2005) in Rockoff's data the experience effect is linear, monotonically increasing with experience. In contrast, Rockoff finds no experience effect after approximately 8 years on maths computation test scores; on vocabulary and maths computation test scores the experience effect increases to a maximum in the first 6 years before decreasing slightly. By contrast, Clotfelter, Ladd and Vigdor (2007) find a substantial, statistically significant impact of teaching experience ranging from 0.057 to 0.118 standard deviations in maths, and between 0.032 and 0.092 standard deviations in reading. Further, they find significant impacts for teachers holding graduate degrees (rather than bachelors' only) and for teachers holding state licensure.

While the US literature finds either very low or no indicative value of teacher characteristics for teacher quality, in the developing economies of Sub-Saharan southern and eastern Africa this may be different for several reasons. First, due to financial constraints, access to other learning-media apart from those provided in schools is potentially more restricted. Thus the impact of teachers might be stronger. Second, as education systems expand to meet the Millennium Development Goals, large numbers of new, less qualified teachers in terms of years of schooling received and duration of teacher training have been recruited. Further, due to financial constraints of developing economies, continual professional development is not available to all, so those teachers receiving it may outperform their peers who do not. Finally, if initial teacher training is less comprehensive in a developing country, improvements due to teaching experience might occur throughout the teaching career or for a longer period of time than in the US.

Few African datasets of teachers are available, most notable of which are the Southern African Consortium for the Monitoring of Educational Quality (SACMEQ) datasets, from mainly English speaking sub-Saharan East African countries, and Program for the analysis of education systems of states and member-governments of the Conference of Ministers of Education of Countries sharing the French language (PASEC) datasets from French speaking West African countries, both of which are repeated cross-sectional data collections. Due to data limitations, the predominant methods of estimation of teacher effectiveness applied in these data are simple OLS (cf. Vegas and De Laat, 2003; Bourdon et al, 2005) and hierarchical/multilevel random effects models.

In Africa, the effects of observable teacher characteristics vary by which countries have been sampled. For example, Bourdon et al (2005) find that contractual teachers do not differ significantly from regular civil servant teachers in Niger. By contrast, Vegas and De Laat (2003) find that contractual teachers gain 0.6 percentage points fewer correct answers on subject competency tests than regular teachers in Togo. Vegas and De Laat also find a negative impact of teacher experience, but at a less significant level.

Fehrler et al (2008) conduct a cross-country analysis of 21 countries using SACMEQ II, collected in 2002/2 and PASEC I data between 1996 and 2000. They find that the 5-point categorical variable indicating academic qualifications from primary to tertiary, is not significant in the PASEC data, but has an effect on pupil test scores of between 0.018 and 0.041 standard deviations in SACMEQ II. Furthermore, they find that teacher subject competency tests have an impact of between 0.21 and 0.32 standard deviations in SACMEQ countries, but are not significant in PASEC countries. The reason for this could be due to the nature of the tests used in the different data sets. Whereas SACMEQ test both teachers and pupils on similar subject competency tests, PASEC test teachers with a fictitious pupil's French dictation and the teachers are required to identify the mistakes.

Zuze (2010) combines SACMEQ II and TIMSS data for Botswana and finds that both teacher academic education and teacher qualification are not significant in SACMEQ and none of the teacher characteristics are significant in the TIMSS data. The only highly significant predictor of pupil achievement in the SACMEQ data is the teacher subject competency test.

Method

In general, education, proxied by achievement *A* of pupil *i* in class *j* and school *k* can be assumed to be a function of innate ability α, family background *H*, school level factors *S* and of the teacher *T*.

$$A_{ijk} = f(\alpha_i, H_i, S_k, T_{jk}) \qquad (3)$$

Suppose we impose the assumption that the effectiveness of the teacher is a summative function of (i) his or her observable subject-specific competency *W,* which is a function of his or her latent subject-matter competency *c,* which in turn is an indirect effect of his or her academic qualifications *q*; and (ii) his or her pedagogic competency ϕ which is a function of his or her teaching experience *x* and teacher training π, this can be formalised as:

$$T_{jk} = W_j\,(c_j\,(q_j)) + \phi_j\,(x_j, \pi_j) \qquad (4)$$

If these assumptions hold, observable characteristics of teachers, such as teaching experience, duration of teacher training, teacher test scores and academic qualifications, should predict teacher quality, thus providing useful guiding evidence for effective policies.

Our estimation approach follows Clotfelter et al. (2007) who argue the teacher-pupil matching problem can be solved by introducing pupil fixed effects capturing both the explained and unexplained effect of the pupil, thus deflating the teacher estimate. Specifically, we apply a variation of the Slater et al (2011) approach, where the achievement A of pupil i in subject $j=\{1 \circ 2\}$ being either maths or reading equates the sum of the vector of the teacher's observable characteristics T_j, and a pupil-fixed effect π. ε represents the error term and τ the respective coefficient:

$$A_{ij} = \tau_j T_{ij} + \pi_i + \varepsilon_{ij} \quad (5),$$

which is subject to the assumptions that

$$E(\varepsilon_{ij} | T_{ij}, \pi_i) = 0 \quad ,$$

$$Var(\Delta\varepsilon_{ij} | T_{ij}) = \sigma^2, j=\{1 \circ 2\} \quad ,$$

$$Cov(\Delta\varepsilon_{ij}, \Delta\varepsilon_{im} | T_{ij}) = 0, j \neq m \quad .$$

Since there are only two subjects per pupil, an inclusion of a pupil fixed effect is equivalent to computing within-pupil differences:

$$A_{i1} - A_{i2} = \tau_1 T_{i1} - \tau_2 T_{i2} + \varepsilon_{i1} - \varepsilon_{i2} \quad (6).$$

Thus each variable and error that is the same in both equations, one for each subject, is eliminated. In an econometric context, this approach nets out all confounders at the individual and school level with only differences at the classroom level remaining. Such confounders will be differences in classroom resources, peer effect and teaching ability for different subsamples within a class. In respect to differences in classroom equipment, such as insufficient amounts of teaching resources, these are unlikely to bias findings for two reasons. Firstly, distributing classroom equipment are strategic decisions made by the school, which due to the method of estimation applied drops out, as the pupil is in the same school for both outcomes A_{ij}. On the other hand, how teachers deal with the teaching resources available lies at the core of what teachers do and thus is a central part of an estimate of a teacher's effectiveness. Similarly, teachers can use pupils as a resource to shape an individual's learning, thus can strategically profit from the peers available to them. Although SACMEQ II provides information on class size, which could function as an indicator of the extent to which potential influential peers might be omitted, this is problematic for two reasons. First, assuming teachers strategically use the peers available to them for their teaching, including class size as a proxying control for a differential peer effect would capture another valuable dimension of the teacher effect thus resulting in a downward bias of the teacher estimate. Second, estimates are only confounded if 'central' peers - i.e. peers an individual interacts with frequently, such as the person the individual sits next to in class - in the respective individual's network are not in the same class for one of the observations ij. SACMEQ II does not provide data on peer networks but does provide information on the composition of the sampled classrooms. Although this allows generating a dummy control variable indicating whether the classroom composition is the same for both subjects, this variable is dropped due to collinearity because of the differences model induced through the pupil-fixed effects model. Finally, we include a variable indicating 'maths' as a control variable due to systematic variation in test difficulty. Further, acknowledging the vast research exploring the impacts of non-school related factors on pupil achievement since the Coleman Report in 1966, we include interactions of pupil gender, pupil socio-economic status, average school distance with 'maths' allowing for gender differences in pupils' maths achievement, allowing for maths achievement to vary by socio-economic status and geographical location (cf. Heyneman, 1976; Saito, 1998; Taylor and Yu, 2009 to mention a few focussing on the African context).

Remaining biases

Sample size

In a random effects approach, Kane and Staiger (2002) claim that small samples of pupils in a school lead to imprecise estimates, because the confidence intervals become wider the fewer observations there are. In a pupil-fixed effects model there is a similar problem in respect to the number of observations, in this case the number of outcome measures. Here there are two outcome measures, one for each subject, on which the pupil-fixed effect is estimated. Two observations is a small number to estimate a true pupil-fixed effect, thus it makes sense to assume that the pupil-fixed effect π_i is estimated with error. In consequence, the remaining estimates T_i, X_{ij} and the error will be biased down, as too much of the total variance will be attributed to the pupil level. Hence, this model runs the risk of producing type II errors.

