Association for the Development of Education in Africa

Biennale on Education in Africa
(Libreville, Gabon, March 27-31, 2006)

Effective Schools and Quality Improvement

Parallel Session B-1
Characteristics of Effective Schools

The Cost-Effectiveness of Inputs in Primary Education: Insights from the Literature and Recent Student Surveys for Sub-Saharan Africa

by Katharina MICHAELOWA and Annika WECHTLER

Working Document
DRAFT

PLEASE DO NOT DISSEMINATE

DOC B-1.2

• Original Version in English •

This document was prepared by ADEA for its Biennial Meeting (Libreville, Gabon, March 27-31, 2006). The views and opinions expressed in this volume are those of the authors and should not be attributed to ADEA, to its members or affiliated organizations or to any individual acting on behalf of ADEA.

The document is a working document still in the stages of production. It has been prepared to serve as a basis for discussions at the ADEA Biennial Meeting and should not be disseminated for other purposes at this stage.

Association for the Development of Education in Africa (ADEA)
International Institute for Educational Planning
7-9 rue Eugène Delacroix
75116 Paris, France
Tel.: +33(0)1 45 03 77 57
Fax: +33(0)1 45 03 39 65
adea@iiep.unesco.org
web site: www.ADEAnet.org

Contents

List of Tables

List of Figures

Acknowledgments

The authors wish to thank Jean Bourdon for his thoughtful comments, and for his invaluable help with providing access to some of the relevant literature, as well as with data preparation and discussion. Moreover, they are grateful to Jean-Marc Bernard, Adriaan Verspor and Marta Encinas-Martin for useful suggestions, and to Mioko Saito for her support with the interpretation of some of the SACMEQ variables.

Acronyms and abbreviations

BEPC	Brevet d'études du premier cycle (exam after lower secondary education)
CONFEMEN	Conference of Francophone Ministers of Education
EFA	Education for All
GDP	Gross Domestic Product
GLS	Generalized least squares
MIT	Massachusetts Institute of Technology
MLA	Monitoring Learning Achievement
MLE	Maximum likelihood estimation
OECD	Organization for Economic Co-operation and Development
PASEC	Programme d'analyse des systèmes éducatifs de la CONFEMEN
PIRLS	Progress in International Reading Literacy Study
PISA	Program for International Student Assessment
SACMEQ	Southern and Eastern Africa Consortium for Monitoring of Educational Quality
TIMSS	Third International Mathematics and Science Study
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNESCO-UIS	UNESCO Institute for Statistics
UNICEF	United Nations Children's Fund

1. ABSTRACT

1. The 2003 ADEA biennial meeting on the *Challenge of Learning* already collected considerable evidence for a set of cost-effective "high priority" inputs for primary education systems in sub-Saharan Africa. In this study, this evidence is complemented with empirical results from recent studies based on promising new methodological approaches, such as random evaluations and natural experiments. Moreover, micro data from large-scale student assessments which has recently become available for both anglophone and francophone Africa, is explored to check the robustness of results in the African context. Generally, the outcomes are in line with earlier policy recommendations.

2. In addition, immaterial school inputs such as accountability, effort and motivation are taken into account. Preliminary evidence suggests that in sub-Saharan Africa, institutional reform should focus on the optimization of contract teacher programs, on raising transparency about resource flows and learning outcomes, and on other measures to enhance parents' monitoring capacity.

3. Generally, it seems that educational policy reform should be a mix of an introduction of incentive based mechanisms for teachers and principals (carefully adjusted to the local context), and the provision of basic school inputs, especially learning materials such as textbooks.

2. EXECUTIVE SUMMARY

4. The increasing availability of student survey data, the development of new statistical and econometric methodologies and the expansion of computing capacities has led to a huge increase in scientific evaluations of the determinants of education quality in recent years. Education quality is thereby measured in terms of student achievement on standardized tests, which reflects the cognitive knowledge acquired through the education process. In line with international policy priorities as codified in the Education for All (EFA) objectives and the Dakar Framework for Action, for sub-Saharan Africa, evaluation efforts currently concentrate on the primary level. In addition to various national level evaluations, three programs have been launched on a larger scale: The UNESCO/UNICEF Monitoring Learning Achievement (MLA), the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and the Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC) now jointly cover most countries on the continent.

5. For a general review of results, SACMEQ and PASEC are of particular interest because they use comparable (or identical) tests in all their countries, which allow us to jointly analyze different country cases as well as to draw comparisons across countries. While the comparison of test items and thus a direct comparison of achievement levels across programs is not yet possible, the relationship between inputs and outcomes can be compared.

6. A joint analysis of PASEC and SACMEQ data in a common education production function framework allows us to estimate the impact of educational inputs on student achievement in 22 sub-Saharan African countries, and to compare our results with those of earlier empirical studies for education systems in Africa and other world regions. The analysis is carried out within the context of a general review of the wider literature on determinants of education quality and the cost of educational reform, including insights from industrialized countries as far as they may be of interest for the African context.

7. The discussion is divided in two parts: First the analysis of traditional school inputs such as learning materials, classroom resources, school equipment and the number of (qualified) teachers, and second the analysis of immaterial inputs such as accountability, effort and motivation of various relevant actors within the education system, especially teachers and principals.

8. With respect to traditional school inputs, studies for industrialized countries show a rather bleak picture indicating that there are hardly any promising policy measures which ensure significant achievement gains. For developing countries, in which the initial resource availability is much more limited, more significant effects can be observed, although even here, outcomes of different studies with a different regional or national focus, different methodologies, or different datasets often yield contradictory results. Nevertheless, a consensus is emerging about certain policy priorities to ensure basic education quality at affordable cost.

9. The intensive discussions about education quality before and after the ADEA Biennial Meeting 2003 already led to the definition of certain "essential inputs", and the underlying ranking of policy measures has not changed since then. While new studies using interesting empirical research methods based on natural experiments and randomized evaluations have become available, their outcomes are either in line with earlier results or based on very limited evidence. Moreover, the

outcomes of these studies appear to be similarly inconsistent across studies as prior retrospective analysis.

10. Our own new empirical results based on a traditional retrospective analysis of student achievement in PASEC and SACMEQ countries are broadly in line with earlier analysis. With few exceptions, no major differences can be observed between francophone and anglophone education systems, especially if differences in the sampling methodology are duly taken into account.

11. It turns out that investment in pedagogical resources, especially textbooks for the core subjects of reading and math, can still be considered as an efficient policy measure. If budget constraints are very strong, one book may be provided to every second student, especially in higher grades where taking the book back home does not seem to be as important as for very young students. The analysis of textbook peer effects reinforces this result, but emphasizes the relevance of the distribution of textbooks across schools which must ensure that each school receives a 50% book coverage (rather than 50% of schools being fully equipped while others receive nothing).

12. Obviously, it must be ensured that the books actually reach the students, and that they are effectively used in the classroom. Moreover, teacher manuals have been shown to be helpful complements of textbooks, which may, in fact, increase the chances that textbooks themselves are effectively used.

13. Another priority should be the reduction of repetition rates. Grade repetition blocks considerable resources as it artificially expands the number of students in the education system. At the same time, the frequently stated hypothesis of a positive effect of grade repetition on student learning is clearly rejected by the empirical evidence. At best, students achieve a transitory gain with respect to the rank within their class. At worst, their performance gains are effectively reduced or they drop out from school altogether. Thus reducing repetition rates appears as a cost reducing and yet quality enhancing measure of education policy. The resources set free by reduced repetition rates can be used for other, more promising programs to assist weak students. In India, for instance, local women were successfully employed as teaching assistants in disadvantaged areas with small one-teacher schools, to help these students catch up with their peers. This was shown to be a particularly effective and relatively low-cost measure for students with important learning deficiencies.

14. More differentiation is required when assessing human resources, i.e. teacher numbers and qualification. Human resources are a very expensive input. Higher teacher numbers reduce class size and student-teacher ratios, but this is one of the policies where empirical results based on student surveys have been most inconsistent. At a cross-country level, it can be observed that among the world's best performers in terms of student achievement, are countries in the Nordic region with the smallest average class size world-wide, while others are Asian countries with student-teacher ratios close to those typically found in sub-Saharan Africa.

15. For sub-Saharan Africa, there is some evidence for a negative impact of class size beyond a threshold of approximately 60 students. Below this threshold, the evidence currently available suggests that reducing class size should not be a policy priority.

16. With respect to teacher education and training, the focus should be on quality rather than duration. In anglophone Africa, where the duration of formal education and teachers' subject matter knowledge are much more clearly correlated than in francophone Africa, longer duration of teachers'

academic education significantly enhances student learning. However, the effect is only moderate in size and has to be carefully weighed against the equally high cost generally involved with salaries for teachers with higher educational attainment. Similar considerations are in order with respect to pre-service and in-service training. From a cost-benefit perspective, short but well designed and practice oriented programs appear to be most promising.

17. Finally, it appears highly relevant to ensure the maximum use of formal instruction time for effective teaching. Double shift teaching seems to have a detrimental impact in this respect. As there is ample evidence for a rather modest negative impact of high student-teacher ratios, double shift teaching should generally be avoided. It should be noted that this reasoning does not apply to multi grade teaching which generally does not show any negative effect.

18. Effective teaching time can also be increased by improving students' attendance. Apart from the well-known requirement of adjusting the academic year to harvesting seasons, attendance can be increase by simple health care measures. In this context, de-worming has been shown to be particularly cost-effective.

19. And last but not least, effective teaching time can be increased by reducing teachers' absences. In some cases, simple administrative measures like the reorganization of teacher remuneration (so that teachers do not need to collect their pay from a far away district officer) may be very effective. In general, however, more effective control mechanisms seem to be required.

20. This creates the link to the relevance of non-material inputs such as accountability, effort and motivation. Such inputs cannot simply be bought or provided. They require incentives, which in turn require appropriate institutions, i.e. a set of rules and regulations which reward socially positive and punish socially negative behavior. While it is difficult to directly observe teachers' and principals' accountability to parents and the society (or to the local community), and whether teachers are effectively doing the work they are paid for, the underlying institutions can be observed.

21. The institutional features most frequently discussed in the economic literature are private-sector participation, decentralization of responsibilities, and the role of standardized exams. In this context, the debate is largely driven by research for industrialized countries in which the potential to improve student performance can be shown to be much higher for institutional reform than for any reform based on a quantitative or qualitative improvement of physical inputs. With particular relevance for developing countries, notably in sub-Saharan Africa, we may add the design of appropriate teacher contracts, and simple control and reward mechanisms like those (potentially) related to the visit of a school inspector.

22. Notably, research for PASEC countries shows that teachers on non-civil servant fixed term contracts miss their classes significantly less often then their colleagues. Unfortunately, contract teacher programs generally combine various features with partly contradicting consequences for student achievement. Three different potential effects can be distinguished: (1) an incentive effect of the teaching contract, (2) a selection effect (changed demand for and supply of new teachers), and (3) a dynamic effect on teachers' attrition. Preliminary evidence suggests that the (positive) incentive effect works best if contract teachers are employed by parents and local communities, rather than by public authorities. Indeed, this should enhance teachers' accountability and parents' incentive for effective monitoring. Theoretically, this system could be generalized by channeling public funds for teacher remuneration via local communities and parents' associations.

23. Other aspects of decentralization and increased local autonomy (both for parents, and for schools and teachers) may also be beneficial for student learning. In particular, any kind of measures to enhance transparency about resource flows and learning outcomes appears to be valuable. Standardized national exams already exist and their results, complemented by some information on student socio-economic background as well as school drop-out and repetition, could become a valid low-cost indicator for school effectiveness. This would also be a first step towards even more comprehensive institutional change.

24. Whether full fledged competition between schools could be a successful and realistic approach for primary education systems in the majority of sub-Saharan African countries remains rather questionable, however. Trying to promote competition while ignoring some preconditions unlikely to be met in the African context entails the serious risk of segregation and increased distributional inequalities without reaping the expected efficiency benefits. In particular, for school competition to be successful, effective options for parental choice need to exist, which is certainly not the case in large parts of the countries concerned, especially in rural areas. Moreover, complete cost coverage would have to be ensured, which might be extremely difficult as it would have to include hidden cost such as unofficial fees requested by school administrators or opportunity cost for traveling especially for children contributing to their families' household activities.

25. Overall, it appears that incentive based approaches may have a relevant impact in sub-Saharan Africa. While full fledged competition between schools does not seem to be a promising policy option for the time being, more effective monitoring and control, and direct incentives for teachers based on accountability towards parents and local communities can be implemented straight away. The political appeal of these measures is that they may bring along considerable improvement in student learning without relevant direct financial implications. In some cases, like in the case of contract teacher programs, budgetary implications can even be strongly positive. At the same time, these measures can bring about some political cost as they may face opposition by relevant stakeholders such as teacher unions. These political costs can be minimized by creating appropriate policy packages which, for instance, jointly introduce more autonomy and control.

26. As both physical resources and immaterial inputs such as accountability, effort and motivation appear to be relevant for education quality, the improvement of student learning in sub-Saharan Africa requires that institutional reform and traditional input oriented policies should go hand in hand. Moreover, carefully designed strategies based on a mix of resource based measures and institutional reform could greatly enhance cost-effectiveness and therefore have a significant impact on education quality even under given budgetary constraints.

3. INTRODUCTION

27. The increasing availability of student survey data, the development of new statistical and econometric methodologies and the expansion of computing capacities has led to a huge increase in scientific evaluations of the determinants of education quality in recent years. Education quality is thereby measured in terms of student achievement on standardized tests, which reflects the cognitive knowledge acquired through the education process. In line with international policy priorities as codified in the Education for All (EFA) objectives and the Dakar Framework for Action, for sub-Saharan Africa, evaluation efforts currently concentrate on the primary level. In addition to various national level evaluations, three programs have been launched on a larger scale: The UNESCO/UNICEF Monitoring Learning Achievement (MLA), the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and the Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN[1] (PASEC) now jointly cover most countries on the continent. General information on these programs is available from Chinapah (1997) for MLA, Ross (1998) and Murimba (2005a) and (2005b) for SACMEQ, and PASEC (1999) and CONFEMEN (2005) for PASEC.

28. For a general review of results, SACMEQ and PASEC are of particular interest because they use comparable (or identical) tests in all their countries, which allow us to jointly analyze different country cases as well as to draw comparisons across countries. While the comparison of test items and thus a direct comparison of achievement levels across programs is not yet possible, the relationship between inputs and outcomes can be compared. As SACMEQ data have only recently become publicly available, to our knowledge, this study presents the first attempt to jointly explore results for francophone and anglophone Africa in a common education production function framework. This will be done within the context of a general review of the wider literature on determinants of education quality, including insights from industrialized countries as far as they may be of interest for the African context.

29. While earlier studies typically focused on the impact of school resources, such as learning materials, classroom resources, school equipment and the number of (qualified) teachers, the discussion has recently moved on to questions of setting the right incentives and finding the appropriate institutional structure for effective learning. Moreover, there has been a new emphasis on peer effects among students which may alter the interpretation of earlier results and raise questions with respect to the efficient distribution of available inputs, and with respect to the allocation of students to schools.

30. In this study, we will first resume the available evidence on the impact of traditional inputs on student achievement and discuss whether these results are in line with our findings from SACMEQ and PASEC for primary education in sub-Saharan Africa (Section 4). We will then discuss what could be learned from the analysis of institutional and incentive related factors and again compare these findings with results from PASEC and SACMEQ (Section 5). Finally, we will consider the role of peer effects (Section 6).

31. As setting policy priorities requires knowledge about both benefits and costs, as far as this is possible, a (crude) appreciation of the budgetary implications will also be provided for the measures discussed in each of the sections. Section 7 will present the conclusions.

[1] Conférence des ministres de l'éducation des pays ayant le français en partage

4. THE IMPACT OF TRADITIONAL SCHOOL RESOURCES ON STUDENT LEARNING

4.1 A general overview

32. In this section, we can draw from a considerable number of previous studies including excellent literature reviews such as UNESCO (2004), Hanushek (2003), Glewwe and Kremer (2006) and the papers prepared for the ADEA's 2003 high level conference on "The Quest for Quality: Learning from the African Experience" (in particular Mingat, 2003, Michaelowa, 2003 and Verspoor, 2003). As outlined in most of the literature reviews, generally, the results of different empirical studies are highly inconsistent, and the overall picture is rather bleak in terms of truly promising policy options. In fact, many of the studies raise doubts about the relevance of traditional inputs in the schooling production function all together (Hanushek, 2003; Glewwe et al., 2004, Glewwe and Kremer, 2006). Although there have been large improvements in the levels of school resources around the world, no corresponding improvement of student learning could be observed. As Hanushek (2003: F67) puts it:

> *"Class sizes have fallen, qualifications of teachers have risen, and expenditures have increased. Unfortunately, little evidence exists to suggest that any significant changes in student outcomes have accompanied this growth in resources devoted to schools."*

33. To begin with, aggregate international data do not show a strong relationship between test performance and the patterns of expenditure across countries (see Hanushek, 2003: F72, Figure 2 and UNESCO, 2004: 63; Mingat and Suchaut, 2000; or PASEC, 2005a for a selection of African countries). While, to a certain extent, in a cross-country setting this could be explained by unobserved differences in other country characteristics, it is stunning that even when observing a single country over time, there is no clear macro-level evidence that added resources have a positive impact on learning achievement. Such comparisons over time are possible for the USA where data on student performance have been collected over the last 45 years, as well as for a limited number of other OECD countries. A comparison of change in real expenditures per student and other inputs with the test scores in mathematics and science for primary and secondary education in eleven OECD countries from 1970 to 1994, shows that test scores fell in seven of the eleven countries, even though every country increased the resources devoted to schools (UNESCO, 2004: 60). This suggests that the efficacy of input-based policies depends decisively on the effective use of resources - rather than simply on their availability (Hanushek, 2003: F69)) - and more broadly on the social and economic incentives associated with formal learning.