Staff-composition

Another source of bias stems from the staff composition of each school. As already described, a pupil-fixed effect model computes differences of $\tau_{i1} \in T_{i1}$ and $\tau_{i2} \in T_{i2}$ so that if $\tau_{i1} = \tau_{i2}$ the variable is omitted for observation *i* due to collinearity. Thus, estimates are only obtained for schools in which $\tau_{i1} \neq \tau_{i2}$. This could lead to a type II error due to the composition of the respective school; Assuming clustering of teachers of similar quality in schools, it is likely that these teachers have similar characteristics, for example duration of teacher training, receiving in-service training or a teacher's academic education. Similarly, if an education system is dominated by one gender, the estimate for a gender effect could be limited to a very small amount of schools that have teachers of both genders. Due to this school-staff-composition related endogenous bias, the findings presented here cannot claim to answer whether, for example, in-service training is effective in the entire education system – this can only be achieved through random assignment of teachers to schools. Instead this design claims to estimate teacher quality as the average impact of a difference in teachers' observable characteristics on a difference in pupil test scores within a school.

Two typical econometric approaches to correcting for selection of teachers to schools are to model a selection equation or to apply a propensity score model. Both approaches model the likelihood of an individual, i.e. teacher, being in a certain group, i.e. school. The explanatory variables being, for example, the observable characteristics of the individual, teaching values, teaching practices, teaching goals, etc. This approach gives rise to a number of problems. First, in a cross-sectional design, the variables listed above are collected at the time the individual is working in a particular school. Therefore variables indicating teaching goals, teaching practices and teaching values, as well as socio-economic status of the teacher are likely to violate the assumption of conditional independence of the predictors, as these are likely to be the result of working in the particular school. A longitudinal design could circumvent this issue, if applying a selection approach to model transitions of teachers from one school to the next. Second, given a pupil-fixed effects model, if this were to be combined with one of the selection approaches named above, then including a selection correction term may not have the wished effect, due to the problem outlined above: assuming teachers of similar quality have similar characteristics, then the likelihood of teachers being at the same school will be similar. Thus the differences model induced by the pupil-fixed effect may not sufficiently correct for selection. Further, the questions relating to eliciting teacher's teaching values, goals and practices, etc, need to be designed in a way that obtains sufficient variation in the data. In SACMEQ II data, the quality of such variables is poor for two reasons. First, the respective variables are based on three-level Likert items corresponding to 1 "Not very important", 2 "Of some importance" and 3 "Very important". Second, the questions themselves, to the eye of an educator, do not provoke variation in the data. For example, in respect to the goals of reading, the following have to be scored from 1 to 3 according to their importance:

- Making reading enjoyable
- Extending pupils' vocabulary
- Improving word attack skills
- Improving pupils' reading comprehension
- Developing a lasting interest in reading

It does not surprise that these items create very little variation, as these are all core goals of learning to read.

Aslam and Kingdon (2011) mention an additional source of endogenous bias that can occur if unobserved teacher behaviour is correlated with both the outcome and the observable teacher characteristics. For example, teachers with more training could have higher motivation that could be proxied by planning lessons in advance or giving regular homework. As argued above, there are both theoretical and data quality issues that prevent an inclusion of variables "delving into the black-box of teaching" (Aslam and Kingdon, 2011, p. 559).

Attenuation

There are two sources for bias due to attenuation. First, teacher observable characteristics are potentially subject to random measurement error ($\tau^{hat}_{ij} = \tau_{ij} + \omega_{ij}$, $\tau_{ij} \in T_{ij}$) Adapting (6) illustrates that in such cases estimates of τ^{hat}_{ij} will be underestimated.

$$\Delta A_{ij} = \tau_1 \Delta T_{ij} + \tau_2 \Delta \omega_{ij} + \Delta \varepsilon_{ij} \qquad (7)$$

Common measurement error methods such as Regression Calibration or Simulation Extrapolation depend on multiple covariates of the "true" variable, for example repetitive measurements of the respective variable, in order to estimate a latent, error free variable that is then included in the estimation. In this particular case, the only available repetitive measures in the data are already being used to compute the differences induced by the fixed-effects model, so that none of the common measurement error models can be applied. Again, this source of attenuation bias could lead to a type-II error.

Attenuation bias can also exist when estimating the within-school variation of total teacher effectiveness, due to sampling error. Aaronson et al (2007) argue that such sampling error can be accounted for approximately by subtracting the estimated average sampling variance from the estimated variance. I therefore compute the adjusted standard deviation by taking the square root of the square of the standard error of the mean of the variable indicating 'subject' and subtract this from the square of the respective estimated standardized coefficient:

$$\text{Adjusted S.D.}_j = \sqrt{(\text{Beta}_j^2 - \text{SEM}_j^2)}$$

Data

SACMEQ II is the second survey of grade 6 pupils in primary schools to be conducted by SACMEQ together with UNESCO's International Institute of Educational Planning (IIEP), with data collection taking place between 2000 and 2002. The original sample of SACMEQ II contains teacher-pupil matches from fifteen countries. Sub-samples from Mauritius and South Africa are dropped because these countries do not conduct teacher subject competency tests and Zimbabwe is dropped due to poor quality of the data resulting from an untrustworthy political context. The remaining sample for analysis contains 35,434 pupils across 12 countries, all being educated by 6316 teacher in the final year of primary school (see Table 1 and 2).

Figure 1 illustrates that the teaching professions in each individual country is usually dominated by either the male or female sex; only the samples from Namibia and Tanzania are close to gender-balanced. At the extremes the gender ratio is 7.8 female teachers to one male teacher in the Seychelles and approximately 7 male teachers to one female in Uganda.

Figure 2 summarises the percentages of teachers by subject and gender. This figure clearly illustrates the two predominant teaching policies among these countries: the majority of grade 6 pupils in Botswana, Lesotho and Zambia are instructed by the same teacher in both subjects, whereas in the other nine countries the majority of pupils are instructed by different teachers for mathematics and reading. While in Botswana and Lesotho the proportion of male and female teachers teaching both subjects is similar, the proportion of male reading teachers is higher than the corresponding proportion of female teachers in Lesotho. In Zambia, the proportion of female teachers teaching only one subject is 6.2 percentage points higher; also the proportion of female maths teachers to female reading teachers is approximately 2/3 in contrast to the respective ratio among their male peers being 1/3. In all other countries there is a gender bias towards subjects taught; maths being preferred by male teachers in contrast to reading predominantly being taught by their female peers. This gender bias is most extreme in the Seychelles, where 92 percent of male teachers teach maths compared to 42.3 percent of female teachers. Strong gender gaps also exist in Kenya (24.4 percentage points in maths), Swaziland (16.6 percentage points in maths) and Tanzania (24.8 percentage points in maths).

Figure 3 shows the proportion of teachers by teacher training received and gender. Again there is a tendency of a gender bias towards male teachers being trained less than their female peers. Only in Kenya, Zambia and Zanzibar are the proportions of male and female teachers receiving little or less training similar; in the Seychelles this gender bias is reversed.

Figure 4 illustrates the highest academic qualification obtained of the respective teachers by gender. In most countries, teachers having completed secondary education predominate. Yet, at one extreme, in Lesotho, 56.8 percent of female and 40.3 percent of male teachers have only completed primary education. On the other hand, teachers holding A-level qualifications are predominant in the Seychelles (61.5 percent of female teachers and all male teachers), followed by Swaziland where 63.5 percent of female teachers and 60 percent of male teachers hold A-levels and Uganda, where 34.8 percent of female teachers and 37.3 percent of male teachers hold A-levels. Teachers having completed tertiary education do not exist in the Zanzibar and Malawian sub-samples; in

Mozambique 0.5 percent of female teachers have tertiary education and in Tanzania the respective proportion is 0.9 percent.

The teacher subject competency tests are constructed to be placed on the same scale as the pupil tests, which have a SACMEQ-wide mean of 500 and a standard deviation of 100 points. Thus the average teacher test scores by country depicted in Figure 5 indicate the difference in subject mastery between the teachers in each country and the average SACMEQ pupil. The average reading teacher test score is 733 points compared to 790 points for the maths teacher test scores.

At one extreme, for both teachers of reading and maths, teachers in Zanzibar perform worst compared to all other samples, and outperforming the average pupil by approximately 1.5 standard deviations in reading and 1.8 standard deviations in maths. In contrast Kenyan maths teachers outperform the average pupil by approximately 4.7 standard deviations and teachers in the Seychelles outperform the average pupil in reading by approximately 3 standard deviations.

Tables 3a and 3b report T-tests testing the null hypothesis that the average test scores by subject within each country do not differ. Table 3a considers individuals that only teach one subject and Table 3b considers those individuals that teach both subjects. This reveals that teachers of only one subject differ significantly in Kenya, Mozambique and the Seychelles, and that the null hypothesis is rejected for teachers teaching both subjects in Lesotho, Swaziland and Tanzania.

As an indicator of geographical location of schools, SACMEQ II asks the head teachers to estimate the distance of the school to six items of infrastructure (clinic, tarmac road, library, book shop, secondary school, market). For this research an average of these six items was computed as a more precise indicator of geographical location than the alternative 4-point categorical indicator of urbanness. Figure 6 shows the geographical distribution of teachers by country and gender. The median distance for male teachers is strictly greater in all samples than the corresponding median for their female peers. The upper adjacent values portray both the size of the respective country and a level of development of the respective economy.

Another aspect of geographical variation is the amount of teacher training received (see Figure 7). In both Uganda and Zambia teachers with less than normal training are approximately similarly distributed on average as their peers with having received normal teacher training. To contrast extreme differences in teacher allocation policy, teachers in Zambia having received less than normal training are placed approximately 3.8 times as far as their peers with normal training; in the Seychelles teachers with less than normal training are placed approximately half the distance as their peers with normal training.

Figure 8 shows the percentage of teachers not having received in-service training by country and gender and illustrates the different country's priorities in respect to continual professional development of teachers. The proportions differ substantially between countries with approximately 80 percent of female teachers not having received in-service training in Malawi, Mozambique and Tanzania, compared to 26 percent of female teachers in Botswana.

Finally Figure 9a illustrates the distribution of teaching experience by gender in the sampled countries. While the average amount of teaching experience is 11.22 years there are substantial differences in the range, indicated by the whiskers and upper adjacent values, between countries and between genders within respective countries. Notably, teachers of both gender reach 40 years of teaching experience – the maximum value of this variable. This value is also reached by female teachers in the Seychelles. In contrast the upper adjacent value of their male peers is approximately 7 years. Figure 9b complements the previous with the average amounts of teaching experience by gender. It reveals that female teachers have more experience in Botswana, Lesotho, Malawi, Mozambique, Seychelles, Swaziland and Zambia. In the remaining countries male teachers have more experience with the exception of Namibia where on average both genders have approximately the same amount of experience. To test whether these gender differences are significant, a two-sample independent T-test shows that the null hypothesis of no difference is rejected in all countries except Lesotho, Namibia, Tanzania and Uganda.

In respect to whether effective teachers teach in well performing schools, this is hard to identify with descriptive statistics. To explore this issue, correlations of teacher test scores and average school socio-economic status as well as correlations of teacher test scores and school equipment are computed (not reported here). In most countries these correlations are not or weakly significant (just below $p=.05$) and of low magnitude. Only in Namibia is there a significant correlation of teacher test scores and school equipment of 0.41.

Findings

As shown in Figure 2, the sample consists of twelve countries, in three of which - Botswana, Lesotho and Zambia - the majority of pupils are taught by the same teacher. Due to the chosen pupil-fixed effects model, we are only able to ask whether teacher subject competency scores predict variation in their class average subject attainment. Table 5 shows that teacher test scores do not predict a difference in a teacher's subject-specific teaching ability.

In the remaining countries the majority of pupils are taught by two separate teachers, although there are occasional exceptions to this. The pupil-fixed effects model allows estimation of the average within-school effect of teacher's observable characteristics. The findings in Table 5 support prior research that the explanatory power in both coefficient and level of significance of teacher's observable characteristics in respect to teacher quality differs substantially from one country to another. Notably, none of the observable characteristics predict teacher quality in Lesotho.

Compared to teachers with more than one year of teacher training, teachers with no teacher training do 2 percent of a standard deviation worse in Namibia and 14 percent of a standard deviation worse in Uganda, but 8 percent of a standard deviation better in Botswana. The sample size underlying the latter estimate though is very small, so that it must be interpreted with care. In contrast, teachers with less than one year's teacher training do 4 percent of a standard deviation better than their peers in both Botswana and Namibia. Teachers with no in-service training do 6 percent of a standard deviation worse in Malawi. Teacher test scores account for 6 percent of a standard deviation change in pupil outcome only in the Seychelles.

Compared to teachers with tertiary education, teachers with primary education do not differ significantly from the reference category of teachers with tertiary education in any country apart from Swaziland, where this group performs 5 percent of a standard deviation worse. In Kenya, teachers with secondary education do 9 percent of a standard deviation better and only slightly better than their peers with tertiary education who do 8 percent of a standard deviation better than the comparison group. In Malawi only teachers with secondary education differ significantly, outperforming their peers by 12 percent of a standard deviation. In Mozambique teachers with secondary education also differ significantly from their peers, but perform 12 percent of a standard deviation worse.

Teacher quality also differs by gender in Malawi, Mozambique, Namibia, Tanzania and Uganda, with female teachers outperforming males by 2 to 14 percent of a standard deviation in pupil outcomes in Malawi, Namibia and Tanzania. By contrast, female teachers perform worse than males in Mozambique and Uganda.

Returns to teaching experience

Table 5 reveals that teaching experience is only significant in three countries, accounting for 7, 5 and 43 percent of a standard deviation in Uganda, Zanzibar and Zambia, respectively. The latter is surprisingly high and may be an artefact of the small sample size (N=41) of teachers underlying this estimate. On the other hand, in the US-based literature, experience itself is commonly found to be not significant, yet returns to the initial years (cf. Rivkin et al, 2005; Hanushek, 2005; Rockoff, 2004) are. To test this hypothesis, we replace 'years of teaching experience' in model (5) with dummy variables indicating one, two, three, and four years of teaching experience; the reference category being teachers with five or more years of experience. Corresponding estimates in Table 6 show, in contrary to findings from the US, that there is no consistent growth in teacher effectiveness due to experience. Therefore we interpret these findings as cohort effect, the possible reasons for this being manifold, potentially ranging from different recruitment policies compared to other years to reduced length of school years due to natural crises. Also, the school-staff-composition bias (described above) can be another reason for these findings.

The findings can be classified into three groups. The first group consists of Kenya and Malawi, where returns to teaching experience remain not significant. Then there is Uganda where the cohort with two years of experience outperforms the reference category by 4 percent of a standard deviation, yet the cohort with 3 years of experience performs 8 percent of a standard deviation worse than the reference category, or approximately 12 percent worse than their peers with two years of experience. In the remaining sub-samples there is one cohort that differs significantly from the reference category ranging from outperforming by 6 percent of a standard deviation in Mozambique to underperforming by 8 percent of a standard deviation in Uganda.