34. Research at the micro level is not able to show the desired clear relationship between resource inputs and test scores either (UNESCO, 2004: 65). A meta-analysis of estimated effects of key resources – such as teacher numbers (reflected in class size), teacher education and experience, teacher salaries, school facilities, and financial resources in general – on student performance, based on over 400 estimates of education production functions in the USA and other developed countries,

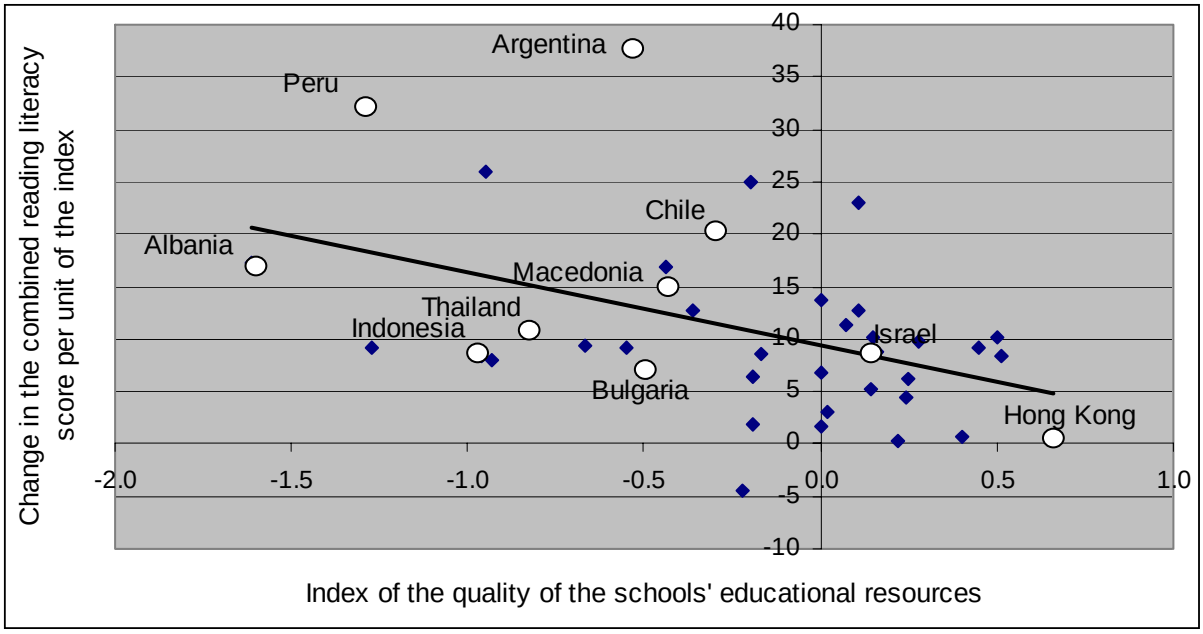
does not show any obvious relationship between increases of particular inputs and increases of student achievement (Hanushek, 2003 : Tables 5 and 7, F76 and F86).

35. One explanation of the missing relationship between resources and student achievement in the USA and other industrialized countries is that schools already dispose of a high level of resources, and therefore operate in an area of diminishing marginal productivity (Hanushek, 2003: F83f.). In other words, the potential success of input-based policies depends on the countries' existing level of resources. The impact of adding further resources diminishes when the initial equipment status is already satisfactory, and in fact, in industrialized countries, variation in school resources is itself so limited that it cannot be expected to explain much of the variation in results (UNESCO, 2004: 67).

36. Correspondingly, one should expect the relationship between resources and outcomes to be much clearer for developing countries. Indeed, looking at 96 production function estimates in less developed countries reveals a somewhat stronger support for the expected positive relationship (Hanushek, 2003: F84). Analyzing 60 studies of education in developing countries, Fuller (1987) also finds that resources were more important determinants of students' achievement in developing countries than in industrialized countries. Fuller and Clarke (1994) reinforce this conclusion taking into account the cross-country differences in socioeconomic and cultural settings even within developing countries.

37. An illustrative example can also be taken from the international PISA study, which includes several lower and middle income countries along with OECD countries. Figure 1 depicts the average impact of additional (or quality improved) pedagogical resources on student learning for each country covered by the survey. Non-OECD countries are marked by a white dot and identified with their name. Countries are ordered by their current level of resource availability (on average across all schools), captured by an index averaging zero in OECD countries. The vertical axis shows the improvement of PISA reading literacy test scores (OECD average scores: 500 points, standard deviation: 100 points) for a change of one standard deviation in resource availability.

Figure 1 The impact of educational resources depending on initial availability



Source: OECD and UNESCO-UIS (2003: 195)

38. Despite much variation, it becomes clear that in high income countries, where the initial stock of pedagogical resources is already high (towards the right hand side of the figure), the impact measure is typically small, while it is much higher for lower income countries. For instance in Hong Kong, where initial endowment is high, adding pedagogical resources is estimated to have zero effect, while in Argentina, where initial endowment is low, increasing resource availability by one standard deviation would lead to a rise in test scores of more than 15% of the international standard deviation. The average relationship between the estimated impact of additional resources and initial availability is indicated by the regression line. Unfortunately, the indicator used for resource availability is rather subjective, as it is based on a general assessment of school principals. Otherwise, the relationship might have come out even more clearly.

39. We conclude that despite rather discouraging evidence on the international level, for developing countries in general, and for most of the very poor sub-Saharan African countries in particular, school resources still play an important role in improving education quality. However, even for these countries, the estimated relationship between school resources and student achievement is far from consistent across studies, so that there is no easy recipe for successful policy interventions.

4.2 A discussion of policy priorities for developing countries

40. Obviously, determining a set of promising policy interventions is much more demanding than simply providing an overall appreciation of input effectiveness. In order to determine promising educational interventions, the competing policy options have to be evaluated and ranked according to both their expected level of effectiveness and their cost. Despite donor commitments in the context of EFA and EFA - Fast Track, each country needs to consider its own budget constraint in order to derive policy options which are not only potentially effective, but also realistic and sustainable. Despite several new studies using interesting estimation techniques (see e.g. Glewwe et al., 2004 and the review by Glewwe and Kremer, 2006), overall evidence on the different interventions has not changed much in recent years, so that we can draw from previous studies to discuss the results. While the magic bullet does not exist, some crude conclusions can be drawn from the evidence available so far:

- Pedagogical resources such as textbooks, teacher guides, wall charts, etc. are relatively low cost inputs with relatively high returns in terms of student achievement. For each of the main school subjects, textbooks should be available (Mingat, 2003; Verspoor, 2003; Lockheed and Verspoor, 1991). Wall charts may be an even cheaper alternative as only one chart is needed per subject for a whole class (Glewwe et al., 2004[2]). But as textbooks are more frequently included in empirical analysis, the result on textbooks can be regarded as more reliable. Results by Glewwe, Kremer and Moulin (2000) suggest that, to ensure benefit for all students, care has to be taken that the level of these books is not only targeted at the very best students. Moreover, it obviously must be ensured that the books actually reach the students, and that they are effectively used in the classroom. Teacher manuals have been shown to be very helpful complements of textbooks, which may, in fact, increase the chances that textbooks themselves are effectively used. Moreover, teacher manuals guide the teacher in a simple and practical way and facilitate the focus on the same curriculum and pedagogical orientation everywhere.
- Buildings and furnishings are generally not found to be very relevant. Equipment should be simple and durable. A usable blackboard is important (Glewwe and Jacoby, 1994), but seems to be available almost everywhere by now.

[2] Note that the authors use the term “flip charts” because different charts are spiral bound together. However, as this term could be misunderstood as a set of empty sheets to write on, we prefer to use the term wall charts here.

- Human resources are a very expensive input. Higher teacher numbers reduce class size and student-teacher ratios, but this is one of the policies where empirical results based on student surveys have been most inconsistent. At a cross-country level, it can be observed that among the world's best performers in terms of student achievement, are countries in the Nordic region with the smallest average class size world-wide, while others are Asian countries with student-teacher ratios close to those typically found in sub-Saharan Africa (OECD and UNESCO-UIS, 2003: 200, Figure °7.7). Nevertheless, for Africa, there is some more consistent evidence for a negative impact beyond a threshold of approximately 60 students per class (Michaelowa, 2001 and Verspoor, 2003). Below this threshold, the evidence currently available suggests that reducing class size should certainly not be a policy priority.

  As far as the quality of teachers in terms of pedagogical and subject knowledge is concerned, it appears that the duration of pre-service and in-service training is much less important than the practical relevance of its content. A long training is not necessarily a good training, but it is definitely an expensive training. Brief but well focused training periods (e.g. a pre-service training of only several months, accompanied by in-service follow-up and tutoring by more experienced teachers) may be effective alternatives to several years of pre-service training during which the teachers often have to be paid without being actually active in class. At the same time, it should be ensured that all teachers effectively receive some training (see e.g. PASEC, 2004). With respect to requirements for teachers' academic qualification, lower secondary attainment (O-levels or BEPC) is generally considered sufficient for teaching primary students. Higher levels of educational attainment are often associated with the expectation of higher pay, but do not seem to significantly enhance students' learning (Mingat and Suchaut, 2000; Mingat, 2003)
- Class management via multi-grade teaching when there are few children in each grade, or via teaching in shifts when there is a lack of classrooms and / or teachers, can help to overcome a situation of resource constraints. It turns out that multi-grade teaching has no general disadvantages. In fact, in industrialized countries, multi-grade teaching has recently been reintroduced in various schools as a promising pedagogical concept. Its pedagogical effectiveness can be enhanced by specific training for the teacher (Juvane, 2005). As the cost-advantage in terms of reduced teaching staff is highly relevant, multi-grade teaching appears as an efficient educational management strategy, especially in remote rural areas with a low population density.

  In contrast, double shifts often show a strongly negative impact on student achievement (Michaelowa, 2001) and often reduce cost to a lesser extent than expected, especially when different teachers are responsible for the two shifts, or when overtime rates have to be paid for those teachers in charge of both shifts (Mingat and Suchaut, 2000). Given the generally rather modest effect of changes in class size, double shifts should only be considered for extremely large class sizes. If an education system decides nevertheless to recur to double shifts, careful consideration should at least be given to ways to avoid significant losses in effective teaching time for either of the shifts (see below).
- Grade repetition blocks considerable resources as it artificially expands the number of students in the education system. At the same time, the frequently stated hypothesis of a positive effect of grade repetition on student learning is clearly rejected by the empirical evidence. At best, students achieve a transitory gain with respect to the rank within their class. At worst, their performance gains are effectively reduced or they drop out from school altogether (Bernard, Simon and Vianou, 2005). Thus reducing repetition rates appears as a cost reducing and yet quality improving measure of educational policy, which should therefore be a high policy priority. The resources set free by reduced repetition rates can be used for other, more promising programs to assist weak students. In India, local women have been successfully employed as teaching assistants in disadvantaged areas with small one-teacher schools, to help these students catch up with their peers. This was shown to be a particularly effective measure for students with important learning deficiencies (Banerjee et al., 2003). In the Indian example examined by the authors, this intervention only resulted in an annual cost of 5$ per child. Even if the employment of pedagogical support staff may be more expensive in some African countries, there is no doubt that it is less expensive than current repetition practices.

Table 1 Setting priorities – comparing policy options and their cost-effectiveness

Policy measure	Assessment	Policy recommendation
Learning materials		
Textbooks and books in general	• Among the most cost-effective inputs for student learning • As subjects and reading capacity increase with the level of education, the number of textbooks provided should also increase	+ Provision of textbooks in core subjects (literacy+math) in the first years of schooling; 3-4 textbooks in more advanced classes + Provide access to libraries
Wall charts	• May be equivalent to textbooks in their effectiveness in the learning process • A single wall chart is less expensive than a class set of textbooks	+ Potentially even more cost-effective alternative to the provision of textbooks (but little empirical evidence available)
Teacher manuals	• Very positive impact on student achievement as they guide the teacher in a simple and practical way and create homogeneity in the teaching	+ Provision of teacher manuals for all core subjects
School equipment		
School buildings	• Little evidence of a significant impact on student performance • Much resources can be wasted on expensive formulas which do not show any effect	+ Simple constructions satisfying basic requirements of acceptable durability and cleanliness + Examination of the possibility of community managed programs
Furnishings	• Cannot be expected to have a considerable influence on student performance • Furnishing can be provided locally at low cost	+ Priority should be given to functional low-cost equipment
Teacher numbers and qualification		
Teacher numbers (student-teacher ratios)	• Rather modest impact of student-teacher ratios on student achievement (low risk of a quality-quantity trade-off) • As teacher remuneration is expensive, a reduction in student-teacher ratios substantially increases cost	+ Class size up to 60 pupils is acceptable + Hiring specialized teachers is unadvisable especially for small schools because it yields financially unacceptable student-teacher ratios without even decreasing class size
Teachers' educational attainment (academic qualification)	• Little evidence of strong learning advantages from enhanced duration of teachers' education • Much higher salaries for teachers with higher educational attainment	+ Accept grade 10 (lower secondary) attainment as a sufficient minimum requirement for primary school teachers
Teachers' professional qualification • pre-service training • on-the-job training	• Professional training is important • Content is more important than duration • Long-term pre-service training is expensive especially if teachers need to be paid without being active in class • If pre-service-training focuses effectively on the practical skills needed in class it can be reduced to a few months; 2 or more years do not lead to significantly higher student achievement • On-the-job training is a relevant complement to pre-service-training and can be implemented at low cost and without negative impact on teachers' presence in class.	+ Provide some professional pre-service training to all teachers + Save resources on duration but focus on content, practical relevance and in-service follow-up (e.g. tutoring) + Pursue inexpensive in-service training strategies like radio broad-casting + Avoid programs which reduce teachers' presence in class

Table 1 (cont.)

Policy Measure	Assessment	Policy recommendation
Organization of student flows and study time		
Double shift organization	• Typically, reduced effective teaching time and unsuitable class hours lead to considerable losses in student learning, especially after the 3rd grade • Positive financial implications as compared to two separate classes, in particular if there is a shortage of classrooms • However, negative financial implications as compared to a single larger class, because even if there is only one teacher for two shifts, he or she must be compensated for overtime, often at higher rates	+ Optimize time schedules in cooperation with parents and make sure that official teaching hours are effectively respected. + Avoid double shifts as long as students can be put in a single class (only with more than 80 students, the negative effect of class size becomes so strong that double shifts usually become preferable)
Multiple level organization (multi-grade teaching)	• Acceptable from a pedagogical point of view: may even enhance student learning • Considerably reduces cost in small schools as students of different grades do not need separate teachers • Enables small schools to accommodate students of all grades of the primary cycle. This may be crucial to avoid drop-out. • Requires specific teacher training for effective implementation	+ Encourage multi-grade teaching in small schools (remote rural areas) + Provide teacher training to ensure effective implementation
Class repetition	• High repetition rates create high private and public costs without increasing educational achievement	+ Reduce repetition rates to <10% + Resources can be devoted to more promising policy measures to assist disadvantaged students
Effective teaching time	• Decisive factor for student learning • Effective teaching time is often much shorter than theoretical teaching time due to deficiencies in organization and monitoring (high absence rates of teachers and students and ineffective management of time in class) • Flexible schedules compatible with children's duties at home (especially harvesting seasons) reduce absence rates and early drop-out without relevant implications on cost • Health checks and medical treatment equally reduce absence rates and drop-out. Financial consequences depend on the concrete measures. Especially de-worming has been shown to be highly cost-effective. • School meals also increase demand for schooling and reduce absence rates. Full meals are, however, relatively expensive, so that these interventions should be well-targeted. A less costly alternative is to provide relevant nutrients in the form of smaller snacks.	+ provision of a sufficient teaching force as well as adequate monitoring to guarantee the full execution of the school term + sanctions for unjustified absences of teachers" + organize a flexible schedule to decrease the children's' opportunity costs of schooling without reducing the scope of teaching + organize regular de-worming to reduce health related irregularities of attendance + organize school meals (or smaller snacks of high nutritional value) to provide incentives for children to attend school

Sources: Mingat (2003), with additions from Verspoor (2003), Michaelowa (2003, 2001), Glewwe et al. (2004), Glewwe and Kremer (2006), Tatto, Nielsen and Cummings (1991).

- Effective teaching time is the most basic resource required for effective learning at school. There are various ways to ensure the presence of both teachers and students. In particular, effective teaching time can be increased by a thoughtful adaptation of the school calendar to agricultural needs (children may be required for harvesting activities), by the provision of school meals, and

by health interventions like de-worming which avoid illness related absences or a lack of concentration in class. Glewwe and Kremer (2006: 26) argue that de-worming is a particularly cheap but highly effective measure in many developing countries. Incentive and control measures are required to reduce teacher absence rates.

An overview of the different measures and their projected cost-effectiveness is provided in Table 1.

4.3. Econometric evidence for francophone and anglophone Africa (PASEC and SACMEQ)

41. We will now examine the evidence from PASEC and SACMEQ data, using a common education production function framework, to assess whether the above discussion drawn from the general literature (and partially from earlier PASEC studies) is consistent with our results. Wherever additional differentiation between individual countries appears relevant, we will also refer to recent analysis by the CONFEMEN secretariat or to Lee, Zuze and Ross (2005), which, to our knowledge, is the first econometric analysis of overall SACMEQ results (unfortunately available only for reading performance, not for math).

42. The SACMEQ data base includes more than 40 000 6th grade students from 13 countries: Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (main land and Zanzibar)[3], Uganda and Zambia. The PASEC data used here includes more than 17 000 5th grade students and the same number of 2nd grade students from eight countries: Burkina Faso, Cameroon, Côte d'Ivoire, Madagascar, Mali, Niger, Senegal and Togo. All surveys were carried out between 1995/96 and 2001/02. For both sets of countries, we estimate the effect of various policy options on student test scores in literacy and mathematics. The policy options discussed include the measures presented in Table 1 as far as the corresponding data is available from the surveys. Moreover, they include some institutional variables, which will be of interest in Section 5. All effects are calculated after controlling for the influence of student socio-economic background, e.g. possessions at home, mothers' and fathers' education, language spoken at home etc.

43. In PASEC, students have been tested twice, once at the beginning, and once at the end of the academic term (pre-test and post-test). This allows us to also control for student performance before the school, teacher and classroom related influences measured for the corresponding term actually start to become relevant. As a comparable variable is not available for SACMEQ, PASEC results are computed twice, with and without the pre-test variable. This procedure ensures that real differences between the two country groups can be distinguished from differences which are merely induced by the introduction of the pre-test variable.

44. A more detailed discussion of the econometric methodology and of the differences between the two surveys, which should be kept in mind when comparing their results, is provided in the appendix. The appendix also includes two detailed tables with regression results for literacy and mathematics respectively (Tables A1 and A2). The following discussion concentrates on the most relevant results.

[3] As separate surveys were carried out for the main land and for Zanzibar these areas are considered in the following analysis as if they were different countries.

45. Most results are in line with the findings for developing countries in general, as outlined above. We do find a positive effect of textbooks everywhere, with, however, a relatively wide variation of the size of this effect. A change from zero textbooks to full coverage of one book per student for the whole class yields improvements in student achievement of between 5-20% of a standard deviation, depending on subjects, grades and regression specification. In 2nd grade, it seems that possessing the book (as compared to sharing it with another student) is more important than in higher grades, especially in French. One could imagine that in lower grades, being able to take the book back home for first reading practice is more relevant than in higher grades.

46. However, higher coefficient estimates and more strongly significant results for individual textbook possession may also be an artifact of the lack of two relevant control variables – parents' literacy and books at home– which were not included in the questionnaire for 2nd grade students. As the expected correlation between these variables and textbook availability is positive, 2nd grade coefficients for textbooks are likely to be biased upwards. Moreover, the generally lower level of initial textbook availability in earlier grades may lead to higher coefficients if there are diminishing returns to overall textbook coverage (for a more general discussion of such non-linearities see Section 6). Thus the distinction between grade levels is more complex here than it might seem at first glance.

47. The question about wall charts was asked only in SACMEQ countries. The coefficient estimate is positive, as expected, but remains insignificant. Teacher manuals are significant in some regressions (only for SACMEQ) and then lead to a positive effect of up to about 7% of a standard deviation in test scores. For PASEC, they become significant in different regression specifications with a lower number of general equipment variables and for a different set of countries (not shown here). But results are clearly less robust than for textbooks.