Relaxing equation (2): complementarities of teacher characteristics

In equation (2) we imposed the assumption that the effectiveness of the teacher is a summative function of his or her observable subject-specific competency *W,* which is a function of his or her latent subject-matter competency *c,* which in turn is an indirect effect of his or her academic qualifications *q*; and his or her pedagogic competency ϕ which is a function of his or her teaching experience *x* and teacher training π, this can be formalised as:

$$T_{jk} = W_j\,(c_j\,(q_j)) + \phi_j\,(x_j, \pi_j) \qquad (2).$$

Yet, it may be more reasonable to assume that this strict division of observable subject-matter competency and pedagogic competency are complementary to one another instead of mutually exclusive. This can be tested by including interaction terms between variables corresponding to either category. These are interactions of the teacher test score, as the direct proxy measure of observable subject-specific competency and either teaching experience, the teacher training and no in-service dummies representing the teaching ability category. Also the latter category is interacted with the dummy variables indicating highest academic qualification attained. The results of this model are summarised in Table 7, which reports the variance of the main effects corrected by the interactions. Again, there are no consistent findings across all countries and the magnitude of variance of the main effects is similar to those of the baseline model in Table 5. Yet, the number of significant interactions, though different across countries, strongly suggests that teacher observable characteristics are better interpreted as complementary goods than as substitutes and lead to some interesting findings. For example in Malawi, accounting for the interactions, the corrected variance of the teacher test score is 126 percent of a standard deviation. Teacher test scores also account for 47 percent of a standard deviation in the Seychelles. Also in Malawi, the corrected variance of not receiving in-service training accounts for 143 percent of a standard deviation change compared to those having received such training. Also, in Swaziland, teachers not receiving in-service training outperform their peers by 29 percent of a standard deviation. In Mozambique, findings now suggest that teachers having completed secondary education outperform their peers by 46 percent of a standard deviation.[2]

Teacher-fixed effects estimates

Findings in Table 5 show that teacher test scores do not predict a difference in subject-specific teaching ability of a teacher thus suggesting that teachers in these three countries do not differ in their teaching ability between subjects. This hypothesis can be tested by replacing the vector of teacher characteristics T_j, with a teacher fixed-effect ϑ_j so that

$$A_{ij} = \vartheta_j + \pi_i + \varepsilon_{ij} \qquad (8).$$

Table 8 shows that differential teaching ability does exist and that on average differential teaching ability accounts for 4 percent of a standard deviation in Botswana and Zambia, and 3 percent of a standard deviation in Lesotho. The findings obtained from corresponding sub-samples of the remaining countries show much greater variation, ranging from 5 percent of a standard deviation in Swaziland, to 12 percent of a standard deviation in Namibia. Interestingly, there seem to be no significant differences in teaching ability in Kenya, Malawi, Tanzania and Uganda.

Finally, model (8) can also be used to obtain within-school estimates of teacher quality for sub-samples that are instructed by a different teacher in each subject. Again, the teacher-fixed effect dummy indicates subject and thus teacher. Teacher effectiveness varies between 4 percent of a standard deviation in Malawi and 17 percent of a standard deviation in Tanzania. There does not seem to be a significant difference in teacher quality in both Zanzibar and Zambia – although sample size of the Zambian sample is low at N = 196 pupil-teacher matches. Also in Namibia, on average teacher quality within-teacher varies more than within-schools. Compared to the existing literature from the developed economy context, most findings presented here are very similar to those in the US, ranging from approximately 9 to 12 percent of a standard deviation. Malawi is the only low

2 To check robustness of findings, one can run a random effects model that allows controlling for confounding factors at the individual, teacher and school level. Please compare Duthilleul and Allen (2005) for this approach. Their findings are similar.

outlier at approximately 4 percent of a standard deviation, as the upper outliers, Tanzania, Uganda and Lesotho are similar to Slater et al. (2011) in the UK.

A negative estimate indicates that reading teachers outperform their peer maths teachers and vice versa. This suggests that within schools on average reading teachers outperform their maths-teaching peers in Kenya, Malawi, Mozambique and Uganda, whereas on average teachers teaching both subjects are better at teaching maths than reading in all countries with significant estimates.

As the standard errors of the means are very small, the adjusted variances (see above) hardly differ from the unadjusted. Existing literature using pupil-and-teacher-fixed effects usually panel by teacher instead of pupil to obtain an estimate of the teacher-fixed effect, and report much larger differences between the unadjusted and adjusted variance. This is to be expected as they obtain an estimate for every teacher. In contrast, the model applied here panels by pupils, thus estimating only at the mean, where the sampling error is small. Therefore this method provides an elegant, cost-effective way to estimating the within-school and within-teacher variance of teacher quality comparable to literature from the developed economy context.

Conclusion and Policy Implications

This paper explores the explanatory power of observable characteristics of teachers in respect to teacher quality. The findings are generally in line with the existing literature that finds such observable characteristics to be weak predictors that differ significantly across countries in their impact on pupil achievement. Surprisingly, in a pupil-fixed effects model, teacher subject competency test scores, the only consistent predictor of teacher quality across African countries, are only significant in the Seychelles. Following the US literature, this paper further explores the effect of the initial years of teaching experience. Contradictory to the findings in the US there is no clear picture of returns to experience, and positive and negative returns for different years within countries suggest that experience may be best interpreted as cohort effects.

On a theoretical note, the bulk of previous papers test teacher characteristics by treating these as mutually exclusive. This paper provides evidence that the available characteristics either refer to two categories, a teacher's subject-matter competency and pedagogic competency, and that these ought to be seen as complementary to one another. This approach is easily operationalised by introducing a vector of interactions between these two categories. Whilst cautioning that this data is too old to be used as a basis of policy making, this approach yields surprisingly strong predictors of teacher quality: for example in Malawi, teacher test scores account for 1.26 standard deviations of variance in pupil achievement, and teachers not having received in-service training outperform their peers by 1.43 standard deviations.

Finally this paper estimates within-school and within-teacher estimates of teacher quality. The latter provides a lower-bound estimate of teaching ability in different countries – in some countries teachers teaching both subjects do not differ significantly, in others teaching ability ranges from approximately 3 to 12 percent of a standard deviation in pupil achievement. Estimates of within-school variation are comparable to those from the US in eight of the sampled countries and three are similar to those in the UK; only Malawi is an outlier at the lower end with a variance of approximately 4 percent of a standard deviation in pupil achievement.

References

Aaronson, D., Barrow, L. and Sander, W. (2007). Teachers and student achievement in the Chicago public high schools, *Journal of Labour Economics,* 25 (1), 95 - 135.

Bourdon, J., Frölich, M. and Michaelowa, K. (2005). *Broadening Access to Primary Education: Contract Teacher Programs and their Impact on Education Outcomes in Africa – An Econometric Evaluation for the Republic of Niger*. Paper presented at the Annual Conference of the Research Committee Development Economics of the German Economic Association on "Pro-Poor Growth", Kiel Institute of World Economics.

Clotfelter, C. T., Ladd, H. F. and Vigdor, J. L. (2006). *Teacher-student matching and the assessment of teacher effectiveness*. Cambridge, Mass.: National Bureau of Economic Research.

Clotfelter, C. T., Ladd, H. F. and Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects, *Economics of Education Review,* 26, 673 - 682.

Coleman, J. S. (1966). *Equality of educational opportunity*. Washington, D.C.: United States National Office of Education

Duthilleul, Y., Allen, R. (2005). *Which teachers make a difference? Implications for Policy Makers in SACMEQ countries*, International Institute of Educational Planning, UNESCO: Paris.

Fehrler, S., Michaelowa, K. and Wechtler, A. (2009). The Effectiveness of Inputs in Primary Education: Insights from Recent Student Surveys for Sub-Saharan Africa, *Journal of Development Studies,* 45 (9), 1545-1578.

Hanushek, E. A. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions, *The Journal of Human Resources,* 14 (3), 351-388.

Hanushek, E. A. (2005). *The market for teacher quality*, Cambridge, Mass.: National Bureau of Economic Research.

Heyneman, S. P. (1976). Influences on academic achievement: A comparison of results from Uganda and more industrialized societies, *Sociology of Education,* 49, 200-211.

Kane, T. J., Rockoff, J. E. and Staiger, D. O. (2006). *What does certification tell us about teacher effectiveness: evidence from New York City*, NBER Working Paper Series Number 12155.

Kane, T. J. and Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures, *Journal of Economic Perspectives,* 16 (4), 91 - 114.

Nye, B., Konstantopoulos, S. and Hedges, L. V. (2004). How Large Are Teacher Effects?, *Educational Evaluation and Policy Analysis,* 26 (3), 237-257.

Raudenbush, S. W. and Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods (2nd Ed.)*, London: Sage.

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005). Teachers, Schools, and Academic Achievement, *Econometrica,* 73 (2), 417-458.

Rockoff, J. E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data, *The American Economic Review,* 94 (2), 247-252.

Saito, M. (1998). Gender vs. socio-economic status and school location. Differences in Grade 6 reading literacy in five African countries, *Studies in Educational Evaluation,* 24 (3), 249-261.

Slater, H., Davies, N. M. and Burgess, S. (2011). Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England, *Oxford Bulletin of Economics and Statistics*

Taylor, S. and Yu, D. (2009). The importance of socio-economic status on determining educational achievement in South Africa, *Stellenbosch Economic Working Papers,* (01/09).