48. With respect to school and classroom equipment, results are also in line with those outlined in Table 1. We spent much time trying to find appropriate indicators using different combinations of classroom furnishings, school facilities and basic equipment, such as chalk and blackboards. The final specification presented in Tables A1 and A2 includes a variety of separate indicators for individual items and facilities, a joint indicator for higher technology equipment, such as computers, television and video projectors, an indicator for the availability of electricity and an indicator of the general condition of the school building.

49. At first glance, looking at SACMEQ regressions, our results seem to present a strong evidence for the relevance of expensive electric equipment. The indicator for higher technical equipment is strongly significant and indicates that adding any high tech item to the existing equipment of a school raises student achievement by more than 12% of a standard deviation. However, this variable must be considered with caution, as it may well suffer from an endogeneity problem: As high tech equipment is an easily visible signal of a rich school environment, wealthy parents and parents with particularly talented children may select these schools in the first place. As most parents can be expected to make their school choice only once (i.e. at the beginning of primary education), controlling for the initial score at the beginning of the year, as possible with PASEC data, will eliminate at least part of this selection effect. Unfortunately, the high tech indicator is not available in PASEC, but electricity, a strongly correlated variable, is. In PASEC, the effect of electricity is significantly positive at the 10% level at 5th grade in French (and close to significant in 2nd grade), but only as long as the pre-test score is not included into the regression. Controlling for the pre-test scores fully eliminates any significant effect.

50. In SACMEQ, the availability of a school or classroom library also appears to be significant, whereby the existence of the library in the classroom itself seems to be more directly beneficial. Not surprisingly, results for reading are higher than for math and make up almost 10% of a standard deviation in literacy scores. The library result is also reflected in two of the PASEC regressions (grade 2). One might take this as yet another indication of the relevance of books in the learning process. Note that libraries also offer a compensation for a scarcity of reading material at home. The variable "books at home", which is introduced as one of the control variables for students' family background, is strongly significant in all literacy and most mathematics regressions. This reinforces the potential relevance of libraries in general, be it at classroom or school level, in the village or town, or in the more flexible form of a "rolling library", which appears to be most cost-effective in scarcely populated rural areas.

51. Nevertheless, it should be noted that coefficients for school libraries shrink considerably and become insignificant when the pre-test scores are controlled for. This suggests that, just as in the case of technical equipment, a self-selection process of good performers into well-equipped schools may bias the results.

52. A similar argument applies to the interpretation of the coefficient for the condition of school buildings. The condition of school buildings – only included in the SACMEQ analysis – reveals a strong and statistically significant positive effect: A change from extremely bad to extremely good conditions leads to an increase of about 10% of a standard deviation of student achievement. However, just as technical equipment and school libraries, the condition of the school building is one of the easily observable characteristics parents may base their school choice on. As the variable is not included in the PASEC analysis, it could not be tested whether the coefficient estimates remain significant when initial knowledge is controlled for. When a related variables providing information about the material the classroom is built with are included in individual PASEC country studies, results generally do not show any relevant positive role of concrete relative to other materials (see e.g. PASEC, 2005c).

53. Otherwise, very few significant effects can be reported. A certain positive effect of the availability of blackboard and chalk can be observed for mathematics, but not in all regressions and only weakly significant in the case of PASEC. In French, the estimates are insignificant and / or even show a negative sign. Toilets, health equipment and fresh water do not show a significant positive effect, either. All in all, evidence for relevant effects of school equipment is rather weak, especially when considering potential selection bias and the more reliable estimates controlling for pre-test scores.

54. Results with respect to class size show the typical insignificant or very small impact on student achievement. In order to take into account possible threshold effects or other non-linearities, the variable is entered into the regression in a quadratic form. In the case of SACMEQ, where the coefficients are significant, the analysis indicates that negative effects start to become evident beyond a class size of 60 students. This result corresponds exactly to earlier results for PASEC in a regression specification for five countries (Michaelowa, 2001) and to the policy conclusions in Table 1. In the regressions specified here, class size is insignificant for the PASEC countries. Another study based on PASEC panel data for students in Senegal, controlling for student fixed effects, does not find any negatively significant effect either (Fehrler, 2005).

55. Teacher qualification is a different issue. For PASEC, neither the indicator of teachers' educational attainment (academic qualification), nor the indicator for professional training is significant at the 5% level. Only for 5th grade math teachers can some positive effect on educational

attainment be discerned, which is significant at the 10% level in one regression, and close to significant in others. In SACMEQ, however, the academic qualification is clearly significant and the professional qualification is significant in all but one regression. Coefficients for academic qualification indicate that the students gain between 2 and 4 percent of a standard deviation in scores when the teacher has attained a one step higher level of education, e.g. lower secondary attainment instead of primary attainment only, or some tertiary instead of simply upper secondary.

56. It is interesting to note the differences between SACMEQ and PASEC countries here. Although the indicator used is almost identical in both surveys, in PASEC, it is much more difficult to find the expected positive results. The problem appears to be that the indicators of both professional training and educational attainment only capture duration while no information is available on quality. Obviously, depending on quality and practical relevance, two different courses of the same duration may have a totally different impact on actual teaching skills. It can be shown that in PASEC, there is no significant positive correlation between the duration of teachers' educational attainment and teachers' knowledge of the subject matter. This implies that the low coefficient estimates for attainment should not be interpreted as an indication of a low impact of increased subject matter knowledge, but rather as an indication of the low quality of the education the teachers themselves received when they attended school (Michaelowa, 2003).

57. To measure actual teacher knowledge, PASEC uses an exercise for teachers in which they have to count the mistakes in a fictitious student dictation. In SACMEQ, a different and exceptionally precise indicator of relevant teacher knowledge is available: Teachers were themselves asked to take the students' tests and marked on the same scale. The average teacher score in literacy is more than two standard deviations above average student scores and is reached only by about 2% of the students.

58. As opposed to PASEC, it can be shown that for SACMEQ countries the correlation between educational attainment and teacher test scores is significant, albeit even here, less pronounced than one might have expected. Estimated correlation coefficients are $\rho=0.21$ for literacy, and $\rho=0.32$ for math. Since we can find a significant correlation only for SACMEQ countries, this may indicate that, on average, the quality of secondary and tertiary educational institutions attended by (future) teachers is better in anglophone than in francophone Africa, at least in the core subjects of literacy and mathematics. This could explain the differing results on the relevance of the academic qualifications. One should be cautious, however, when interpreting these results, because the indicator of teachers' subject matter knowledge in PASEC is much less reliable than the one used in SACMEQ. Moreover, neither in PASEC, nor in SACMEQ are the indicators for teachers' subject matter knowledge available for all countries. This is also the reason why these indicators have not been included directly in our regressions in Tables A1 and A2.

59. In any case, it should be noted that the coefficient estimates of 2-4% of a standard deviation for a full level of education (like the whole upper secondary cycle) are not very high when compared to the cost incurred for this additional education, including the opportunity cost of having the teachers start effectively teaching much later, and the higher pay they can expect with a higher academic qualification. While the linear specification of educational duration used here does not indicate any optimal cut-off point, some prior research on PASEC indicates that this may be below the A-levels or baccalauréat (successful upper secondary completion).

60. It has been shown that teachers holding a baccalauréat are often less motivated than their peers with lower educational attainment, possibly because their higher expectations with regard to their future jobs are not met by the reality of their situation (Michaelowa, 2002). Moreover, even if there was a linear increase in the impact of educational attainment on student achievement, cost in

terms of salaries would increase over-proportionately, with a strong jump related to the completion of the upper secondary final examination. We can thus confirm the results presented in Table 1, that raising entry requirements for the teaching profession to include the successful completion of upper secondary education or beyond should not become a policy priority.

61. As mentioned above, the differences in the significance (or lack of significance) of SACMEQ and PASEC can be observed not only for teachers' academic qualification, but also – in a similar way – for their professional training. In this context, there is no way to directly show from the data that this may be related to a different quality of the courses offered. The correlation between teachers' professional training and subject matter knowledge is not very strong, even in SACMEQ countries, but this is plausible even for very good training modules since professional training could focus on pedagogical rather than academic skills. Most probably, the reason for difficulties in finding significant results in overall PASEC regressions is that professional qualifications vary widely across countries (even within the francophone education systems) and are more or less effective, so that it is very difficult to capture their overall effect.

62. Individual country estimates for PASEC have often shown the relevance of professional training for student achievement (see, in particular, PASEC 2004). In their individual country regressions for SACMEQ, Lee, Zuze and Ross (2005) construct a joint estimate for academic and professional qualification, so that results are not directly comparable. Nevertheless, they also find that the effect varies widely between countries. A positively significant impact is only found for about one third of the countries covered (and insignificant effects otherwise). In this context, it may be argued that duration (the only available measure for professional training) is less relevant than content (Michaelowa, 2003, Bourdon, Frölich and Michaelowa, 2006). If the latter could be adequately measured, we would probably face much less variation of results between individual countries and between country groups.

63. Similar reasoning applies to in-service training (see e.g. Nguyen, Wu and Gillis, 2005: 40f.). The latter is negatively significant in SACMEQ. This is a counter-intuitive result also found for individual country cases in francophone Africa, and often related to training sessions during class hours which then reduce effective teaching time (Bernard and Michaelowa, 2005). It should also be noted, however, that in SACMEQ, the in-service training variable is based only on teachers' own subjective assessment of the efficacy of these courses. In PASEC regressions, the variable reflects the number of courses attended per year, and teacher absence can be directly controlled for (in SACMEQ, only an indirect school level variable is available). In this setting, in-service training has a positive coefficient, which is significant for 5^{th} grade French and implies an improvement of up to 5% of a standard deviation in students' scores for each additional training seminar the teacher has attended per year (during the last five years).

64. Coming to the organization of student flows, our analysis confirms the negative effect of double shift teaching. As the control for pre-test scores generally reduces the overall effect (and makes it statistically insignificant in some regressions), parts of the effect seem to be related to a selection of bad performers in double-shift classes. However, after controlling for initial knowledge, the negative coefficients remain and still indicate losses of often more than 10% of a standard deviation in student test scores for double-shift classes. As opposed to earlier analysis, we do not find any evidence that this effect is weaker in 2^{nd} grade. In fact, 2^{nd} grade mathematics shows the most significant negative results among all PASEC regressions.

65. SACMEQ regressions for 6^{th} grade only indicate losses of up to 6% of a standard deviation in the case of double-shift organization, and the results are significant only in one regression

(even at the 10% level). However, if we look again at the individual country regressions carried out by Zuze, Lee and Ross (2005), we find that in some countries, this variable does not seem to be relevant in current education practice. In fact, the authors include it only in 9 out of 14 regressions, 4 of which show the expected significant negative effect, sometimes with extremely high coefficients corresponding to up to about 30% of a standard deviation of (international) student scores (Kenya and Zambia).

66. As opposed to double-shifts, no significant effect in either direction can be discerned for multi-grade teaching. Unfortunately, this variable does not exist in the SACMEQ database. The reason might be that in SACMEQ, very small schools for which this system is generally most relevant have been excluded from the target population (see appendix, Section 8.2).

67. Grade repetition, introduced to help students to catch up with their peers, can be shown to be counter-productive. The coefficient in all regressions is clearly negative and significant, except for the initial year, in which the repeating student appears to benefit from a short-term advance on his new classmates, at least in grade 5. As indicated by a comparison of regressions with and without control for the pre-test score, this short-term advance is primarily due to the students' higher initial knowledge. In the SACMEQ regressions and most of the PASEC regressions for 2nd grade, even the coefficient for current class repetition on test scores is negative. In many cases, repeaters can be shown to do much worse than their classmates, with an overall performance reduced by 15-20%. Results are robustly negative and highly significant, even in individual country regressions. Coefficient estimates for some countries correspond to student achievement reduced by more than half of a standard deviation of student scores (for Botswana and Mauritius) (Lee, Zuze and Ross, 2005).

68. Caution is required, however, when interpreting these results, since in models without control for initial ability, these figures cannot be interpreted as a causal relationship. A bad student repeating his grade is not necessarily doing badly because he repeats, but rather repeats because he has been doing badly. Nevertheless, as strongly negative results can also be established in PASEC regressions in which initial knowledge is controlled for, they cannot be considered as a mere issue of reverse causation. This confirms the conclusions of Table 1 and of a detailed PASEC analysis on the specific issue of repetition based on information for students followed through several years of their primary education in Burkina Faso, Côte d'Ivoire and Senegal (Bernard, Simon and Vianou, 2005 and PASEC, 1999).

69. Finally, considering the issue of effective teaching time, various indicators would have to be taken into account simultaneously, and only some of them are available in the data. SACMEQ data reveal a strongly negative impact of student absence from class. The SACMEQ indicator for teachers being late shows a strongly negative effect as well, while the effect of teachers' absence is more difficult to capture. The latter is measured either as indicated by teachers themselves (PASEC), or in a more general but also more reliable way, as indicated by school principals (SACMEQ). While overall, coefficients show the expected negative sign, for PASEC 2nd grade we get some counter-intuitive results, which may be related to teachers not reporting their absences truthfully.

70. The effective use of time in the classroom (for teaching the subject matter rather than, e.g., try to establish discipline) was covered by the PASEC teacher questionnaires as well, but is based on an extremely subjective appreciation, and therefore difficult to explore in quantitative analysis. Thus, the last point in Table 1 cannot be fully covered by our regression analysis. Nevertheless, SACMEQ regressions show that the cumulative negative effect of those variables for which reliable evidence is available is already considerable. We will return to the question of how this problem could be addressed when we discuss incentive structures in Section 5.

4.4. Methodological refinements and their impact on overall results

71. While, all in all, the available evidence backs the policy priorities suggested in Table 1, the estimated effect of even the most relevant variables is relatively small. Moreover, doubts may remain with respect to the plausibility of certain results, such as the high threshold under which no negative effects of class size can be discerned. Modern micro-econometric theory suggests that there may be several serious problems related to the estimation techniques typically applied, and to the analysis of retrospective data in general. Recent studies, such as Glewwe and Kremer (2006), suggest that the results are plagued with measurement error, endogeneity problems due to reverse causation and selection bias, omitted variable bias, and an inappropriate specification of the functional form due to existing non-linearities. There is no doubt that some of these problems are mitigated through the control for initial knowledge such as measured by the PASEC pre-test at the beginning of the academic year. Nevertheless, even the inclusion of a pre-test variable is not sufficient if innate ability, other unobserved student characteristics, or imprecisely measured family characteristics influence not only the students' starting point, but also their learning process during the year of observation. With respect to class size, for instance, one might argue that parents and teachers may be aware of certain student characteristics such as ability, social behavior, ambition, etc. that are related to expected performance but unobserved by the survey. Students with higher expected performance may be sent to (and accepted by) better schools with the implication that in these schools, class size will increase. Moreover, teachers may use these criteria when making decisions about the allocation of students to certain classes within schools. One could imagine that for children known as calm, diligent and assiduous, classes with a higher number of students would be deemed more acceptable than for notoriously nervous and fidgety children. In both cases, the negative effect of a higher class size would be underestimated because it would be (partially) outweighed by the unobserved and favorable class composition.

72. Recent empirical studies have acknowledged these problems and tried to avoid them through the use of appropriate econometric estimation techniques. The most well-known example is Angrist and Lavy's (1999) paper on the impact of class size on educational achievement in Israel. In this particular country context, the authors are able to use a fixed national regulation regarding the cut-off point for opening up a new class (the so called "Maimonides' rule") as an instrument for class size. As Maimonides' rule is not subject to influences by student or family characteristics, endogeneity problems can be effectively eliminated. As a result of the improved estimation method, the authors obtain a relatively strong negative effect of increased student numbers.

73. Unfortunately, reliable instruments such as Maimonides' rule are not frequently available elsewhere. We are not aware of any comparable study for poor developing countries, such as countries in sub-Saharan African. Moreover, among the studies explicitly taking into account potential endogeneity problems, results are ambiguous as well. Averett and McLennan (2004) provide a comprehensive review of different studies (including meta-studies) and the methodologies used, and are not able to derive any definite conclusion. As reducing class size is an extremely expensive policy measure, it should certainly not be given priority as long as its benefits cannot be clearly established.

74. Another interesting strand of recent empirical literature uses randomized evaluations to establish more reliable estimations for the impact of various measures of educational policy. The MIT Poverty Action Lab currently conducts such evaluations to estimate the effects of various policy interventions in different developing countries. Some of the most well-known studies refer to the

impact of pedagogical resources on primary education achievement in Kenya (Glewwe et al. 2004, Glewwe, Kremer and Moulin, 2000). The main idea is to disentangle the actual effect of the provision of textbooks, wall charts or other instructional materials from unobserved characteristics of students, families or schools, by randomly assigning these additional resources to different schools and students. The outcomes reflect the total (direct and indirect) change induced by the policy intervention, rather than results for a situation where all other factors are held constant. Observations over several years also enable the authors to estimate long-run effects, taking into account parents' reaction to the provision of additional resources.

75. Unfortunately, while the methodology used appears to be clearly superior to earlier retrospective analysis, the results appear to be even less plausible. In fact, the results of the studies quoted above seem to indicate that there is simply no input with any relevant impact on student achievement – with the exception of a small impact of textbooks, but only for the best students in the sample. Based on these results, the authors reject their own prior results from a retrospective analysis, which indicated a strongly significant impact of textbooks and, in particular, wall charts (Glewwe et al. 2004, Glewwe and Kremer, 2006). While one may be able to explain such a missing link in the context of an industrialized country, where overall availability of educational resources is already high at the outset, explanations are difficult to find in the context of primary education in Kenya. This provokes a search for the potential problems these studies may have despite their generally convincing methodology. One remaining issue may be measurement problems: As admitted by the authors themselves, there are some contradictions between the evidence on actual textbook availability from student interviews and the textbook usage survey (Glewwe, Kremer and Moulin 2000: 18). Obviously, if the resource allocation is not fully clear, this will lead to less precise results and lower coefficient estimates.

76. It should also be noted that other studies based on randomized evaluations in developing countries do reveal the expected positive impact of learning materials (see e.g. Jamison et al., 1981 and Banerjee et al. 2003). Additionally, other studies which attempt to replicate a quasi-experimental design using econometric matching techniques also find significant positive effects, notably for textbooks (Frölich and Michaelowa, 2005). Thus evidence from this new strand of literature is far from conclusive, and, as a whole, does not seem to contradict the exposition of Table 1.

77. It remains that despite the significant impact of certain variables, the overall explanatory power of resource inputs on student achievement is rather small. Note that removing the country dummies, the pre-test score, student socio-economic background and all other controls from PASEC 5^{th} grade regressions, and leaving only those explanatory variables directly related to the policy measures discussed in Table 1, reduces the share of explained variance by about 50%. Correspondingly, the R^2 shrinks to about 30%. While this percentage is non-negligible, a major part of the variance in student scores remains unexplained.

78. This calls for a search for additional policy options complementing simple resource provision. In this context, the consideration of institutions and related incentive structures has gained momentum in the recent literature.

5. INSTITUTIONS AND INCENTIVE STRUCTURES

5.1. Non-physical inputs in educational production

79. While the traditional discussion of school inputs focuses on physical goods such as teachers, books, buildings, desks and benches, the "second generation" educational production function literature focuses on more subtle inputs such as accountability, effort and motivation. The idea is that much of the unexplained variation in student achievement may be brought about by differences in these inputs that have previously been largely neglected by the economic literature. Obviously, their relevance has been widely discussed by educational scientists, sociologists and psychologists, but these discussions did not create the interesting link to the input effectiveness literature. This step is required in order to include relevant insights in the discussion about priorities between different educational policies on the basis of their cost-effectiveness.

80. Inputs like accountability, effort and motivation cannot simply be bought or provided. They require incentives, which in turn require appropriate institutions, i.e. a set of rules and regulations which reward socially positive and punish socially negative behavior. While it is difficult to directly observe teachers' and principals' accountability to parents and the society (or to the local community), and whether teachers are effectively doing the work they are paid for, the underlying institutions can be observed. The institutional features most frequently discussed in the economic literature are private-sector participation, decentralization of responsibilities, and the role of standardized exams (see e.g. Wößmann 2004, Pritchett 2004; and see Evers and Walberg 2002, and Peterson and West 2003 for comprehensive collections on the various aspects of enhancing accountability). With particular relevance for developing countries, notably in sub-Saharan Africa, we may add the design of appropriate teacher contracts, and simple control and reward mechanisms like those (potentially) related to the visit of a school inspector.

81. To a certain extent, these institutional features can also be analyzed empirically along with the physical inputs of the education production function. It should be noted, however, that it is often difficult to find appropriate indicators, that many of these indicators do not belong to the standard set of variables covered by student surveys, and that the concrete forms of implementation vary so much between countries that very detailed information is required to make valid comparisons. Moreover, institutional reform is often multidimensional and / or goes hand in hand with changes in traditional resource availability. And finally, in many respects, the institutional framework is implemented on a national basis (based on national legislation), so that only cross-country evidence for a great number of countries can be expected to yield significant results (Wößmann, 2005).

82. In the following, the different institutional features will be discussed one by one, with a brief literature review and an attempt to provide complementary evidence from PASEC and SACMEQ. Given the problems mentioned above, it will not be possible to cover all aspects, and some aspects will be discussed only in the specific country contexts for which they have become particularly relevant, so that the corresponding data is available.

5.2. Private-sector participation

83. The discussion of whether private schools should play a relevant role in the provision of primary education in developing countries has long been discussed only from the perspective of whether it may help to enhance supply where the public education system fails to satisfy demand. The problems highlighted in this context refer to primary education as a public good, which should thus not require private financing, and to the rise in distributional inequalities that private financing might imply. However, both in theory and in practice, private financing and private management can be separated from each other. Privately managed schools in different countries receive widely varying shares of public subsidies. Voucher systems in countries like Colombia and Chile also show that sending a child to a private school does not necessarily imply that the parents have to pay. When looking at private schooling here, we will take it for granted that primary education needs to be publicly financed. This implies that, neglecting potential second order efficiency gains, we can assume that the particular policy measures discussed here have no budgetary implications, but only implications for educational outcomes. With respect to outcomes, we will neglect the discussion of the impact on access to education, and merely concentrate on consequences for educational quality.

84. There are two issues to consider in this context. First, private schools may have a different mode of functioning, with higher autonomy and a more efficient internal incentive system for teachers and other staff. They may also be more responsive to parents' requests, especially if funding depends on enrolment (as in the case of voucher systems, or if teachers are engaged by parents). If private schools are known for particular cultural values, religious beliefs or ideologies, they may also be able to attract more dedicated staff, as well as students and parents ready to engage themselves for their school. Obviously, in the latter case, there is also a risk of conflict generating ideological and cultural segregation (Kremer and Sarychev 2000, Scheerens 2000: 83).

85. Second, enabling private schools to operate in a country may improve the effectiveness of the whole education system, due to increased competition between schools. For this effect to work, some additional conditions need to be met: In particular, there must be a certain level of transparency regarding the schools' educational outcomes, and there must be a regulatory framework and practical living conditions that enable parents to effectively choose the preferred school for their children. In developing countries, the latter is often problematic, especially in remote rural areas.

86. Nevertheless, it is worthwhile to look at the international evidence on private schooling because the estimated effects are often very strong. As the effect of private school competition on the education system as a whole cannot be estimated within a given country, the analysis requires cross-country studies based on a high number of different country observations. Unfortunately, these studies such as TIMSS, PISA and PIRLS primarily cover industrialized countries and a few middle income countries, and (except PIRLS) focus on secondary rather than primary education. According to Wößmann (2005: 146f.) cross-country evidence for industrialized countries consistently shows the superior overall performance of education systems with a higher share of privately managed schools. As an example, he reports evidence from TIMSS indicating that a 50% increase in the share of students enrolled in privately managed schools corresponds to an improved performance equivalent of half-year's learning in both math and science. Looking at the allocation of government resources to public versus private schools, his result is similarly strong: In mathematics, a 1-percentage point increase in the share of public funding going to private rather than public schools raises student scores by about one fourth of the average learning progress within a year. Other examples, primarily for different regions within the United States, are provided in Hoxby (2003a and 2003b).

87. The estimated impact of private schools on average country performance reflects the overall competition effect on both public and private schools and does not imply that the latter demonstrate a higher performance. Do private schools outperform public schools? This question refers us back to the first issue raised above.

88. The evidence is less conclusive here. On the basis of PISA results for industrialized countries, OECD (2001: 211) reports that in 14 out of 17 countries, private schools reached a higher average achievement than public schools. Positive effects have also been reported for various developing countries, e.g. for Indonesia (James, King and Suryadi, 1996; Bedi and Garg, 2000), for Colombia and Tanzania (Cox and Jimenez, 1991), and for Chile (Mizala, Romaguera and Farren, 2002). However, other authors do not find evidence for any systematic relationship (see e.g. Somers, McEwan and Willms, 2004 for Latin America; Preuschoff and Weiß, 2004 for Germany; and the discussion in Scheerens, 2000: 84).

89. Generally, the problem is that private schools often attract the best students or the students with the highest socio-economic background, and if these factors are not adequately taken into account, the impact of the private management structure tends to be greatly overestimated. Moreover, certain variations in results for different countries are highly plausible since "privately managed" may imply very different degrees of actual autonomy, depending on national regulation for private schools. Public schools in some countries may enjoy more autonomy than private schools in other countries. Indeed at cross-country level, the correlation between any kind of autonomy indicator and the prevalence of private schooling is very low. Using the PISA data for OECD and some middle and lower income countries, none of the available 24 measures of school and teacher autonomy turns out to be significantly related with private schooling at the 5% level (OECD and UNESCO-UIS, 2003: 204). An interesting individual country-case is PISA champion Finland, which reaches high scores on relevant indicators of both school and teacher autonomy, while the share of privately managed schools is negligible.

90. Looking at our data for sub-Saharan Africa, only very crude information is available on whether schools are private or public. This information is derived from a simple yes-or-no question to principals. It is available for all SACMEQ countries, so that it could be included in the regressions presented in Tables 2 and 3. The regression tables show a relatively high coefficient for literacy, indicating that students in private schools outscore students in public schools by 7-10% of a standard deviation in achievement scores. However, this relationship is significant only in one of the two regressions and not significant at all for mathematics. Moreover, if we try to further reduce sample selection bias by moving into (remote) rural areas where self-selection into specific schools is impossible, no impact can be found either.

91. In PASEC, the information is available only for Togo. Both for 5^{th} and for 2^{nd} grade we find similar results: Students in private schools show higher overall performance, but this performance advantage vanishes when socio-economic background and initial knowledge as measured in the pre-test scores is adequately controlled for. Moreover, even without any controls, in rural areas where the possibility of school choice is considerably reduced, private schools do not show any advantage.

92. In contrast, in a more elaborate study, Lassibile and Tan (2003) complement the PASEC data base for Madagascar with external information on school types. The authors control for self-selection using the Heckman two-step procedure and do find some positive impact of private schools.

93. Independently of whether schools are publicly or privately managed, effective competition can be encouraged by a link between funding and enrolment in a particular school. Considerable evidence is available for the functioning of corresponding voucher systems in Latin America. For the extensive voucher system in Chile, evidence is not yet conclusive (Glewwe and Kremer, 2006; Wößmann, 2005). One reason may be that it does not truly cover all cost. For Colombia, King, Orazem and Wohlgemuth (1999), Angrist et al. (2002), and Angrist, Bettinger and Kremer (2004) find a strong positive effect on participants. According to Angrist et al. (2002), participants outscored their peers who did not benefit from the vouchers by an equivalent of the learning progress during a full academic year. Voucher eligibility was determined by a lottery thereby generating a natural experiment which lets these results appear particularly reliable.

94. The problem remains that any kind of competition will be very difficult to establish in remote rural communities. This is a relevant constraint for large parts of sub-Saharan Africa. Parental choice between different alternative educational institutions is a necessary condition for competition to take place. In many rural communities, the opportunity of choosing a school is limited to those parents who can send their children to relatives in other regions of the country. And even where this option exists, it may have little relevance for primary school aged children, but rather for secondary or tertiary students. Finally, even in cities where school choice is theoretically possible, it is often not possible in practice, especially for poorer families.

95. Other important preconditions of functioning competition are not reached in sub-Saharan Africa, either. The issue of transparency with respect to schools' educational outcomes was already mentioned above. Currently, transparency about existing schools and reliable information on quality is so limited that parents cannot base their decisions on any objective criteria (see Sections 5.3 and 5.4).

96. In addition, it should be kept in mind that the above discussion was based on the assumption of full public cost coverage. In practice, this may be difficult to ensure for at least two reasons: First because there are many hidden costs (like unofficial fees requested by school administrations, or the opportunity cost of traveling longer distances, in particular for children expected to also contribute to household activities). And second because there is typically not enough control over the administration at national, regional and local level to make sure that subsidies actually reach the designated recipients (and on time).

97. Several conclusions can be drawn from the above discussion. First, no clear answer can be given to the question whether private or public schools perform better. While it is highly probable that certain incentive based management structures increase school performance, they are not necessarily implemented more easily in private schools than in public schools. Second, competition between schools is more a matter of public regulation than a matter of private versus public schools, i.e. competition may well be encouraged between public schools alone (see Hoxby, 2000a). And third, true competition with a plausible chance for success in terms of raising educational quality requires certain preconditions difficult to meet in the African context.

98. Trying to promote competition while ignoring these necessary conditions entails the risk of higher segregation and distributional inequalities without reaping the expected efficiency benefits. While, in an adequate environment, enhancing competition between schools can strongly increase efficiency, it seems that policy making in sub-Saharan Africa should focus on alternative institutional reforms. Such reforms are often targeted at issues which have so far been characterized as preconditions of successful competition (e.g. transparency, autonomy). However, they also represent important reform measures for their own right, and are therefore relevant here, even if full fledged competition between schools is not considered as an advisable policy objective for the time being.

5.3. Decentralization of responsibilities

99. Decentralization of responsibilities to local communities and increased autonomy of schools may be one way to change the incentive system within schools, even in rural areas, and within the educational administration as a whole. In an industrialized country context, decentralization of responsibilities typically refers to school autonomy. Only if schools have some freedom to change their policies can they be innovative and improve service provision.

100. The analysis in OECD and UNESCO-UIS (2003: 202ff. and 364-367) shows that autonomy is a multidimensional concept and that the different dimensions are more or less relevant for student achievement. Sizeable positive correlations with student achievement can be observed especially for schools' responsibility for budgetary allocations within school, the choice of textbooks, disciplinary policies and courses offered. The correlation coefficients range from 0.65 for budgetary allocation to 0.37 for establishing disciplinary policies (p. 203).

101. Clearly, not all dimensions of autonomy are positive. For instance, the cross-country correlation between average student achievement on the one hand and teacher autonomy to approve students' admission to school on the other hand, is negative with a correlation coefficient of -0.42 (p. 367). In sub-Saharan Africa, some remote rural schools are so autonomous that they actually feel abandoned by public authorities. They do not receive any support; public authorities sometimes do not even know that they exist and consequently do not care for what they do. This is certainly not the type of autonomy which should be aspired to.

102. Again, a necessary condition for school autonomy to successfully enhance student achievement is transparency, so that students and parents are able to evaluate the outcomes of their schools' decision making. It should be noted that even industrialized countries face considerable difficulties in this respect. If we now imagine an environment in which many parents are illiterate and have never had any contact to segments of the labor market requiring formal education, it becomes clear that the challenge is much greater (see also Tekleselassie, 2005).

103. It may be for this reason that in a developing country context, the discussion about the decentralization of responsibilities typically focuses on the involvement of local communities in educational decisions and monitoring. This can be considered as a necessary first step, or at least a relevant complement, to increased school autonomy. In principle, local communities should have the best knowledge about students' needs, they should have a high incentive to monitor teachers, principals, and administrative officers, and they should be best placed to do so because they can most directly observe the effort of the former (Glewwe and Kremer, 2006: 44).

104. Glewwe and Kremer (2006: 44ff.) review the available evidence on corresponding decentralization efforts in developing countries. Most notably, a local community empowerment program in Uganda led to regular monitoring of grant transfers, which in turn increased the share of government resources effectively reaching the local schools from 20% in 1995, to 80% in 2001 (Reinikka and Svensson, 2003). Positive effects have also been reported from community schools in francophone Africa, in which the parents themselves selected the teacher and paid his or her salary (see Section 5.4). If a lump-sum payment were made by the national government to the parents and communities in the first place, the incentive system could be maintained without the currently prevailing negative impact on distribution (the poorest communities have to finance their schools while the better-off receive full public financing) and on demand (given the strong externalities of

primary education, demand will be socially suboptimal if it is privately financed). Similar arguments apply to parents' participation in the construction and maintenance of school buildings and classrooms, or to their contribution to school equipment. As noted by Miguel and Gugerty (2005), it is necessary to consider the concrete local environment in order to design the most successful decentralization strategies.

105. In SACMEQ and PASEC, community or parental involvement is measured primarily in terms of contributions to school equipment (SACMEQ), or the director's appreciation of how easily they could be mobilized for such purposes (PASEC). These variables are positively significant in all SACMEQ regressions as well as in PASEC 5th grade mathematics regressions, as long as the initial knowledge of students is not controlled for. However, as the variable's significance cannot be shown in any of the regressions including the pre-test score, our results might again simply reflect sample selection bias. Students with better scores tend to have parents who can be more easily mobilized for school issues and who also have the resources required.

106. SACMEQ further includes a variable indicating whether parents and / or the community are involved in the payment of exam fees, additional teacher salaries or bonuses. This variable is fully insignificant, probably because the realities reflected in a yes or no to this question can vary considerably - from a veritable influence on teacher pay thereby creating accountability, to obligatory payments for certain services.

107. Finally, in earlier PASEC regressions on five countries, the existence of an active parent-teacher organization was considered as an additional variable for parental involvement. Results were similarly weak and insignificant (Michaelowa, 2000: 31). As there is no other, more convincing indicator to analyze the issues of decentralization and school or teacher autonomy, our data do not allow us to draw any firm conclusions with respect to this topic.

5.4. Standardized exams

108. Just as decentralization and community or parental participation in school and teacher related decision making, standardized exams are an institutional feature that may facilitate monitoring by parents and communities. If results are published, standardized external exams increase the accountability of teaching staff and school management, and provide the relevant information to set up a more refined performance based incentive system (Bishop and Wößmann, 2004). According to Bishop (1997) and Wößmann (2005), results from international student achievement tests like PISA and TIMSS suggest that the average achievement gains due to the introduction of standardized external school-leaving exams are large, possibly larger than a full grade-level equivalent. While prior results are based primarily on industrialized or middle-income countries, Wößmann indirectly suggests that developing countries may also benefit from such an institutional reform.

109. In most African countries today, partly as a heritage of colonial education systems, standardized primary school leaving examinations already exist (Kellaghan and Greaney, 2003: 5). There would therefore be little extra cost, if the results of these exams were effectively used for monitoring purposes (Bernard, 2004: 9f.). Currently, especially at primary level, this is rarely the case. The first step would be to create a national data base with a school and region based break-up of results. Potentially, even development over time could be monitored, but this would require adequate anchor items for the tests in different years and an adjusted evaluation mehodology. In any case, with or without comparisons over time, the creation of such a data base would be valuable to inform

national level policy making, especially in the context of the elaboration of national sector strategies. Access to this kind of information would allow policy makers to identify problem areas and to review the allocation of resources (Kellaghan and Greaney, 2003: 6ff.).

110. It may be objected that the use of these tests as a monitoring tool is suboptimal because they are designed primarily as an instrument of student selection for further studies. Test items may thus go beyond primary school curricula and evaluate competencies required for successful studies at secondary level. Moreover, the general quality of existing tests is often reported as rather weak, with a focus on the recall or recognition of factual knowledge rather than on competencies needed for real life. Nevertheless, the majority of test items usually covers the curriculum and even if the test design is not optimal, this will only make their results somewhat less precise, but by no means useless.

111. The issue becomes more complex when standardized exams are used for comparisons between schools and performance based incentive systems (Kellaghan and Greaney, 2003: 14f.). In this case, the data already collected are clearly insufficient. In order to make sensible comparisons between schools which can be evaluated with respect to teacher performance and the management performance of principals, one has to be able to control for the school environment and the students' socio-economic background. However, the additional information required might be relatively easy to collect. For the case of primary education in Germany, Jürges and Schneider (2005) find that a simple complement to the tests in the form of a short student questionnaire would be sufficient to provide the relevant controls. Information collected from principals and teachers (who might have strong incentives to cheat) would not be necessary. As schools tend to be selective in that they hinder less performing students to actually take the exams in the first place (via grade repetition and pressure for school change or drop-out), the share of students taking the exam should also be assessed.

112. In order to create transparency and to allow local backstopping and incentive systems to work, the evaluation results net of the effect of socio-economic background and prior selection must be accessible to parents and local communities. In principle, this information would then also enable local, regional or national authorities (wherever the responsibility may lie) to establish performance-based salary scales for teachers. While both economic theory and various empirical studies suggest that the general level of teachers' salaries does not have a relevant impact on their performance (for a discussion, see e.g. Michaelowa, 2001), higher pay linked to higher performance can be shown to have a significant effect (Lavy, 2004 and 2002).

113. If several teachers are involved in the final result because the students went through various grades with different teachers, group incentives may be the most appropriate instrument. Repeated testing to assess added value by individual teachers is not necessary in this case. Moreover, group incentives for all teachers (and the principal) in a given school help to sustain a positive working climate as opposed to a situation of rivalry among colleagues which has been observed, for instance, in the context of the individualized pay scheme in Ethiopia (which further encouraged rivalry through the introduction of peer assessments among teachers) (Tekleselassie, 2005: 625f.).