Vegas, E. and De Laat, J. (2003). *Do differences in teacher contracts affect student performance? Evidence from Togo*, Washington DC: The World Bank.

Zuze, T. L. (2010). *Human Resource Inputs and Educational Outcomes in Botswana's Schools: Evidence from SACMEQ and TIMMS*, Stellenbosch Economic Working Papers, Number 16/10.

Tables

Table 1: Pupils per Country

Country	Freq.	Percent
BOT	3,322	9.38
KEN	3,268	9.22
LES	3,155	8.90
MAL	2,333	6.58
MOZ	3,120	8.81
NAM	5,048	14.25
SEY	1,484	4.19
SWA	3,139	8.86
TAN	2,841	8.02
UGA	2,642	7.46
ZAM	2,568	7.25
ZAN	2,514	7.09
Total	**35,434**	**100.00**

Table 2: Teacher Training Received by Gender

	No training			Little training			More training		
Country	**F**	**M**	**Total**	**F**	**M**	**Total**	**F**	**M**	**Total**
BOT	12	13	25	3	3	6	339	141	480
KEN	3	6	9	5	6	11	268	331	599
LES	12	7	19	10	7	17	177	53	230
MAL	9	17	26	46	96	142	102	107	209
MOZ	40	329	369	74	261	335	475	880	1,355
NAM	16	19	35	23	49	72	583	495	1,078
SEY	2		2	6		6	96	13	109
SWA	3	13	16	7	3	10	330	179	509
TAN	1		1	21	12	33	243	308	551
UGA		13	13	5	16	21	41	231	272
ZAM	4	2	6	6	6	12	281	157	438
ZAN	28	10	38	53	62	115	291	195	486

Table 3a: T-test of Difference of Teacher Test Scores by Subject within Country

Country	Mean(M)	N(M)	Mean(R)	N(R)	T-test (Difference)
BOT	757.38	42	762.81	42	-.11
KEN	902.19	315	751.22	303	3.38***
LES	874.18	17	839.17	17	.27
MAL	793.95	174	759.68	182	1.03
MOZ	848.36	1005	798.43	1034	1.97*
NAM	751.93	579	746.15	579	.30
SEY	879.38	55	806.07	57	5.39***
SWA	849.93	247	787.09	245	1.47
TAN	923.77	274	833.75	287	1.91
UGA	913.13	118	768.15	190	1.70
ZAM	1667.14	19	1570.16	23	.15
ZAN	929.94	298	903.72	314	.58

M = Maths; R = Reading; *p=.05, **p=.01, ***p=.001

Table 3b: T-test of Difference of Teacher Test Scores within same Teachers within Country

Country	Mean(M)	N(M)	Mean(R)	N(R)	T-test (Difference)
BOT	747.67	217	758.46	210	-1.63
KEN	-- no observations due to missing data --				
LES	742.20	123	720.45	110	2.63**
MAL	765.10	7	631.48	5	2.24
MOZ	-- no observations due to missing data --				
NAM	710.15	18	710.81	9	-0.02
SEY	-- too few observations --				
SWA	806.08	23	756.43	19	2.08*
TAN	756.94	9	696.65	6	2.00*
UGA	890.09	5	722.01	4	1.43
ZAM	744.68	203	758.14	212	-1.80
ZAN	694.12	13	569.44	12	1.83

M = Maths; R = Reading; *p=.05, **p=.01, ***p=.001

Table 4: T-test of teaching experience by gender

Country	Mean(F)	N(F)	Mean(M)	N(M)	T-test
BOT	11.85	354	8.45	157	4.87***
KEN	10.85	276	14.74	343	-3.98***
LES	19.74	199	15.83	67	1.86
MAL	9.99	157	7.52	220	3.70***
MOZ	15.55	589	9.91	1470	8.39***
NAM	11.64	622	11.27	563	0.72
SEY	15.99	104	4.06	13	3.67***
SWA	12.19	340	8.13	195	5.08***
TAN	15.53	265	13.70	320	1.84
UGA	8.41	46	8.63	260	-0.12
ZAM	14.86	291	10.31	165	2.21*
ZAN	15.83	372	13.44	267	2.05*

*M = Male; F = Female; *p=.05, **p=.01, ***p=.001*

Table 5: Baseline Model of Teacher Characteristics

	Same Teacher		Different Teachers									
	Test Score	N_i	Test Score	No Teacher Training	1 year or less Training	No In-Service Training	Teaching Experience	Highest Academic Education: Primary	Secondary	A - Levels	Female Teacher	N_i
BOT	X	6115	X	0.08*[1] (2.11)	0.04*[1] (2.40)	X	X	X	X	X	X	448
KEN	[2]	32	X	X	X	X	X	X	0.09** (2.72)	0.08* (2.40)	X	6278
LES	X	6055	X	X	X	X	X	X	X	X	X	244
MAL	X	394	X	X	X	-0.06* (-2.09)	X	X	0.12* (2.12)	X	0.14*** (3.92)	4053
MOZ	[2]	[2]	X	X	X	X	X	X	-0.12* (-2.23)	X	X	5919
NAM	X	689	X	-0.02* (-2.25)	0.04*** (4.97)	X	X	X	X	X	X	9330
SEY	X	218	0.06*** (3.92)	X	X	X	X	X	X	X	-0.03* (-2.39)	2698
SWA	-0.13***[1] (-3.57)	1172	X	X	X	X	X	-0.05* (-2.30)	X	X	X	4941
TAN	X	293	X	X	X	X	X	X	X	X	X	5032
UGA	X	266	X	-0.14*** (-5.24)	X	X	0.07*** (3.98)	X	X	X	-0.05* (-2.02)	4160
ZAM	X	4877	X	X	X	X	0.43**[1] (3.01)	X	X	X	X	206
ZAN	-0.25***[1] (-4.13)	350	X	X	X	x	0.05* (-2.22)	X	X	X	X	4014

X = not significant; [1] = estimated using a very small sample of teachers; standardised coefficients are reported; * p<0.05, ** p<0.01, *** p<0.001; [2] = no observations due to missing data

T-statistic in parentheses

Table 6: Testing for an effect of teaching experience in the initial years

	Years of Experience One	Two	Three	Four	Test Score	No Teacher Training	One year or less Training	No In-Service	Highest Academic Education Primary	Secondary	A - Levels	Female Teacher	N_i
BOT	X	X	-0.14***[1] (-3.44)	X	X	0.14**[1] (3.33)	X	X	X	X	X	X	448
KEN	X	X	X	X	X	X	X	X	X	0.08* (2.29)	0.07* (2.06)	X	6278
LES	0.17*[1] (2.15)	X	X	X	X	X	X	X	X	X	X	X	244
MAL	X	X	X	X	X	X	X	X	X	0.12* (2.19)	X	0.13*** (3.51)	4053
MOZ	X	X	0.06*** (3.99)	X	X	X	X	X	-0.09* (-2.26)	-0.16** (-3.03)	X	X	5919
NAM	X	X	-0.02* (-2.08)	X	X	-0.02** (-2.82)	0.04*** (4.90)	X	X	X	X	X	9330
SEY	X	-0.03* (-2.07)	X	X	0.05** (3.23)	X	X	X	X	X	X	X	2689
SWA	X	X	X	-0.03* (-2.27)	X	X	X	X	X	X	X	X	4941
TAN	X	-0.05** (-3.06)	X	X	X	X	X	X	X	X	X	X	5032
UGA	X	0.04* (2.55)	-0.08*** (-3.52)	X	X	-0.13*** (-5.06)	X	X	0.09*** (3.52)	X	0.13* (2.05)	-0.04* (-1.99)	4160
ZAM	X	X	-0.29**[1] (-3.00)	X	X	X	X	X	X	X	X	X	206
ZAN	0.08** (3.29)	X	X	X	0.06* (2.35)	X	X	X	X	0.05* (2.08)	X	X	4014

X = not significant; [1] = estimated off small sample of teachers; standardised coefficients are reported; * p<0.05, ** p<0.01, *** p<0.001; t-statistic in parentheses

Table 7: Corrected Estimates due to including interactions between Subject Matter Competency and Teaching Ability