114. It should be noted that cheating may remain a non-negligible problem. Such problems have been reported even for industrialized countries (see e.g. Jacob and Levitt, 2003, and Gay, 1990). With a focus on Africa, Kellaghan and Greaney (2003: 18) distinguish between various forms of common "examination corruption". Examples are: intended leakage of exam content, collusion between students and exam supervisors, side-payments for examiners, substitution of scripts, intimidation of markers, and the falsification of data files. If students are requested to answer additional questions on socio-economic background they may face pressure to indicate that they all are very poor. Cheating problems already exist today, but must be expected to increase when tests become

relevant, not only for students' further educational perspectives (as they already do today), but also for teachers' career prospects and salaries. The fight against the misuse of standardized exams therefore needs to be reinforced.

115. It should be noted that transparency of learning outcomes and the introduction of performance-based schemes may face political opposition, especially by teacher unions. However, if the control for background variables is fair, the system well-explained and the overall impact on salaries positive rather than negative, political cost can be kept to a minimum. Obviously, the direct financial implications depend directly on the impact of the resulting pay scheme on average salaries. Moreover, the cost for controls required to counter cheating can be expected to increase. Costs for all other parts of the intervention, such as the establishment of the statistical database, are relatively low or negligible.

116. Unfortunately, neither PASEC nor SACMEQ provide information on performance-based pay schemes. Moreover, there is no information about any use being made of existing standardized exams to create an appropriate statistical basis for such an endeavor.

5.5. Teacher contracts

117. The issue of teacher salary schemes is directly related to the general question of teacher contracts. Since the late 1990s, this question has become an issue of concrete policy reform in many African countries, so that data and initial analysis is available, which can be discussed in more detail here. This discussion will be based primarily on the results presented in Bourdon, Frölich and Michaelowa (2006).

118. "Contract teachers", often also called "voluntary teachers", are teachers who are no longer engaged in traditional civil servant positions, but on the basis of fixed-term contracts. These contracts typically imply considerably lower salaries and a sharply reduced duration of professional training. While in most countries, these teachers have been employed by public authorities, parents of school-aged children have often also resorted to private initiatives, opening their own schools with privately engaged teachers, on contracts at considerably lower rates than those foreseen in the public sector. Both groups of teachers fall in the category of "contract teachers".

119. Financially, these programs have led to a considerable relaxation of budgetary restrictions, which enabled the countries concerned to hire a large number of additional teachers. Table 2 presents some evidence on the spread of these programs and the savings involved with respect to salaries.

120. Overall, the employment of contract teachers has led to a boost of primary enrolment which has already become evident in recent national education statistics. At the same time, stakeholders in the education system generally fear an important loss in education quality. They argue that the relaxation of professional training requirements and the loss in teacher job satisfaction, supposedly implied by reduced salaries and less job security, will necessarily lead to a collapse of the education system in the long run.

121. Theoretically, we have to distinguish between different potential effects: (1) an incentive effect of the teaching contract, (2) a selection effect (changed demand for and supply of new teachers), and (3) a dynamic effect.

Table 2 Distribution of primary teachers according to their statute

Country	Teacher remuneration *(relative to GDP per capita)*			Distribution across statutes		
	Civil Servants	Contract teachers		Civil servants	Contract teachers	
		Public	*Private*		*Public*	*Private*
Benin (2002)	5,2	2,1	1,3	55%	16%	29%
Burkina Faso (2002)	5,8	5,6	2,2	64%	24%	12%
Cameroon (2002)	5,3	1,4	0,8	35%	20%	45%
Chad (2002)	8,2	-	2,3	32%	0%	68%
Congo (2003)	2,4	0,9	0,6	42%	4%	54%
Côte d'Ivoire (2001)	4,8	-	-	87%	0%	13%
Guinea (2000)	3,5	1,1	-	52%	30%	18%
Mali (2000)	5,8	1,5	0,9	71%	8%	21%
Niger (2000)	8,9	3,5	-	46%	50%	4%
Senegal (2003)	5,7	2,6	-	44%	41%	15%
Togo (2001)	6,4	3,3	1,3	35%	30%	35%
Mean	**5,6**	**2,4**	**1,3**	**51%**	**20%**	**29%**

Notes: *Public*: under contract with public authorities; *Private*: under contract with parents or local communities. On average, in OECD countries, primary school teachers' salary corresponds to 1.3 times GDP per capita; in Germany and UK the factor is 1.5, and in the US, it is 1.1 to name but a few examples. (However, one should be cautious with direct comparisons given the scarcity of human capital in developing countries.)

Source: World Bank, Africa Region (quoted from Bernard, Tiyab and Vianou, 2004: 5), OECD (2002: Table D6.1).

122. The first of these, i.e. the incentive effect, is frequently overlooked in the political debate, but is of particular interest in this section on institutional reform options. The direction of this effect cannot be determined theoretically. On the one hand, the unfavorable conditions of new teacher contracts could be regarded as unfair and demotivating, and short-term contracts could prevent personal investments in pedagogical training and school specific human capital. On the other hand, for contract teachers, further employment prospects depend on performance and, among other things, parents' satisfaction, so that from this perspective, the contract statute could be expected to have a positive incentive effect.

123. As far as the selection effect is concerned, the changed employment conditions could lead to a different composition of teacher candidates. On the one hand, we would expect a lower number of highly skilled candidates due to the inferior contract conditions. On the other hand, the frequently reduced entry requirements could reduce entry costs and increase the attractiveness of (temporary) teaching positions. The higher demand for teachers would lead us to expect a lower quality of the marginal (newly employed) teacher.

124. The dynamic effect, finally, refers to a potential change of teacher behavior over time. In particular, the inferior contract conditions may induce a reduced retention period of teaching staff. This effect could lead to a different distribution of job experience before and after the reform.

125. PASEC data allow us to estimate a combination of the overall incentive and selection effect without, however, enabling us to distinguish between these effects or between any of the underlying indirect effects. At the same time, a consideration of the dynamic effect cannot be provided

so far because the distribution of job experience in the teaching population has not yet reached its equilibrium due to the relatively recent introduction of the new program. A recent discussion of the effects of salary changes on teacher attrition in the US suggests that professional development schemes with the perspective of higher salaries after some years of experience may reduce attrition especially of female teachers (Imazeki, 2005). However, differences in the socioeconomic and labor market context must be expected to be highly relevant here, so that we will have to wait for further evaluation results of the contract teacher program in African states before we are able to draw any firm conclusions on dynamic effects. We therefore confine our discussion to the (joint) incentive and selection effect.

126. In the regressions presented in Tables 2 and 3, teachers' employment on a contractual (non-civil servant) basis has been simply introduced as a dummy variable. The pooled effect for the eight countries is not conclusive: In 2nd grade, it appears to be negative, but significant only at the margin and only in one regression; and in 5th grade, it seems to be positive, but significant only in those regressions in which the pre-test score is not controlled for. The magnitude of the effects is similar, with opposite signs.

127. As the concrete specification of contract teacher programs, including academic entry requirements, the type and duration of professional training, salary schemes and the predominance of public or private employers (public authorities versus parents and local communities) vary considerably between countries, the difficulties to find consistent overall effects do not come as a surprise. We will therefore have a look at individual country evaluations based on the same data, which have been carried out by different authors. PASEC (2003, 2004, 2005b and 2005c) provides some regression estimates for Guinea, Togo, Niger and Mali. Bourdon (2005) reproduces the results for Mali using propensity score matching. Vegas and de Laat (2003) reexamine the PASEC data for Togo. Michaelowa (2001, 2002) analyzes the evidence from earlier PASEC surveys in Burkina Faso, Cameroon, Côte d'Ivoire, Madagascar and Senegal. And finally, Bourdon, Frölich and Michaelowa (2006) use a propensity score matching approach to reexamine the data for Niger.

128. In their case study of Niger, Bourdon, Frölich and Michaelowa (2006) find little evidence of any significant overall effect of the contract teacher status. At first glance, there appears to be a significant deterioration of education quality for grade 2, but this effect vanishes when the authors restrict the sample to younger teachers so that they can control for job experience. Comparing the different specifications, it appears that the negative effect of contract teachers in the full sample is largely an artifact of their limited job experience. PASEC (2005b) finds a significant negative effect in 5th rather than in 2nd grade. However, the simple (clustered) linear regression analysis used seems to lead to somewhat less robust results - as indicated by the coefficients and varying significance levels of various specifications presented in the appendix of their study.

129. Similarly mixed results have also been found for Togo and Guinea. In Togo, the effect of contract teachers on student achievement was shown to be significantly negative for both grades, and in Guinea, a significantly negative impact was found for grade 2 (PASEC, 2005b, 2004, 2003, and Vegas and de Laat, 2003). At the same time, the PASEC studies suggest that this negative impact may be driven to some extent by the impact of reduced teacher training. In fact, if the sample in Togo is split between teachers with and without initial training, the contract teacher statute is no more significant in the former. While Guinean teachers all receive some training, its structure and duration was reformed from the first to the second cohort, and indeed, the second cohort does not show any significantly negative (but partially even a slight positive effect) on students' learning.

130. Interestingly, in other countries, contract teacher programs have sometimes led to an obvious improvement of results. A positive impact of contract teachers, significant for grade 2, was found in Mali, and this result was rather robust to the use of different estimation methods and model specifications (PASEC, 2005c; Bourdon, 2005). Moreover, an joint study on the five countries Burkina Faso, Cameroon, Côte d'Ivoire, Madagascar and Senegal for the mid-1990s (a time period in which contract teachers were still rather rare) also found a significant and positive effect (Michaelowa, 2001b). As opposed to Guinea, Niger and Togo, where at least about half of the contract teachers are employed by public authorities, in all these cases the majority of contract teachers were engaged by parents or local communities. It seems that these teachers are relatively highly motivated and also miss classes less often than their peers employed in the civil service (Michaelowa, 2002).

131. Chaudhury et al. (2005: 4) examine the effect of contract teacher programs on absence rates in Bangladesh, India, Indonesia, Peru and Uganda. They do not find the positive effects reported above. However, except for India, only public schools were taken into account, so that teachers directly engaged by parents and local communities are excluded from the sample.

132. All in all the available evidence suggests that the new type of contracts may have a relevant impact on incentives. While a teacher engaged as a civil servant faces hardly any incentives to increase his or her effort, the situation is different for teachers engaged on a short-term contract which may or may not be prolonged. However, this feature alone does not seem to drive the overall effect, and the specific conditions of the relevant national programs need to be considered with care. Most importantly, it seems that if teachers are directly engaged by parents or the local community, they will feel the recognition of their effort and they can also be very directly held responsible for their work. The evidence available so far suggests that this may outweigh lower pay and adverse working conditions.

133. As contract teacher programs have been shown to be highly relevant tools in progress towards universal primary education, it seems to be important to further examine under which conditions their positive impact on enrolment can go hand in hand with an equally positive effect on education quality, and under which conditions these effects remain stable over time. While contract teacher programs appear to have been worthwhile policies in all countries, the quality-quantity trade-off apparent in some countries might still be considerably reduced. This requires further research based on a distinction between the different types of contract teachers.

5.6. School inspection

134. The final institutional feature, discussed here because of its common prevalence in most African education systems, is a simple advice and control mechanism. It has been highlighted earlier that effective control mechanisms appear highly relevant in the African context, to a certain extent simply to ensure teachers' presence at school. In a regression analysis of five PASEC countries, Michaelowa (2001) finds a strong and positive effect of an inspector's visit at some point during the ongoing academic year. While there appears to be a negative effect of such control measures on teacher job satisfaction, the overall effect of the inspector's visit on student achievement remains positive.

135. However, the available evidence is again not fully conclusive. Bernard (1999) indicates that inspector's visits may be more effective if they fulfill an advisory rather than a control function. Clearly, the effectiveness of the inspector's visit also depends on (i) the mandate he or she has,

including the power to attach positive or negative consequences to the results of the inspection, (ii) the loss of teaching time due to exchange between the inspector and the teachers or the inspector and student families, and (iii) the extent to which collusion between teachers and inspectors must be expected.

136. In our data for eight PASEC countries, the positive effect of inspections is found significant for 5th grade mathematics, and close to significant (or, in one case, significant at the 10% level) for 5th grade French. The magnitude of the effect is not huge but non-negligible, as it corresponds to up to about 10% of a standard deviation in student scores. Just as in Bernard (1999), this result cannot be replicated for 2nd grade.

137. For SACMEQ countries, the inspection variable is insignificant. To a certain extent, this may be explained by the fact that, asking for the frequency of visits within the previous calendar year, the measure is not necessarily related to the academic year concerned. Other teachers may have been concerned, so that the indicator for effective teacher control becomes less precise. In Malawi and Mauritius where evaluations took place later than in other SACMEQ countries, this problem might be particularly strong. However, it should also be noted that the roles of inspections in francophone and anglophone Africa are not the same, so that it could be worthwhile to examine in some more detail the concrete incentives related to this control mechanism in the different regions.

138. Overall, it appears that incentive based approaches may have a relevant impact in sub-Saharan Africa. While competition between schools should not be an immediate priority, more effective monitoring and control, and direct incentives for teachers based on accountability towards parents and local communities can be implemented straight away. The political appeal of these measures is that they may bring along considerable improvement in student learning without relevant direct financial implications. In some cases, budgetary implications can even be positive. At the same time, these measures may bring about some political cost as they may face opposition by relevant stakeholders such as teacher unions. These political costs can be minimized by creating appropriate policy packages which, for instance, jointly introduce more autonomy and control.

139. It remains that the impact of institutional change is not more easily captured empirically than the effect of physical school inputs. On the contrary, finding appropriate indicators and distinguishing between various effects induced by one and the same policy intervention appear to be even more challenging here. Moreover, with respect to developing countries, little research has been done in this area so far. While promising approaches exist, there is no magic bullet, either.

6. A CONSIDERATION OF PEER EFFECTS

140. A final issue which has received much consideration in recent empirical studies analyzing student achievement is the role of peer effects. Most studies concentrate on the effect of ability grouping or the effect of grouping by socioeconomic background (see e.g. Hoxby, 2000b; Angrist and Lang, 2004). The emphasis is then generally on the appropriate econometric technique to distinguish proper peer effects from other effects such as the effect of one's classmates' socio-economic background on the financial support provided to the school.

141. While these papers provide valuable insights into general issues of distribution and learning effectiveness related to it, they are not directly discussing educational inputs and thus beyond the scope of our analysis. However, although this is less frequent in the literature, peer effects can also be considered directly in the context of school inputs. And in fact, the appropriate consideration of peer effects may have a relevant impact on estimation results with respect to their overall effect on student learning. Moreover, the consideration of such input related peer effects may provide an indication of how the inputs in question should most efficiently be distributed within classes and between schools. This will be demonstrated at the example of textbooks for French and math in francophone Africa.

142. For PASEC, the regression results on textbooks in Tables 2 and 3 are split between the effect of a student's own textbook possession and the availability of textbooks among his or her classmates. This represents a very simple way to distinguish between the direct and the peer effect. While the peer effect is generally insignificant in grade 2, it is significant or close to significant in all regressions for 5^{th} grade French, with coefficient estimates being 5-10 times as high as coefficient estimates for the direct effect. In contrast, for 2^{nd} grade French, direct effects appear to be almost twice as high as the peer effects. This gives us an initial indication that if resources are insufficient to provide reading books to all students, upper grade students rather than school beginners should be requested to share their books. As already mentioned in Section 4.3., personally owning the reading book (i.e. also being able to regularly take it home) seems to be relatively more important for students in initial grades.

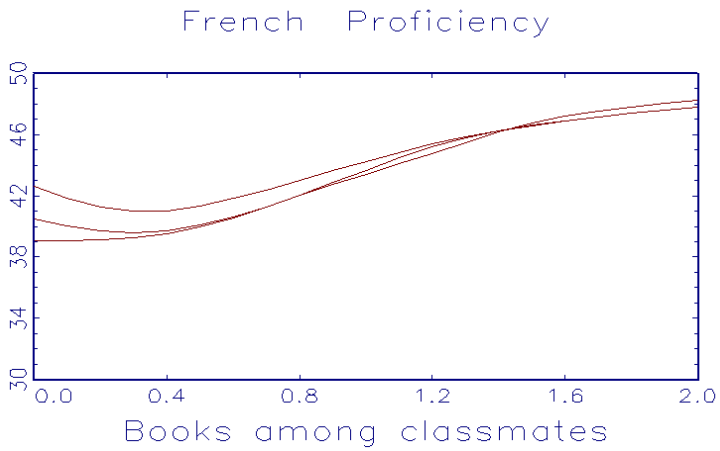
143. As textbook peer effects can be shown to be frequently significant and with high coefficients, it also becomes clear that leaving them out of the regression model will generally lead to a serious underestimation of the impact of textbooks as a whole. While the individual effect and the peer effect are positively correlated, the coefficient in a regression in which only the direct effect is taken into account will typically not capture the overall effect.

144. In the literature, production function models vary in the way in which the textbook variable is included. If the share of books in the classroom is considered (like in our SACMEQ regressions), the effect of classmates' textbooks is taken into account, even though it is not separated from the direct effect. But seriously downward biased estimates must be expected, when only private textbook ownership is entered into the model. To a certain extent, this problem may explain why the literature appears to be somewhat inconsistent regarding both the size and the significance of the effects.

145. Frölich and Michaelowa (2005) examine the textbook related peer effects for PASEC 5th graders in some more detail using a non-parametric matching estimator. This estimation technique has the advantage not to rely on any assumptions about the underlying functional relationship between the different variables of interest. This implies that, apart from the estimation of the overall (direct and indirect) effect of textbooks on achievement, potential changes of this effect depending on the initial distribution can be observed.

146. As there may be cross effects of the book in one discipline on achievement in the other, both French and math textbooks are included in the analysis for each subject. Figure 2 shows expected French achievement as a function of own books and classmates' book coverage.

Figure 2 French proficiency and the availability of textbooks for French and math



Source: Frölich and Michaelowa (2005: 26)

147. Student achievement in French is plotted on the vertical axis while the horizontal axis shows textbook availability among the student's classmates in the range from zero (none of the classmates have any book) to two (all classmates have both the French and the math book). The Figure shows that student achievement rises considerably from the left to the right, indicating that the number of books among classmates has a fairly strong effect.

148. Figure 2 shows three lines, each of which depicts a student with a different number of own books. The lowest line represents a student without any book and the top line represent a student with both books. As the distance between these lines is almost negligible as compared to the achievement differences from the left to the right of each line, the direct effect represents only a small part of the overall effect. In fact, quantifying this relationship yields the result that the peer effect makes up about 90% of the overall effect. This result is close to identical in mathematics (not shown here), and for both subjects in a rural sub-sample of the student population.

149. To understand how the peer effect can be so high, we need to understand how it works. Obviously, children can share books, and if two children share one book the peer effect can be up to 50% if there is no specific advantage of owning the book. If three children share a book, the peer effect could be up to 66%. Sharing alone makes it difficult, however, to explain a peer effect as high as 90%. It is therefore plausible that two additional mechanisms may be at work: First, those students having textbooks could show a higher performance which in turn spills over on their classmates

(traditional peer effect), and second, teachers could change their teaching practice depending on overall textbook availability.