	Country											
	BOT[1]	KEN	LES[1]	MAL	MOZ	NAM	SEY	SWA	TAN	UGA	ZAM[1]	ZAN
Test Score	X	X	X	1.26	X	X	0.47	X	X	0.21	9.51[2]	X
No Training	0.04	X	X	X	-0.08	X	X	-0.04	X	X	X	X
Little Training	X	X	X	X	X	0.18	X	X	X	X	X	X
No In-service	X	X	X	1.43	-0.08	X	0.08	0.29	X		X	X
Experience	X	X	X	X	-0.01	X	X	X	X	-0.16	9.51[2]	X
Primary	X	-0.01	X	X	X	X	X	X	X		X	X
Secondary	X	0.09	X	0.12	0.46	X	X	X	-0.26		X	X
A-Levels	X	0.09	X	X	X	X	X	X	X	0.32	X	X
Female Teacher	-0.09	X	X	0.16	X	X	-0.03	X	X	X	X	X
N_i	448	6278	244	4053	5919	9330	2689	4941	5032	4160	206	4014

X = not significant; [1] = estimated off small sample of teachers; [2] = variables linked through interaction; total effects of variables reported that include the effect of interactions

Table 8: Estimates of Variation of Teacher Quality

		Country											
		BOT	KEN	LES	MAL	MOZ	NAM	SEY	SWA	TAN	UGA	ZAM	ZAN
Variation of Teacher Quality	Within-School	0.095*** (0.02)	-0.093*** (0.01)	0.174*** (0.03)	-0.039** (0.01)	-0.108*** (0.01)	0.106*** (0.01)	0.121*** (0.01)	0.109*** (0.01)	0.14*** (0.01)	-0.169*** (0.01)	X	X
	N_i	448	244	6055	4163	6161	9330	2748	4990	5337	4277	212	4322
	Within-Teacher	0.044*** (0.01)	X	0.028** (0.01)	X	[2]	0.122*** (0.02)	0.079* (0.03)	0.05** (0.01)	X	X	0.04*** (0.01)	.096* (0.02)
	N_i	6195	32	6055	421	[2]	689	218	1172	293	266	4887	423
Adjusted Estimates	Within-School	0.095	0.092	0.171	0.038	0.108	0.106	0.121	0.109	0.14	0.169	X	X
	Within-Teacher	0.043	X	0.026	X	[2]	0.12	0.073	0.049	X	X	0.039	.094

X = not significant; [2] = sample size too small; standardised coefficients are reported; Standard Error of the Mean in parentheses; * p<0.05, ** p<0.01, *** p<0.001

Figures

Figure 1: Gender Ratio (Female over Male)

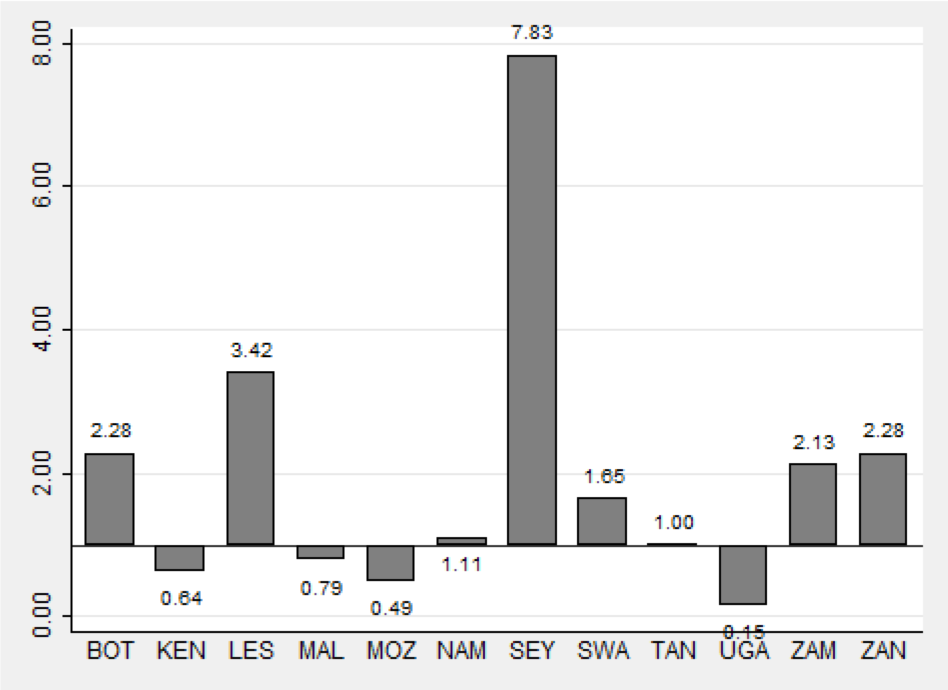


Figure 2: Proportion of Teachers by Subject taught and Gender

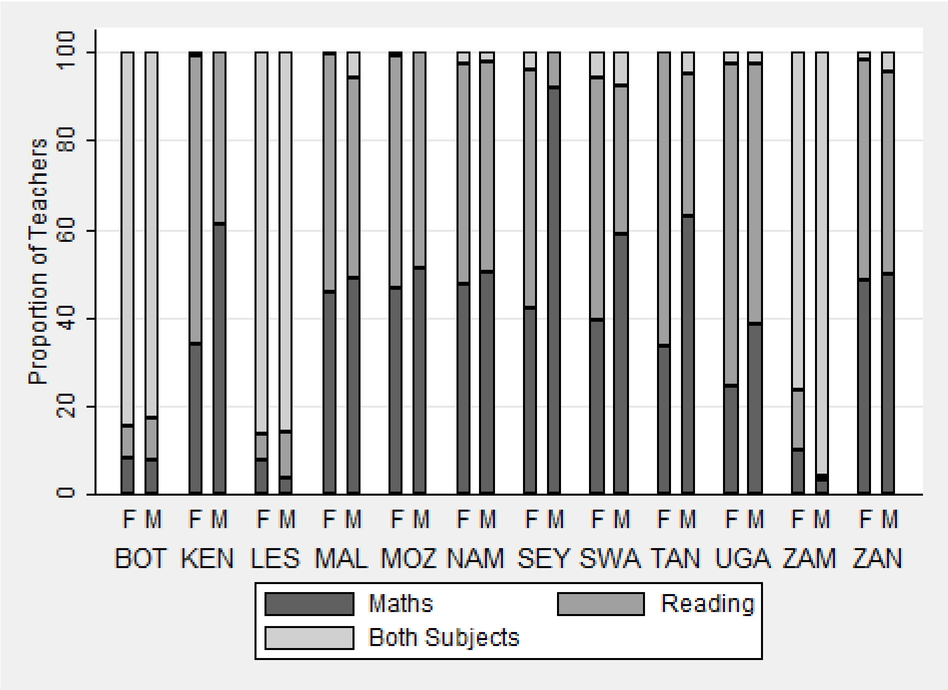


Figure 3: Proportion of Teachers by Teacher Training and Gender

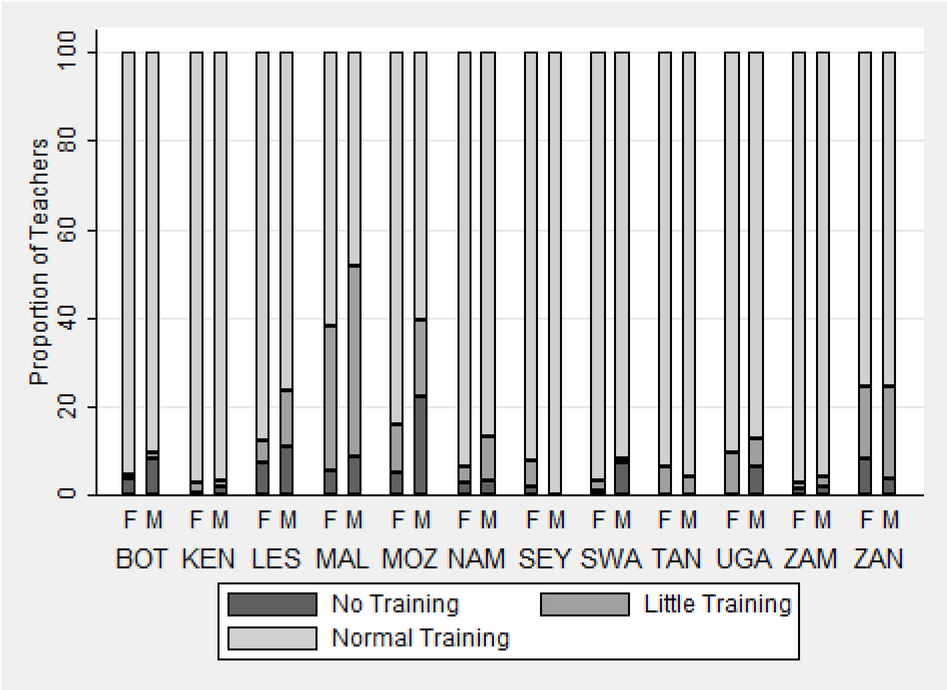


Figure 4: Proportion of Teachers by Highest Academic Qualification and Gender

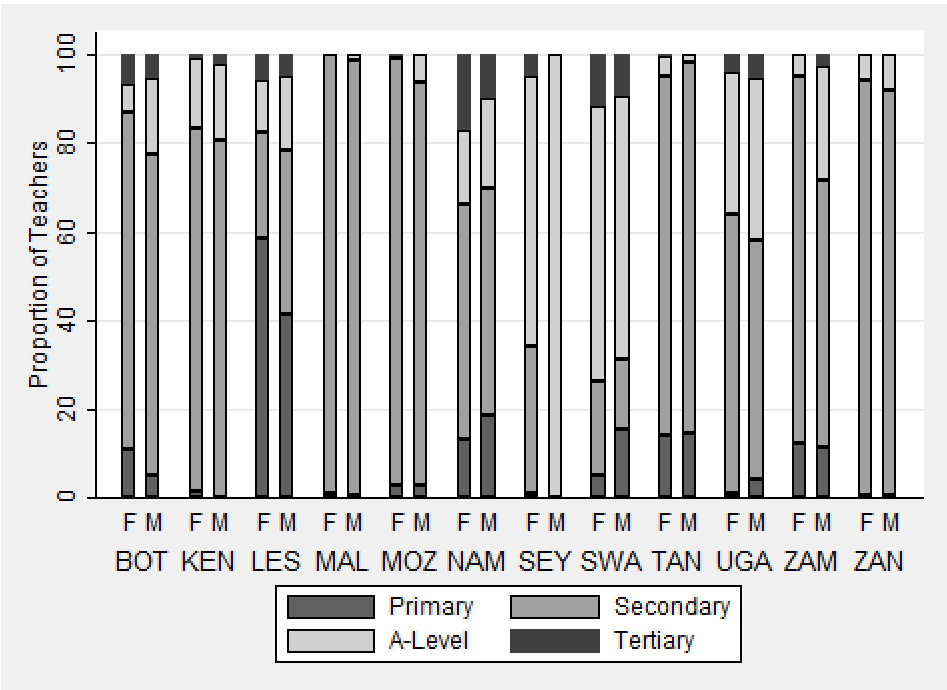
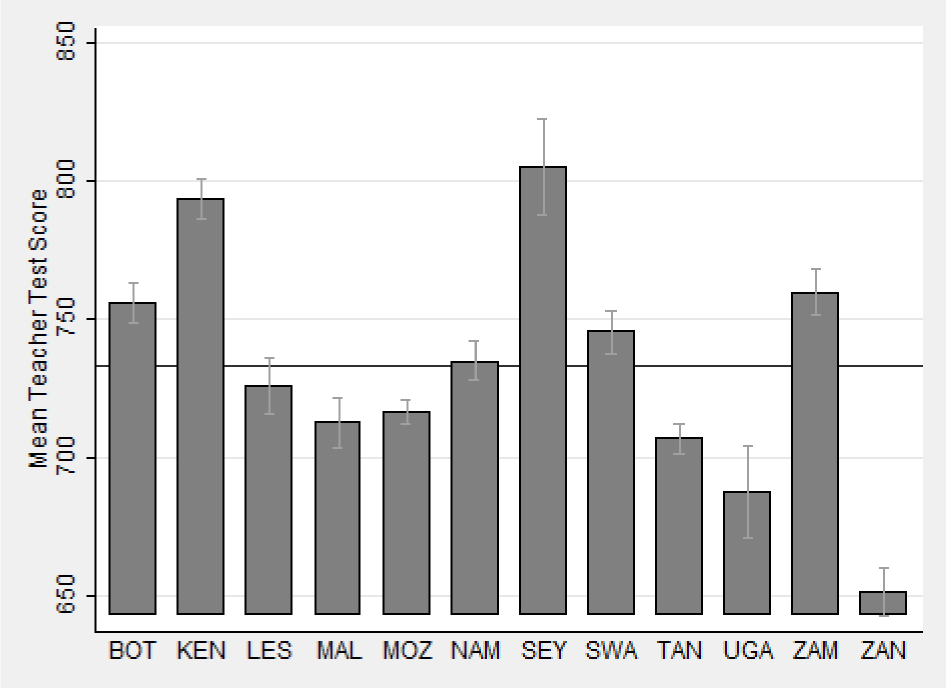
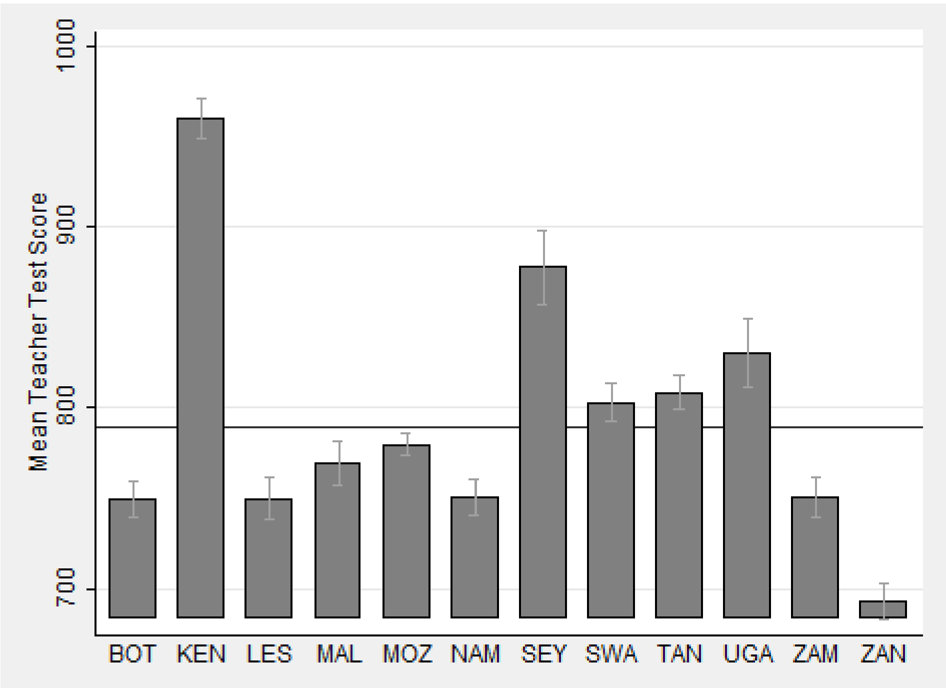


Figure 5a: Teacher Test Score (Reading)