150. The latter is consistent with the slope of the curve in Figure 2. The slope is steepest in the middle indicating that the effect of an additional textbook is greatest when a certain number of books already exists, but before the class is already almost fully equipped. It is well conceivable that teachers will actually use the textbooks only if there is a minimum general coverage. However, if almost every student already has a book, adding additional books will no longer have much impact on teaching practice.

151. The conclusion for policy makers should be that the distribution of resources between schools also matters. Providing two schools with textbooks to cover 50% of the students in each appears to be more effective than providing full coverage in one school, and no books at all for the other. In contrast, the distribution of textbooks within the class does not seem to matter much for average performance since students share the books anyway.

7. CONCLUSIONS

The optimal policy package reconsidered

152. After reviewing the evidence on the impact and cost of traditional school inputs, the more recent literature on immaterial inputs such as accountability, effort and motivation, and finally the relevance of input related peer effects, no easy recipe can be given for a successful reform of education systems in sub-Saharan Africa. Generally, it seems that policy reform should aim at a mix of an introduction of (carefully adjusted) incentive based mechanisms for teachers and principals, and the provision of basic school inputs, especially learning materials.

153. The intensive discussions about education quality before and after the ADEA Biennial Meeting 2003 included the attempt to define a set of essential inputs ensuring basic education quality at affordable cost. With respect to physical inputs, the ranking of policy measures has not changed since then. While new studies using interesting empirical research methods based on natural experiments and randomized evaluations have become available, their outcomes are either in line with earlier results or based on very limited evidence. Moreover, the outcomes of these studies appear to be similarly inconsistent across studies as prior retrospective analysis.

154. Investment in pedagogical resources, especially textbooks for the core subjects of reading and math, can still be considered as an efficient policy measure. If budget constraints are very strong, one book may be provided only to every second student, especially in higher grades where taking the book back home does not seem to be as important as for very young students.

155. Another priority should be the reduction of repetition rates. Repetition induces high cost because the system has to cope with an increased overall number of students. Moreover, repetition increases early drop-out. And finally, the effects of repetition on student learning have consistently been shown to be negative, rather than positive, at least in the long run. An example from India shows that there may be other, much more promising measures to help disadvantaged students with particular learning problems.

156. With respect to teacher education and training, the focus should be on quality rather than duration. In anglophone Africa, where the duration of formal education and teachers' subject matter knowledge are much more clearly correlated than in francophone Africa, longer education for teachers significantly enhances student learning. However, the effect is only moderate in size and has to be carefully weighed against the equally high cost generally involved with salaries for teachers with higher educational attainment. Similar considerations are in order with respect to pre-service and in-service training. From a cost-benefit perspective, short but well designed and practice oriented programs appear to be most promising.

157. Finally, it appears highly relevant to ensure the maximum use of formal instruction time for effective teaching. Double shift teaching seems to have a detrimental impact in this respect. As there is ample evidence for a rather modest negative impact of high student-teacher ratios, double shift teaching should generally be avoided.

158. Effective teaching time can also be increased by improving students' attendance. Apart from the well-known requirement of adjusting the academic year to harvesting seasons, attendance can be increase by simple health care measures. In this context, de-worming has been shown to be particularly cost-effective.

159. And last but not least, effective teaching time can be increased by reducing teachers' absences. In some cases, simple administrative measures like the reorganization of teacher remuneration (so that teachers do not need to collect their pay from a far away district officer) may be very effective. In general, however, more effective control mechanisms seem to be required.

160. This creates the link to the relevance of functioning incentive systems. Notably, in several countries, teachers on non-civil servant fixed term contracts have been shown to miss their classes significantly less often then their colleagues. While contract teacher programs combine various features with partly contradicting consequences for student achievement, preliminary evidence suggests that the incentive effect works best if these teachers are employed by parents and local communities, rather than by public authorities. Indeed, this should enhance teachers' accountability and parents' incentive for effective monitoring. Theoretically, this system could be generalized by channeling public funds for teacher remuneration via local communities and parents' associations.

161. Other aspects of decentralization and increased local autonomy (both for parents, and for schools and teachers) may also be beneficial for student learning. In particular, any kind of measures to enhance transparency about resource flows and learning outcomes appears to be valuable. This could also be a first step towards even more comprehensive institutional change.

162. At the same time, full fledged competition between schools does not seem to be a realistic approach for primary education systems in sub-Saharan African. For the time being, the focus should be on the optimization of contract teacher programs and on other measures to enhance parents' monitoring capacity.

8. APPENDIX

Evaluation designs, econometric methodology and regression results

163. This appendix presents the details of the empirical results from PASEC and SACMEQ regressions discussed in this study. Moreover, it discusses the regression methodology and differences in the survey design which should be kept in mind when differences between the two groups of countries (francophone and anglophone Africa) are analyzed. Some of the issues discussed in this appendix have already been briefly mentioned in the main body of the text, but will now be complemented with additional details and explanations.

8.1. Regression methodology

164. For both sets of countries, the dependent variable used in our regressions is the test score in literacy and mathematics. This test score is coded on a scale with mean 500 and standard deviation 100 for SACMEQ. For PASEC, the simple percentage of correct answers (0-100%) is used with a cross-country average of: 37% (French) and 39% (math) and a standard deviation of 20.5% for 5th grade; and an average of 47.2% (French) and 45% (math) and a standard deviation of 27% and 26% respectively for 2nd grade.

165. All countries within each country group are considered jointly in a single regression. This has the considerable advantage that, due to the high total number of observations, even very small effects can be distinguished. Country differences are captured by country fixed effects.

166. We use two different econometric models to estimate the education production functions. For both SACMEQ and PASEC, model type A is the usual hierarchical linear (or multi-level) model with school random effects (for textbook expositions see e.g. Raudenbusch and Bryck, 2002 or Goldstein, 2003). Estimations are carried out with generalized least squares (GLS) with the exception of SACMEQ regressions because the availability of sampling weights makes maximum likelihood estimation (MLE) computationally more attractive in the multi-level framework.

167. Model A has the advantage of providing a clear distinction between the explanations of the variance within and between schools. However, the true standard errors may be underestimated if sub-clusters exist (such as classes within schools for SACMEQ - see below; or groups of students living in the same area or doing their homework together), which lead to a variance structure different from the one explicitly specified. As a robustness check, we therefore introduce a model type B using the Stata survey sampling command. The underlying estimation procedure takes into account the nested structure of the data without separate computations of the variances at the different levels. However, it has the advantage of being robust with respect to any type of sub-clustering, as it uses the Huber-White sandwich estimator for the estimation of the variance-covariance matrix. For details, see the Stata handbook on survey sampling (StataCorp., 2003: 38f.). For a comparison of the different methodologies and their results, see Brown and Micklewright (2004).

8.2. Differences in PASEC and SACMEQ evaluation design

168. Some conceptional differences in PASEC and SACMEQ evaluation methodology and survey design may have a non-negligible impact on estimates of regression coefficients and standard errors, as well as on their interpretation:

Pre-test scores

169. Only PASEC tests students twice, one at the beginning and once at the end of the year. The inclusion of a pre-test score in a regression functions has important implications. First, it is a relevant control variable for general ability and the influence of student background which might not have been fully captured otherwise. Its inclusion can avoid (or reduce) omitted variable bias when estimating the effects of relevant policy measures. Second, it changes the interpretation of all coefficients as the control for score at the beginning of the term implies that the coefficients of all other variables reflect the influence on students' progress over the year, rather than on students' final skills. This is why econometric models including a pre-test score are also known as "added value models". And third, many teacher and classroom related variables change over the years, so that a precise estimation of their impact is only possible for the ongoing term. For example, the student may have got a high performing teacher for the current term, but had bad teachers before. Now since the overall skills of this student are influenced by all these teachers the positive influence of the last teacher will be blurred in any model in which initial student skills (before they got this teacher) cannot be taken into account.

Student weights

170. Only SACMEQ includes student weights, which can be used in the regression in order to ensure that the overall results are truly representative. Most PASEC surveys are designed to be representative surveys of schools, but it is not taken into account that the probability of any particular student to be part of the sample also depends on the size of the school. For Togo, Mali and Niger, they are not representative for schools, either, because they were designed to study specific policy measures (i.e. contract teachers and double shift teaching). This may result in some selection bias, and there are no weights to adjust for the non-random selection ex-post.

Target population and exclusions

171. An obvious difference is that PASEC and SACMEQ focus on different grades. Clearly, differences must be expected between students' learning in the early grades (like PASEC 2nd grade) and later grades. However, the differences between PASEC 5th grade and SACMEQ 6th grade appear to be less substantial. While drop-out increases from year to year, overall completion rates are higher in SACMEQ than in PASEC countries, so that the effect of sampling students from a one-year higher grade should be about compensated. Another concern could be that in many countries, 6th grade is the last year of primary education, which may make it an atypical year, difficult to compare with other years. However, it turns out that in most SACMEQ countries, primary schooling includes one more year and ends only after 7th grade. Thus in this respect, there does not seem to be a major problem for comparisons between PASEC and SACMEQ.

172. Obviously, this only holds for a comparison of regression coefficients. If, at some later point of time, anchor items will allow us to directly compare student achievement between the two country groups, the extra year of study will have to be taken into account.

173. A more relevant issue for our current analysis is that PASEC is sampling students within a single class for each school while SACMEQ is randomly drawing students from the overall 6th grade population within each school in the sample. This implies that for a given number of students drawn in each school and grade (typically 20 students in both surveys), in SACMEQ, we have more variation between teacher and classroom environments, but with only few students to whom this information can be directly related. Conversely, in PASEC, where we have information on 20 students actually taught by the same teacher in exactly the same environment. These differences lead to different degrees of precision for our econometric estimates at the different levels (schools, teachers / classrooms, and students).

174. In SACMEQ regressions, schools are the only level explicitly considered in the hierarchical models, and the primary sampling units in the survey regressions. In PASEC, the hierarchical level and the primary sampling unit considered is the classroom. The overall impact is difficult to predict. In any case, for SACMEQ, simple two-level hierarchical estimation models which do not take into account any sub-group clustering within schools appear to be problematic. This is the reason for the introduction of an alternative specification using Stata's survey sampling procedures which are robust to any correlations within primary sampling units.

175. Finally, neither in SACMEQ nor in PASEC all schools are included in the defined target population. In PASEC, sampling relies on school mappings available at the ministries of education, which, in some countries, exclude private schools. In SACMEQ, small schools with less than 15 or 20 students, schools for students with special needs and, in some cases, "inaccessible" schools were removed from the initial target population. While in SACMEQ countries these exclusions never went beyond 5%, their exclusion may still have an impact on the estimated role of certain variables such as class size, teachers' absence etc.

176. For further details on sample design procedures for SACMEQ, see SACMEQ (2004, section F). For PASEC, a similar brochure is in process and should be available in 2006.

8.3. Regression results

177. Finally, without being related to different sampling procedures, one more difference between our data for SACMEQ and PASEC should be kept in mind when interpreting regression results: Overall sample size is quite different for the two country groups. In SACMEQ, 14 countries are covered while only 8 countries are covered by PASEC (other country data are available since recently, but could not yet be integrated here). In terms of observations for individual students, this leads to a total sample size for SACMEQ which is more than twice as high as in PASEC. Obviously, this influences the precision of coefficient estimates in our regressions.

178. Table A1 displays the results for literacy, and Table A2 presents the results for mathematics. Each table includes ten regressions, two for SACMEQ (6^{th} grade only, model A and B), four for PASEC 5^{th} grade (model A and B, with and without pre-test) and four for PASEC 2^{nd} grade.

Table A1: Determinants of primary student literacy achievement in sub-Saharan Africa

SACMEQ	Regr. 1 English (1) grade 6 A: 2-level model		Regr. 2 English (1) grade 6 B: survey regr.		PASEC	Regr. 3 French grade 5 A: 2-level model		Regr. 4 French grade 5 B: survey regr.		Regr. 5 French grade 5 A: 2-level model		Regr. 6 French grade 5 B: survey regr.		Regr. 7 French grade 2 A: 2-level model		Regr. 8 French grade 2 B: survey regr.		Regr. 9 French grade 2 A: 2-level model		Regr. 10 French grade 2 B: survey regr.	
Variable and Range	Coef.	P>|z|	Coef.	P>|t|	**Variable and Range: modifications**	Coef.	P>|z|	Coef.	P>|t|	Coef.	P>|z|	Coef.	P>|t|	Coef.	P>|z|	Coef.	P>|t|	Coef.	P>|z|	Coef.	P>|t|
Initial score at the beginning of term (pre-test score)										0,44	0,00	0,47	0,00					0,51	0,00	0,51	0,00
Learning materials																					
Pupil possesses a textbook for reading						0,62	0,02	0,44	0,09	0,40	0,08	0,24	0,28	3,34	0,00	3,21	0,00	1,95	0,00	1,89	0,00
Availability of reading book (0=none, 0,3=shared with several peers, 0,5=shared with one peer, 1=own book)	3,89	0,01	6,46	0,01	share of books among the pupil's classmates	2,63	0,06	3,01	0,02	1,88	0,12	1,99	0,09	1,94	0,33	1,91	0,37	1,00	0,57	0,72	0,70
Class is equipped with wall chart	1,12	0,49	1,91	0,38																	
Teacher has access to a teacher's manual for reading	6,54	0,00	4,84	0,04		0,57	0,48	0,65	0,44	0,18	0,79	0,14	0,84	-0,21	0,88	-0,48	0,73	0,30	0,81	0,16	0,90
School equipment																					
Condition of school building (1=needs complete rebuilding - 5=good conditions)	1,67	0,05	1,87	0,04																	
School is equipped with a library	4,13	0,07	4,58	0,04		-0,08	0,95	-0,12	0,92	-0,88	0,37	-0,89	0,34	3,99	0,03	4,50	0,02	1,97	0,21	2,42	0,14
School is equipped with a first aid kit	3,63	0,12	3,88	0,12	sick-room with some equipment	0,37	0,84	0,72	0,67	0,28	0,86	0,57	0,69	-0,11	0,98	0,55	0,84	0,60	0,84	1,35	0,51
School has access to water	-3,11	0,22	-2,08	0,40		-0,46	0,49	-0,45	0,54	-0,89	0,13	-0,89	0,18	-0,32	0,77	-0,23	0,84	-0,24	0,80	-0,08	0,94
Pupil-toilet-ratio	0,00	0,98	0,00	0,80	toilets available (dummy)	0,37	0,58	0,18	0,81	0,46	0,42	0,32	0,63	0,20	0,85	0,37	0,75	-0,18	0,85	-0,08	0,94
School is equipped with electricity	-0,52	0,85	0,23	0,93		1,51	0,07	1,52	0,06	0,41	0,56	0,29	0,67	2,20	0,12	2,45	0,11	0,01	0,99	-0,11	0,93
School is equipped with technical resources: radio, tv,vcr,computer (0-4)	14,08	0,00	12,72	0,00																	
Class is equipped with library	9,65	0,00	3,54	0,10																	
Classroom is equipped with blackboard and chalk (or equivalent alternatives)	-6,92	0,01	-2,50	0,52		1,11	0,32	0,92	0,38	0,91	0,35	0,75	0,38	2,02	0,23	2,23	0,24	2,05	0,15	2,02	0,25
Teacher numbers and qualification																					
Class size	3,98	0,00	0,84	0,01		0,00	0,94	-0,02	0,62	0,01	0,66	0,01	0,87	-0,07	0,33	-0,07	0,37	-0,05	0,42	-0,06	0,42
Class size squared	-0,03	0,00	-0,01	0,01		0,00	0,66	0,00	0,89	0,00	0,29	0,00	0,27	0,00	0,61	0,00	0,54	0,00	0,48	0,00	0,37
Teacher academic qualification (1=primary-5=tertiary)	2,30	0,00	4,11	0,00	(0=below primary - 6=at least 3 years of tertiary)	0,30	0,28	0,28	0,34	0,03	0,89	0,02	0,94	-0,44	0,33	-0,42	0,32	-0,25	0,51	-0,19	0,62
Teacher professional qualification (1=no teacher training - 6=three years or more)	2,32	0,00	1,05	0,26	teacher has received at least some professional training (dummy)	-0,28	0,77	-0,39	0,64	0,09	0,91	0,06	0,93	1,03	0,42	1,37	0,29	0,77	0,48	0,99	0,37
Assessment of the efficacy of in-service training by teacher (1=no such training received - 5=very effective)	-1,83	0,00	-1,55	0,01	average number of in-service training courses per year (during the last five years) (3)	0,98	0,02	1,09	0,01	0,64	0,07	0,67	0,04	0,22	0,71	0,10	0,87	0,32	0,53	0,23	0,67
Organization of student flows and study time																					
School organization in shifts	-2,93	0,45	-5,83	0,10		-2,51	0,05	-2,44	0,09	-1,76	0,10	-1,58	0,23	-2,45	0,09	-1,54	0,28	-1,59	0,20	-1,11	0,39
Multi-grade teaching						-0,48	0,74	-1,22	0,51	1,03	0,41	0,80	0,66	-1,71	0,41	-1,49	0,47	-0,95	0,60	-0,78	0,70
Pupil repeats current grade (3)	-11,16	0,00	-12,11	0,00		1,17	0,00	0,97	0,02	0,46	0,10	0,19	0,60	-0,89	0,02	-0,75	0,18	-1,28	0,00	-1,15	0,01
Pupil's overall grade repetition (1=never - 4=three or more times)	-7,67	0,00	-10,83	0,00	current grade	-1,97	0,00	-2,17	0,00	-1,38	0,00	-1,30	0,00	-4,39	0,00	-5,02	0,00	-3,06	0,00	-3,12	0,00
Pupil is never missing school	16,15	0,00	14,81	0,01																	
Pupil has no health problems	-0,12	0,98	-3,92	0,45																	
Teacher arrives late (1=never - 3=often)	-6,06	0,01	-5,81	0,04																	
Teacher absenteeism (1=never - 3=often)	-3,08	0,09	-3,47	0,07	absence in days per month (0-25)	-0,08	0,42	-0,08	0,33	-0,04	0,63	-0,04	0,56	0,20	0,17	0,19	0,30	0,14	0,28	0,12	0,42
Number of lost official school days in the previous school year	0,06	0,63	0,06	0,66																	

(Table A1 cont.)