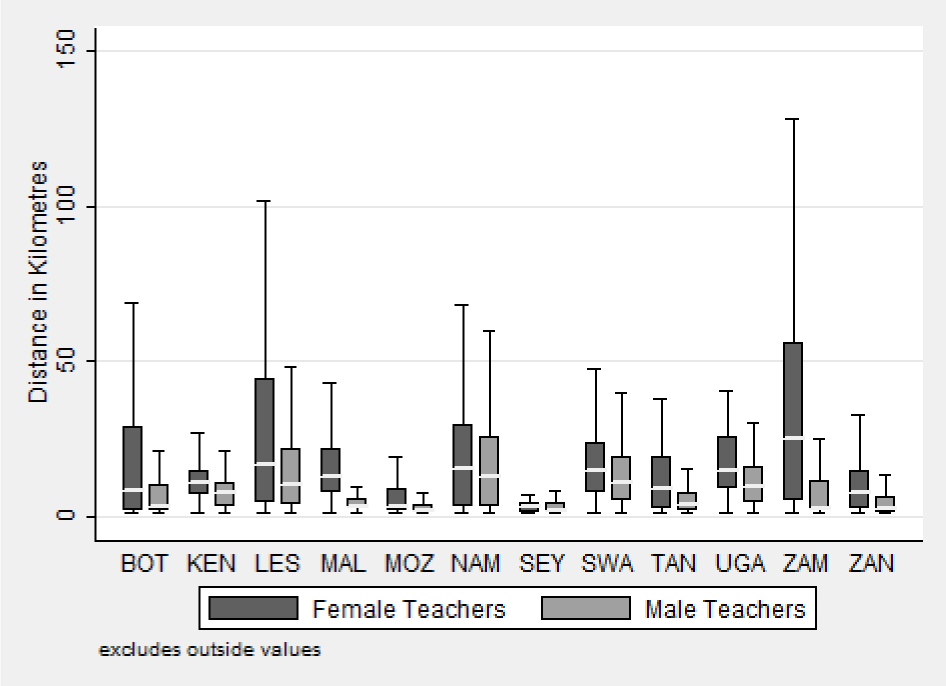
Note: Confidence Intervals at p=.05

Figure 5b Teacher Test Score (Maths)



Note: Confidence Intervals at p=.05

Figure 6: Average School Distance by Gender



Boxes signify IQR; adjacent values are 25th percentile – 1.5 IQR and 75th percentile + 1.5 IQR

Figure 7: Ratio of Average School Distance for "less than normal training" over "normal training"

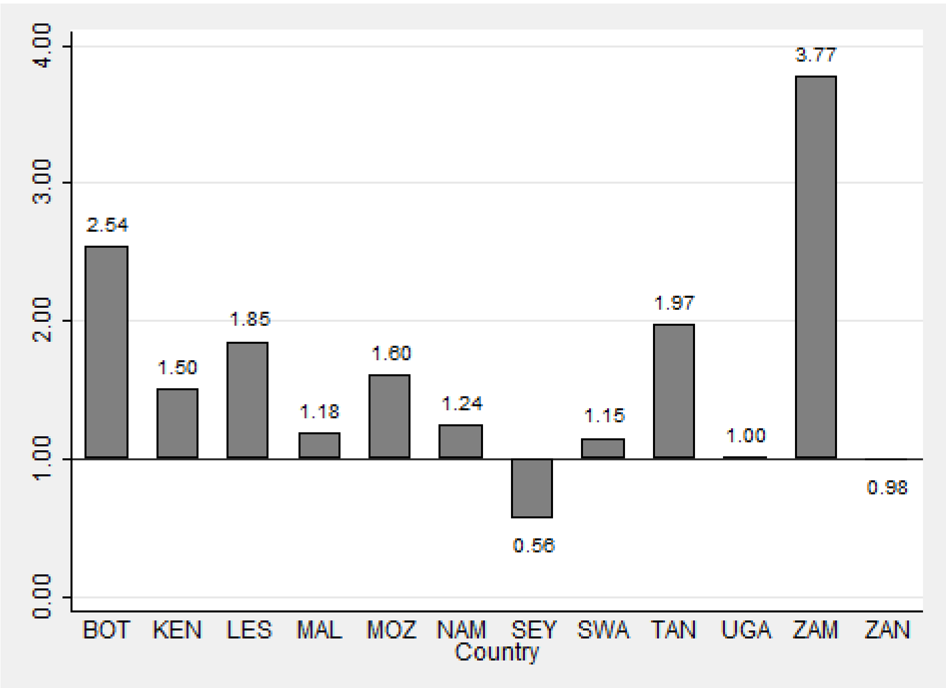


Figure 8: Percentage of Teachers not having received in-service training

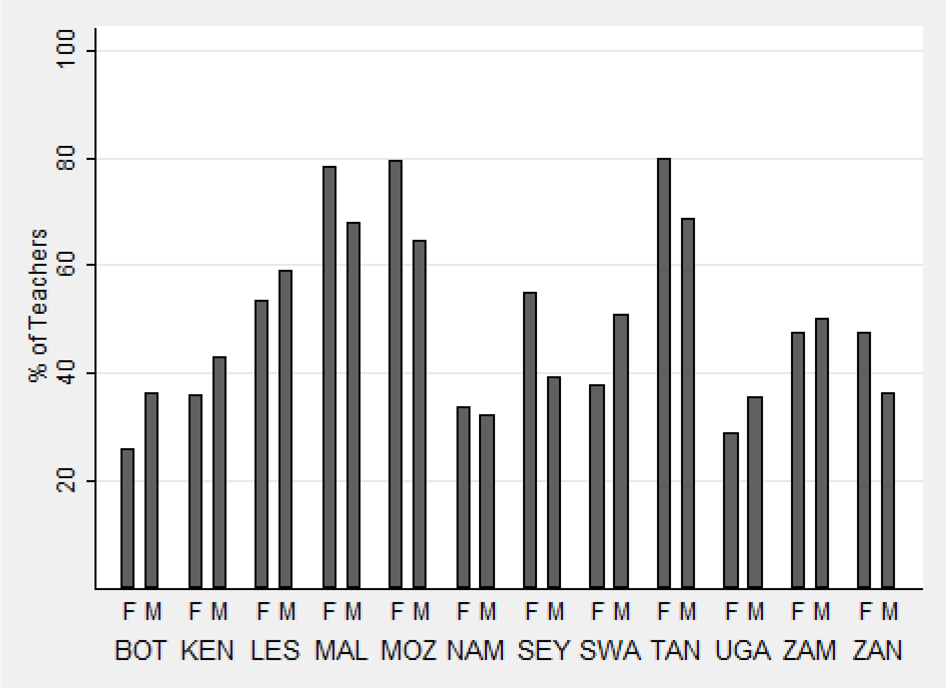
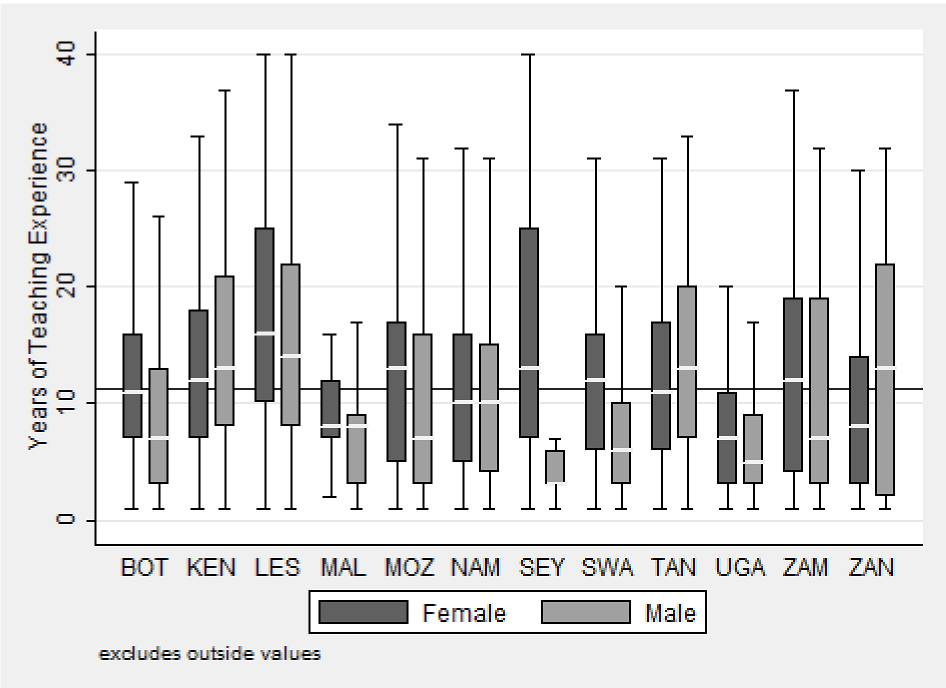


Figure 9a: Teaching Experience by Gender



Boxes signify IQR; adjacent values are 25th percentile – 1.5 IQR and 75th percentile + 1.5 IQR

Figure 9b: Average Teaching Experience by Gender