Institutional variables																				
School type (1=government, 2=privat)	9,60	0,02	7,13	0,12																
Parents' or community's contribution to class equipment of furniture, books and other materials (0=none-4)	2,14	0,01	2,28	0,01	parents easily mobilized for school issues (dummy)	0,72	0,27	0,63	0,37	0,02	0,97	-0,15	0,82	0,05	0,96	-0,02	0,98	-0,32	0,72	-0,39	0,67
Parents' or community's payment of exam fees, additional teacher salaries or bonuses (0=none -5)	0,54	0,61	0,34	0,78																
Teacher works on a non-civil servant contract						2,10	0,03	2,55	0,01	1,14	0,18	1,30	0,15	0,00	1,00	-0,08	0,96	-1,69	0,18	-2,27	0,10
Teacher gets advice from principal at least once a year	-5,76	0,02	-3,06	0,41	frequent exchange among teachers	-0,78	0,30	-0,60	0,46	-0,71	0,28	-0,53	0,44	-0,64	0,61	-0,54	0,71	0,25	0,82	0,36	0,77
Teacher considers promotion opportunities as very important	-5,92	0,00	-0,82	0,72																
School inspection in the year 2000	-0,88	0,60	-2,33	0,29	in the year of the survey	1,33	0,11	1,51	0,08	1,01	0,16	1,17	0,11	-0,93	0,45	-1,41	0,27	-0,47	0,66	-0,60	0,59
Controls					**Controls**																
Student characteristics and family background																					
Pupil is female	4,03	0,00	4,20	0,00		-0,46	0,02	-0,29	0,26	-0,10	0,57	0,08	0,74	-0,01	0,98	0,70	0,10	0,34	0,19	0,81	0,03
Pupil's age in months	-0,27	0,00	-0,22	0,00	in years	-0,69	0,00	-0,78	0,00	-0,49	0,00	-0,54	0,00	0,40	0,00	0,37	0,11	-0,13	0,31	-0,06	0,78
Pupil's home possessions (e.g. newspaper, tv, fridge, etc.; 0-14)	0,98	0,00	0,70	0,01	(0-8)	0,19	0,00	0,18	0,00	0,14	0,00	0,14	0,00	0,48	0,00	0,47	0,00	0,30	0,00	0,28	0,00
Pupil' housing conditions (3=bad - 16=good)	1,69	0,00	1,27	0,00																	
Pupil's meals per day (1=none at all - 12=3x every day)	3,43	0,00	3,55	0,00	(0-3)	1,22	0,00	0,87	0,02	0,57	0,00	-0,16	0,64	0,39	0,30	0,37	0,63	0,20	0,55	-0,36	0,59
Parental education (2=none - 12=both some post-secondary)	2,35	0,00	2,67	0,00	are literate)	0,82	0,00	1,10	0,00	0,44	0,00	0,54	0,02								
Number of books at pupil's home (0-250)	0,07	0,00	0,09	0,00	Pupil has some books at home (3)	1,70	0,00	2,01	0,00	1,13	0,00	1,32	0,00								
Pupil speaks English at home	23,86	0,00	27,04	0,00	French (3)	1,58	0,00	1,25	0,03	0,92	0,00	0,66	0,23	2,89	0,00	2,51	0,00	1,22	0,00	1,68	0,03
Pupil gets help with homework						-0,63	0,01	-0,72	0,03	-0,49	0,02	-0,65	0,03	1,90	0,00	1,98	0,00	0,61	0,05	0,25	0,61
Socio economic status of classmates (1-15)	6,92	0,00	6,58	0,00	(0-8)	0,57	0,01	0,62	0,00	0,37	0,04	0,40	0,03	0,60	0,15	0,75	0,08	0,06	0,86	0,13	0,73
Teacher characteristics																					
Teacher is female	-3,15	0,03	0,35	0,86		0,41	0,58	0,70	0,41	-0,28	0,66	-0,10	0,90	0,62	0,57	0,77	0,49	0,65	0,49	0,76	0,44
Teacher job experience (in years)	0,22	0,01	0,19	0,19		0,14	0,00	0,14	0,01	0,11	0,00	0,11	0,02	-0,04	0,55	-0,04	0,58	-0,04	0,47	-0,04	0,56
Teacher speaks teaching language at home (3)						-0,17	0,71	-0,11	0,82	0,26	0,51	0,29	0,47	-0,12	0,87	0,02	0,98	-0,32	0,60	-0,27	0,67
Teacher speaks local language						0,68	0,36	0,44	0,55	0,64	0,31	0,39	0,55	2,00	0,09	2,05	0,09	0,62	0,54	0,56	0,58
Pedagocial tools																					
Frequency of reading test (1=none - 6=once or more per week)	24,27	0,00	17,05	0,02																	
Squared frequency of reading test	-2,77	0,00	-2,07	0,01																	
Frequency teacher corrects reading-homework (1=no homework - 5=always)	1,62	0,00	1,10	0,05																	
Other controls																					
School participates in a pilot project, exchange program etc.						0,40	0,59	0,44	0,57	-0,10	0,88	-0,07	0,91	1,30	0,27	1,19	0,39	1,37	0,18	1,41	0,20
School size (number of pupils)	-0,01	0,00	0,00	0,68		-0,01	0,00	0,00	0,01	0,00	0,01	0,00	0,06	0,00	0,08	-0,01	0,02	0,00	0,06	-0,01	0,02
School location (1=isolated-4=city)	3,53	0,02	4,98	0,00	dummy (0=rural, 1=urban)	2,32	0,00	2,39	0,00	1,99	0,00	2,10	0,00	2,18	0,10	2,12	0,12	1,92	0,09	1,98	0,09
Class environement (e.g. disturbance, theft,etc.; 0=never-17)	0,19	0,55	0,34	0,38																	

(Table A1 cont.)

Country fixed effects																				
Botswana	262,66	0,00	338,90	0,00	Burkina Faso	39,66	0,00	41,35	0,00	23,22	0,00	24,74	0,00	49,23	0,00	48,87	0,00	42,39	0,00	43,12	0,00
Kenya	292,13	0,00	372,21	0,00	Cameroon	49,36	0,00	50,48	0,00	27,58	0,00	28,18	0,03	58,18	0,00	58,01	0,00	44,33	0,33	44,97	0,42
Lesotho	194,33	0,00	286,19	0,00	Côte d'Ivoire	43,58	0,00	45,32	0,01	24,25	0,38	25,44	0,60	47,46	0,43	47,25	0,47	37,13	0,01	37,95	0,01
Malawi	226,11	0,00	304,04	0,00	Madagascar	38,13	0,30	39,49	0,26	21,73	0,25	22,74	0,19	50,18	0,68	50,27	0,60	37,52	0,02	38,61	0,04
Mauritius	166,02	0,00	257,09	0,00	Mali	5,32	0,00	7,17	0,00	1,61	0,00	3,56	0,00	8,70	0,00	7,19	0,00	13,52	0,00	13,54	0,00
Mozambique	288,04	0,00	376,36	0,00	Niger	23,95	0,00	25,54	0,00	15,11	0,00	17,19	0,00	34,07	0,00	34,77	0,00	33,53	0,00	35,29	0,00
Namibia	193,70	0,00	276,00	0,00	Senegal	28,77	0,00	30,16	0,00	19,70	0,01	21,46	0,01	34,57	0,00	34,18	0,00	31,27	0,00	32,46	0,00
Seychelles	227,13	0,00	301,20	0,00	Togo	10,82	0,00	12,78	0,00	6,32	0,00	8,67	0,00	14,35	0,00	14,12	0,00	17,22	0,00	18,01	0,00
South Africa	199,21	0,00	290,50	0,00																	
Swaziland	268,12	0,00	351,00	0,00																	
Tanzania	318,80	0,00	401,80	0,00																	
Uganda	254,72	0,00	336,27	0,00																	
Zambia	206,73	0,00	286,71	0,00																	
Zanzibar	242,59	0,00	326,95	0,00																	
n	37383		37383			13612		13612		13551		13551		13621		13621		13500		13500	
Specified strata (countries)			14					8				8				8				8	
Specified PSUs (schools)	2074		2074			862		862		862		862		874		874		874		874	
R-squared, between (2)	43,4%					76,7%				82,4%				62,2%				71,9%			
R-squared, within (2)	4,7%					4,4%				22,2%				2,8%				24,1%			
R-squared, total (2)	78,2%		40,7%			54,8%		54,8%		64,0%		64,1%		39,6%		39,7%		53,8%		53,9%	

(1) Exceptions: Portuguese for Mozambique, Swahili for Tanzania (mainland) and Zanzibar.

(2) Pseudo R-squared in case of Maximum likelihood estimations (Regression 1). The R-squared refers to a model with constant (omitting one country fixed effect). Note that the R-squared between schools appears very high because it includes the impact of the country dummies.

(3) Variable missing for one country and imputed using the linear regression on related variables in the cross-country sample.

Table A2: Determinants of primary student mathematics achievement in sub-Saharan Africa

SACMEQ	Regr. 11 Math grade 6 A: 2-level model		Regr. 12 Math grade 6 B: survey Regr.		PASEC	Regr. 13 Math grade 5 A: 2-level model		Regr. 14 Math grade 5 B: survey Regr.		Regr. 15 Math grade 5 A: 2-level model		Regr. 16 Math grade 5 B: survey Regr.		Regr. 17 Math grade 2 A: 2-level model		Regr. 18 Math grade 2 B: survey Regr.		Regr. 19 Math grade 2 A: 2-level model		Regr. 20 Math grade 2 B: survey Regr.	
Variable and Range	Coef.	P>|z|	Coef.	P>|t|	**Variable and Range: modifications**	Coef.	P>|z|	Coef.	P>|t|	Coef.	P>|z|	Coef.	P>|t|	Coef.	P>|z|	Coef.	P>|t|	Coef.	P>|z|	Coef.	P>|t|
Initial score at the beginning of term (pre-test score)										0,41	0,00	0,42	0,00					0,60	0,00	0,51	0,00
Learning materials																					
Pupil possesses a textbook for math						0,93	0,00	0,93	0,00	0,58	0,01	0,60	0,03	1,91	0,00	1,69	0,00	1,20	0,00	1,13	0,01
Availability of math book (0=none, 0,3=shared with several peers, 0,5=shared with one peer, 1=own book)	5,94	0,00	8,43	0,00	share of books among the pupil's classmates	1,01	0,47	1,17	0,47	0,60	0,63	0,65	0,66	2,92	0,10	2,82	0,15	1,74	0,27	1,62	0,33
Class is equipped with wall chart	0,80	0,65	0,55	0,80																	
Teacher has access to a teacher's manual for math	1,88	0,25	6,74	0,00		0,63	0,38	0,84	0,33	0,62	0,34	0,83	0,29	0,91	0,47	1,13	0,36	1,15	0,30	1,19	0,33
School equipment																					
Condition of school building (1=needs complete rebuilding - 5=good conditions)	2,44	0,01	2,42	0,01																	
School is equipped with a library	2,98	0,22	2,70	0,25		0,82	0,48	0,43	0,70	-0,40	0,70	-0,78	0,45	1,51	0,31	2,00	0,18	-0,21	0,87	0,51	0,70
School is equipped with a first aid kit	2,77	0,27	1,83	0,52	sick-room with some equipment	2,23	0,25	2,57	0,19	2,03	0,24	2,33	0,13	1,66	0,56	2,49	0,43	3,07	0,22	3,76	0,17
School has access to water	-4,81	0,08	-3,03	0,27		-0,44	0,53	-0,57	0,46	-0,59	0,34	-0,69	0,32	-0,77	0,40	-0,80	0,39	-0,63	0,44	-0,71	0,40
Pupil-toilet-ratio	0,00	0,67	0,00	0,75	toilets available (dummy)	-0,14	0,84	-0,28	0,72	0,23	0,70	0,15	0,83	-0,34	0,71	-0,35	0,70	0,12	0,88	0,04	0,96
School is equipped with electricity	-5,84	0,04	-4,28	0,10		1,08	0,20	1,23	0,15	0,22	0,77	0,25	0,73	1,49	0,20	1,52	0,21	-0,08	0,94	0,03	0,98
School is equipped with technical resources: radio, tv,vcr,computer (0-4)	13,67	0,00	12,90	0,00																	
Class is equipped with library	4,27	0,01	4,20	0,08																	
Classroom is equipped with blackboard and chalk (or equivalent alternatives)	6,24	0,03	3,57	###		0,58	0,62	0,44	0,69	0,02	0,98	0,01	0,99	0,25	0,86	0,08	0,96	2,20	0,08	1,83	0,10
Teacher numbers and qualification																					
Class size	3,29	0,00	-0,01	0,98		0,02	0,60	0,01	0,82	0,01	0,69	0,00	0,95	-0,06	0,31	-0,05	0,41	-0,03	0,52	-0,03	0,55
Class size squared	-0,03	0,00	0,00	0,77		0,00	0,26	0,00	0,32	0,00	0,29	0,00	0,36	0,00	0,66	0,00	0,66	0,00	0,71	0,00	0,61
Teacher academic qualification (1=primary-5=tertiary)	1,75	0,04	3,20	0,01	(0=below primary - 6=at least 3 years of tertiary)	0,42	0,14	0,49	0,09	0,34	0,18	0,42	0,12	-0,26	0,49	-0,26	0,46	-0,01	0,96	0,00	0,99
Teacher professional qualification (1=no teacher training - 6=three years or more)	2,37	0,00	1,87	0,05	teacher has received at least some professional training (dummy)	-0,60	0,53	-0,59	0,51	-0,64	0,45	-0,69	0,41	0,14	0,90	0,23	0,84	-0,40	0,67	-0,27	0,78
Assessment of the efficacy of in-service training by teacher (1=no such training received - 5=very effective)	-0,14	0,76	-0,20	0,76	average number of in-service training courses per year (during the last five years) (2)	0,52	0,22	0,61	0,17	0,37	0,32	0,40	0,27	0,46	0,34	0,48	0,34	0,49	0,25	0,52	0,25
Organization of student flows and study time																					
School organization in shifts	0,57	0,89	-2,48	0,45		-2,28	0,08	-2,50	0,11	-1,14	0,32	-1,32	0,34	-3,45	0,00	-2,70	0,02	-2,29	0,03	-2,13	0,05
Multi-grade teaching						-1,48	0,32	-1,72	0,34	-0,09	0,95	-0,08	0,96	-1,97	0,26	-1,40	0,48	-0,20	0,90	-0,07	0,97
Pupil repeats current grade (2)	-8,66	0,00	-8,88	0,00		2,16	0,00	1,71	0,00	0,77	0,01	0,27	0,51	0,30	0,46	0,57	0,26	-0,55	0,13	-0,18	0,68
Pupil's overall grade repetition (1=never - 4=three or more times)	-4,82	0,00	-9,32	0,00	number of grades repeated before current grade	-1,58	0,00	-1,94	0,00	-1,10	0,00	-1,30	0,00	-4,74	0,00	-4,93	0,00	-3,39	0,00	-3,53	0,00
Pupil is never missing school	28,15	0,00	24,29	0,00																	
Pupil has no health problems	2,87	0,58	1,96	0,79																	
Teacher arrives late (1=never - 3=often)	-7,95	0,00	-7,04	0,01																	
Teacher absenteeism (1=never - 3=often)	0,16	0,94	-0,43	0,83	absence in days per month (0-25)	-0,14	0,16	-0,16	0,11	-0,10	0,27	-0,12	0,16	0,22	0,07	0,23	0,09	0,11	0,32	0,12	0,26
Number of lost official school days in the previous school year	-0,06	0,69	-0,02	0,87																	

(Table A2 cont.)

Institutional variables																		
School type (1=government, 2=privat)	3,67	0,42	1,08	0,84														
Parents' or community's contribution to class equipment of furniture, books and other materials (0=none-4)	1,81	0,04	2,09	0,02	parents easily mobilized for school issues (dummy)	1,33	0,04	1,38	0,05	0,58	0,32	0,40	0,53	0,61	0,48	0,38	0,66	0,53	0,49	0,40	0,59
Parents' or community's payment of exam fees, additional teacher salaries or bonuses (0=none -5)	-0,82	0,47	-0,47	0,73																	
Teacher works on a non-civil servant contract						1,44	0,16	1,99	0,08	0,70	0,44	0,98	0,35	-0,45	0,71	-0,54	0,67	-1,03	0,35	-1,24	0,32
Teacher gets advice from principal at least once a year	0,96	0,70	-0,77	0,84	frequent exchange among teachers	0,66	0,39	0,72	0,41	0,57	0,40	0,53	0,49	-1,05	0,31	-1,06	0,36	0,57	0,53	0,38	0,70
Teacher considers promotion opportunities as very important	-6,67	0,00	-1,85	0,47																	
School inspection in the year 2000	-0,30	0,87	-3,49	0,18	in the year of the survey	1,79	0,04	2,13	0,02	1,55	0,04	1,94	0,02	0,67	0,51	0,46	0,66	-0,18	0,84	-0,17	0,85
Controls					**Controls**																
Student characteristics and family background																					
Pupil is female	-8,48	0,00	-7,93	0,00		-1,56	0,00	-1,20	0,00	-0,71	0,00	-0,30	0,20	-1,54	0,00	-0,92	0,03	-0,73	0,01	-0,47	0,18
Pupil's age in months	-0,27	0,00	-0,22	0,00	in years	-0,43	0,00	-0,27	0,02	-0,41	0,00	-0,25	0,01	0,81	0,00	0,81	0,00	-0,04	0,78	0,00	0,98
Pupil's home possessions (e.g. newspaper, tv, fridge, etc.; 0-14)	1,19	0,00	0,92	0,00	(0-8)	0,10	0,06	0,07	0,22	0,07	0,13	0,04	0,37	0,45	0,00	0,45	0,00	0,25	0,00	0,26	0,00
Pupil' housing conditions (3=bad - 16=good)	1,02	0,00	0,59	0,02																	
Pupil's meals per day (1=none at all - 12=3x every day)	2,81	0,00	3,11	0,00	(0-3)	0,97	0,00	0,85	0,03	0,08	0,67	-0,48	0,19	0,59	0,13	0,23	0,71	-0,07	0,83	-0,59	0,27
Parental education (2=none - 12=both some post-secondary)	1,45	0,00	1,80	0,00	Parental literacy (0=none - 2=both)	0,63	0,00	0,98	0,00	0,42	0,00	0,50	0,02								
Number of books at pupil's home (0-250)	0,05	0,00	0,08	0,00	Pupil has some books at home (2)	1,15	0,00	1,56	0,00	0,22	0,37	0,38	0,32								
Pupil speaks English at home	22,25	0,00	24,02	0,00	French (2)	1,14	0,00	1,07	0,05	0,52	0,05	0,84	0,12	1,86	0,00	2,11	0,00	0,90	0,04	2,45	0,00
Pupil gets help with homework						-0,33	0,15	-0,37	0,29	-0,05	0,80	-0,15	0,63	2,33	0,00	2,56	0,00	0,96	0,00	0,92	0,04
Socio economic status of classmates (1-15)	4,88	0,00	4,32	0,00	(0-8)	0,36	0,10	0,36	0,07	0,20	0,31	0,19	0,31	0,73	0,03	0,89	0,01	0,81	0,01	0,90	0,00
Teacher characteristics																					
Teacher is female	1,27	0,40	-1,49	0,51		-0,99	0,20	-0,78	0,36	-1,59	0,02	-1,54	0,04	1,70	0,06	1,78	0,06	1,85	0,02	1,87	0,02
Teacher job experience (in years)	0,30	0,00	0,23	0,07		0,10	0,04	0,11	0,04	0,09	0,03	0,10	0,03	-0,08	0,18	-0,08	0,19	-0,08	0,11	-0,07	0,16
Teacher speaks teaching language at home (2)						-0,48	0,30	-0,55	0,27	-0,05	0,91	-0,02	0,97	0,23	0,70	0,31	0,62	0,13	0,81	0,22	0,69
Teacher speaks local language						0,23	0,76	0,05	0,95	0,32	0,64	0,19	0,80	1,96	0,05	2,03	0,05	0,86	0,32	1,06	0,24
Pedagocial tools																					
Frequency of math test (1=none-6=once or more per week)	-0,68	0,91	-10,50	0,25																	
Squared frequency of math test	0,00	1,00	1,11	0,26																	
Frequency teacher corrects math-homework (1=no homework - 5=always)	3,13	0,00	3,31	0,00																	
Other controls																					
School participates in a pilot project, exchange program etc.						0,40	0,60	0,51	0,53	-0,55	0,41	-0,46	0,51	0,39	0,69	0,24	0,83	-0,10	0,91	-0,27	0,76
School size (number of pupils)	-0,01	0,00	0,00	0,43		0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,01	0,00	0,66	0,00	0,87	0,00	0,64	0,00	0,99
School location (1=isolated-4=city)	0,99	0,53	2,68	0,10	dummy (0=rural, 1=urban)	1,27	0,12	1,29	0,11	0,92	0,20	1,04	0,16	1,50	0,17	1,40	0,20	0,47	0,63	0,62	0,53
Class environement (e.g. disturbance, theft, etc.; 0=never-17)	0,41	0,23	0,52	0,16																	

(Table A2 cont.)

Country fixed effects																				
Botswana	341,48	0,00	422,23	0,00	Burkina Faso	44,28	0,00	42,33	0,00	31,19	0,00	30,03	0,00	41,10	0,00	40,89	0,00	25,62	0,00	29,32	0,00
Kenya	396,38	0,00	483,11	0,00	Cameroon	46,04	0,27	43,85	0,44	27,02	0,00	25,63	0,02	44,77	0,07	44,58	0,11	19,52	0,00	24,38	0,02
Lesotho	277,64	0,00	374,14	0,00	Côte d'Ivoire	35,25	0,00	33,13	0,00	19,08	0,00	17,52	0,00	28,93	0,00	28,84	0,00	11,76	0,00	15,64	0,00
Malawi	306,58	0,00	387,15	0,00	Madagascar	56,24	0,00	53,78	0,00	34,68	0,01	33,10	0,07	52,11	0,00	52,38	0,00	21,26	0,02	27,95	0,48
Mauritius	317,79	0,00	416,53	0,00	Mali	11,25	0,00	8,97	0,00	9,33	0,00	7,87	0,00	4,92	0,00	3,67	0,00	6,29	0,00	6,97	0,00
Mozambique	384,28	0,00	480,83	0,00	Niger	27,28	0,00	24,94	0,00	19,91	0,00	18,83	0,00	26,21	0,00	26,93	0,00	13,18	0,00	17,26	0,00
Namibia	262,78	0,00	352,72	0,00	Senegal	33,53	0,00	31,20	0,00	27,72	0,01	26,73	0,02	29,73	0,00	29,88	0,00	23,82	0,31	26,94	0,18
Seychelles	306,44	0,00	386,36	0,00	Togo	13,26	0,00	11,28	0,00	9,56	0,00	8,42	0,00	7,61	0,00	7,44	0,00	9,25	0,00	10,02	0,00
South Africa	287,46	0,00	385,07	0,00																	
Swaziland	348,13	0,00	437,59	0,00																	
Tanzania	383,03	0,00	471,52	0,00																	
Uganda	371,31	0,00	454,44	0,00																	
Zambia	289,95	0,00	375,09	0,00																	
Zanzibar	328,31	0,00	423,73	0,00																	
n	36023		36023			13866		13866		13791		13791		13464		13464		13414		13414	
Specified strata (countries)			14					8				8				8				8	
Specified PSUs (schools)	2002		2002			842		842		842		842		860		860		860		860	
R-squared, between (1)	40,8%					74,8%				80,0%				69,3%				74,4%			
R-squared, within (1)	1,5%					2,9%				22,4%				2,3%				26,1%			
R-squared, total (1)	79,0%		38,5%			51,8%		51,9%		61,5%		61,6%		40,0%		40,1%		53,3%		53,7%	

(1) Pseudo R-squared in case of Maximum likelihood estimations (Regression 11). The R-squared refers to a model with constant (omitting one country fixed effect). Note that the R-squared between schools appears very high because it includes the impact of the country dummies.

(2) Variable missing for one country and imputed using the linear regression on related variables in the cross-country sample.

9. REFERENCES

Angrist, J., Bettinger, E., Blohm, E., King, E. and Kremer, M. 2002. "Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment" in: *American Economic Review, 92, (5), 1535-1558*.

Angrist, J., Bettinger, E. and Kremer, M. 2004. "Long-Term Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia" *BREAD Working Paper, 79, Harvard.*

Angrist, J. and Lang, K., 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program" in: *American Economic Review 94, (5), 1613-1634.*

Angrist, J. and Lavy, V. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Children's Academic Achievement" in: *Quarterly Journal of Economics, 114, (2), 533-576.*

Averett, S. and McLennan M. 2004. *Exploring the Effect of Class Size on Student Achievement: What Have We learnt over the Past Two Decades,* in Geraint Johnes and Jill Johnes (eds), Cheltenham, Edward Elgar, International Handbook on the Economics of Education.

Banerjee, A., Cole, S., Duflo, E. and Linden, L. 2003. *Remedying Education: Evidence from Two Randomized Experiments in India*, Poverty Action Lab Paper (4), MIT downloaded on 27/10/05 at http://www.prathamusa.org/docs/Balsakhi_Impact_MIT_Pratham.pdf

Bedi, A. and Garg, A. 2000. "The Effectiveness of Private Versus Public Schools: The Case of Indonesia" in: *Journal of Development Economics*, 61, (2), 463-494.

Bernard, J.-M. 1999. *Les enseignants du primaire dans cinq pays du Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN: Le role du maître dans le processus d'acquisition des élèves*, report for the ADEA Working Group on the Teaching Profession, French Section, Paris: ADEA.

Bernard, J.-M. 2004. "Scolarisation primaire universelle et le pilotage par les résultats dans le contexte africain: quels indicateurs?", Paper presented at the international AFEC *conference Le droit à l'éducation: quelle effectivité au sud et au nord?*, 9-12 March, Ouagadougou.

Bernard, J.-M. and Michaelowa, K. 2005. *Managing the Impact of PASEC Projects in Francophone Sub-Saharan Africa*, in: Kenneth Ross and Ilona Jürgen-Genevois (eds): Cross-National Studies of the Quality of Education: Planning Their Design and Managing Their Impact, Paris: IIEP/UNESCO. Forthcoming document.

Bernard, J.-M., Simon, O. and Vianou, K. 2005. Le redoublement: mirage de l'école africaine?, Dakar: PASEC/CONFEMEN.

Bernard, J.-M., Tiyab, B. and Vianou, K. 2004. Profils enseignants et qualité de l'éducation primaire en Afrique subsaharienne francophone: Bilan et perspectives de dix années de recherche du PASEC, Dakar : PASEC/CONFEMEN.

Bishop, J. 1997. "The Effect of National Standards and Curriculum Based Examinations on Achievement" in: *American Economic Review*, 87, (2), 260-264.

Bishop, J. and Wößmann, L. 2004. "Institutional Effects in a Simple Model of Educational Production" in: *Education Economics*, 12, (1), 17-38.

Bourdon, J. 2005. Les maîtres contractuels sont-ils aussi efficaces que les maîtres titulaires? Essai d'appariement sur les données de l'enquête PASEC sur le Mali, IREDU (University of Burgundy).

Bourdon, J., Frölich, M. and Michaelowa, K. 2006. *Broadening Access to Primary Education: Contract Teacher Programs and Their Impact on Education Outcomes in Africa – an Econometric Evaluation for the Republic of Niger*, in: Lukas Menkhoff (ed.): Pro-Poor Growth: Issues, Policies, and Evidence, Berlin: Duncker & Humblot. Forthcoming document.

Brown, G. and Micklewright J. 2004. Using International Surveys of Achievement and Literacy: A View from the Outside. Montreal: UNESCO Institute for Statistics.

Chaudhury, N., Hammer, J., Kremer, M. Muralidharan, K. and Rogers, H. 2005. "Missing in Action: Teacher and Health Worker Absence in Developing Countries", in: *Journal of Economic Perspectives*, forthcoming, downloaded on 21/11/05 at http://post.economics.harvard.edu/faculty/kremer/webpapers/MissinginAction.pdf.

Chinapah, V. 1997. *Monitoring Learning Achievement*. UNESCO-UNICEF follow-up to Jomptien. Paris: UNESCO.

Cox, D. and Jimenez, E. 1991. "The Relative Effectiveness of Private and Public Schools" in: *Journal of Development Economics*, 34, (1), 99-121.

CONFEMEN 2005. *Accueil PASEC* downloaded on 15/11/05 at http://www.confemen.org/rubrique.php3?id_rubrique=3.

Evers, W. and Walberg, H. (Eds) 2002: *School Accountability*, Stanford: Hoover Institution Press.

Fehrler, S. 2005. "Erklärung von schulischem (Miss-) Erfolg im Senegal mit Hilfe von fixed effects Modellen" bachelor thesis, Center for Applied Statistics and Economics, Berlin: Humboldt-University.

Frölich, M. and Michaelowa, K. 2005. "Peer Effects and Textbooks in Primary Education: Evidence From Francophone Sub-Saharan Africa" *HWWA Discussion Paper,* 311, Hamburg: Hamburg Institute of International Economics.

Fuller, B. 1987. "What School Factors Raise Achievement in the Third World?" in: *Review of Educational Research*, 57, (3), 255-292.

Fuller, B. and Clarke, P. 1994. "Raising School Effects While Ignoring Culture? Local Conditions and the Influence of Classroom Tools, Rules, and Pedagogy" in: *Review of Educational Research* 64, (1), 119-157.

Gay, G. 1990. "Standardized Tests: Irregularities in Administering of Test Affect Test Results" in: *Journal of Instructional Psychology*, 17, (2), 93-103.

Glewwe, P. and Jacoby, H. 1994. "Student Achievement and Schooling Choice in Low-Income Countries: Evidence from Ghana" in: *Journal of Human Resources*, 29, (3), 834-864.

Glewwe, P., Kremer, M. and Moulin, M. 2000. *Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya,* Development Research Group, Washington DC: World Bank. Downloaded on 27/10/05 at http://post.economics.harvard.edu/faculty/kremer/webpapers/Textbooks_Test_Scores.pdf.

Glewwe, P., Kremer, M., Moulin, S. and Zitzewitz, E. 2004. "Retrospective vs. Prospective Analysis of Flip Charts in Kenya" in: *Journal of Development Economics*, 74, (1), 251-268.

Glewwe, P. and Kremer, M. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries" in: *Handbook on the Economics of Education* (Elsevier), second draft as of April 2005. Forthcoming document.

Goldstein, H. 2003. *Multilevel Statistical Models*, 3rd Edition, London: Edward Arnold.

Hanushek, E.A. 2003. "The Failure of Input-Based Schooling Policies" in: *The Economic Journal*, 113, (485), F64-F98.

Hoxby, C. 2000a. "Does Competition among Public Schools Benefit Students and Taxpayers?" in: *American Economic Review*, 90, (5), 1209-1238.

Hoxby, C. 2000b. "Peer Effects in the Classroom: Learning from Gender and Race Variation" Working Paper, 7867, Cambridge: NBER

Hoxby, C. (ed.) 2003a. *The Economics of School Choice*, NBER Conference Report, Chicago: University of Chicago Press.

Hoxby, C. 2003b. "School Choice and School Competition: Evidence from the United States" in: *Swedish Economic Policy Review*, 10, (3) 9-65.

Imazeki, J. 2005. "Teacher Salaries and Teacher Attrition" in: *Economics of Education Review*, 24, (4), 431-449.

Jacob, B. and Levitt, S. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating" in: *Quarterly Journal of Economics*, 118, (3), 843-877.

James, E., King, E. and Suryadi, A. 1996. "Finance, Management, and Costs of Public and Private Schools in Indonesia" in *Economics of Education Review*, 15, (4), 387-398.

Jamison, D., Searle, B., Galda, K. and Heyneman, S. 1981. "Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement" in: *Journal of Educational Psychology*, 73, (4), 556-567.

Jürges, H. and Schneider, K. 2005. *Performance-Based Pay for Teachers? Fair Rankings of Teacher Quality*, presentation at the Annual Conference of the German Economic Association, Bonn, 27-30 September.

Juvane, V. 2005. *Redefining the Role of Multi-Grade Teaching*, presentation at the FAO, IIEP/UNESCO and ADEA ministerial seminar on "Education for Rural People in Africa: Policy Lessons, Options and Priorities, Addis Ababa, 7-9 September, downloaded on 25/10/05 at www.fao.org/sd/erp/addisababa/Papers/Multigrade%20Teaching%20(WGTP)_ENG.doc

Kellaghan, T. and Greaney, V. 2003. *Monitoring Performance: Assessment and Examinations in Africa*, Background Paper prepared for the ADEA Biennial Meeting 2003, Paris: ADEA.

King, E., Orazem, P. and Wohlgemuth, D. 1999. "Central Mandates and Local Incentives: The Colombia Education Voucher Program" in: *World Bank Economic Review*, 13, (3), 467-491.

Kremer, M. and Sarychev, A. 2000. *Why Do Governments Operate Schools?*, Harvard University. Unpublished document.

Lassibile, G. and Tan, J.P. 2003. "Student Learning in Public and Private Primary Schools in Madagascar" in: *Economic Development and Cultural Change*, 51, (3), 699-717.

Lavy, V. 2002. "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement" in: *Journal of Political Economy*, 110, (6), 1286-1317.

Lavy, V. 2004. "Performance Pay and Teachers' Effort, Productivity and Grading Ethics" Working Paper, 10622, Cambridge: NBER.

Lee, V., Zuze, T. and Ross, K. 2005. "School Effectiveness in 14 Sub-Saharan African Countries: Links With 6th Graders' Reading Achievement" in: *Studies in Educational Evaluation*, 31, (2-3), 207-246.

Lockheed, M.E. and Verspoor, A.M. 1991. *Improving the Quality of Primary Education in Developing Countries*, Washington: World Bank.

Michaelowa, Katharina (2000): "Dépenses d'éducation, qualité de l'éducation et pauvreté: l'exemple de cinq pays d'Afrique francophone", *OECD Development Centre Technical Paper*, 157, Paris: OECD Development Centre.

Michaelowa, K. 2001. "Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations" in: *World Development*, 29, (10), 1699-1716.

Michaelowa, K. 2002. "Teacher Job Satisfaction, Student Achievement, and the Cost of Primary Education in Francophone Sub-Saharan Africa", *HWWA Discussion Paper*, 188, Hamburg: Hamburg Institute of International Economics.

Michaelowa, K. 2003. *Determinants of Primary Education Quality: What Can We Learn From PASEC forFrancophone Sub-Saharan Africa*, Background Paper prepared for the ADEA Biennial Meeting 2003, Paris: ADEA.

Miguel, E. and Gugerty, M. 2005. "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya" in: *Journal of Public Economics*, 89, (11), 2325-2368.

Mingat, A. 2003. *Analytical and Factual Elements for a Quality Policy for Primary Education in Sub-Saharan Africa in the Context of Education For All*, Background Paper prepared for the ADEA Biennial Meeting 2003, Paris: ADEA.

Mingat, A. and Suchaut, B. 2000. *Les systémes éducatifs africains: une analyse économique comparative*, Brussels: De Boeck Univertsité.

Mizala, A., Romaguera, P. and Farren, D. 2002. "The Technical Efficiency of Schools in Chile" in: *Applied Economics*, 34, (12), 1533-1552.

Murimba, S. 2005a. "The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ): Mission, Approach and Projects" in: *Prospects*, 35, (1), 75-89.

Murimba, S. 2005b. "The Impact of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)" in: *Prospects*, 35, (1), 91-108.

Nguyen, K., Wu, M. and Gillis, S. 2005. *Factors Influencing Pupil Achievement in SACMEQ II – Botswana: An Application of Structural Equation Modelling*, paper presented at the International Invitational Educational Policy Research Conference, 28 September-2 October, Paris, IIEP/UNESCO.

OECD. 2001. *Knowledge and Skills for Life*, First Results from the OECD Programme on International Student Achievement (PISA), Paris: OECD.

OECD. 2002. Education at a Glance, Paris: OECD.

OECD and UNESCO-UIS. 2003. *Literacy Skills for the World of Tomorrow, Further Results from PISA 2000*, Paris: OECD.

PASEC. 1999. *Les facteurs de l'efficacité dans l'enseignement primaire: Les résultats du programme PASEC sur neuf pays d'Afrique et de l'Océan indien*, Rapport de synthèse du Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC), Dakar: CONFEMEN.

PASEC. 2003. *Le Programme de formation initiale des maîtres et la double vacation en Guinée*, Dakar: CONFEMEN.

PASEC. 2004. *Recrutement et formation des enseignants du premier degré au Togo: Quelles priorités?*, Dakar: CONFEMEN.

PASEC. 2005a. *Rapport Mauritanie*, Dakar: CONFEMEN. Forthcoming document.

PASEC. 2005b. *Les enseignants contractuels et la qualité de l'enseignement de base I au Niger : Quel bilan?*, Dakar : CONFEMEN. Forthcoming document.

PASEC. 2005c. *Impact du statut enseignant sur les acquisitions dans le premier cycle l'enseignement fondamental public au Mali*, Dakar : CONFEMEN. Forthcoming document.

Preuschoff, C. and Weiß, M. 2004. *Schülerleistungen an staatlichen und privaten Schulen im Vergleich*, in: Schümer, G., Tillmann, K.J. and Weiß, M. (eds). Die Institution Schule und die Lebenswelt der Schüler, Wiesbaden: VS-Verlag für Sozialwissenschaften, 39-71.

Peterson, P. and West, M. 2003. *No Child Left Behind? The Politics and Practice of School Accountability*, Washington: Brookings Institution Press.

Pritchett, L. 2004. *Access to Education*, in: Björn Lomborg (ed.): Global Crisis, Global Solutions, Cambridge: Cambridge University Press, 175-234.

Raudenbush, S. and Bryk, A. 2002. *Hierarchical Linear Models: Applications and Analysis Methods*, 2nd Edition, Sage: Thousand Oaks.

Reinikka, R. and Svensson, J. 2003. "The Power of Information: Evidence from a Newspaper Campaign to Reduce Capture" *Policy Research Working Paper*, 3239, Washington: World Bank.

Ross, K. 1998 (ed.). *SACMEQ Policy Research, The Quality of Education: Some Policy Suggestions Based on a Survey of Schools* (various country reports). Paris: IIEP/UNESCO.

SACMEQ (2004): *Data Archive for the SACMEQ I and SACMEQ II Projects*, Version 4, CD-ROM, Paris: IIEP/UNESCO

Scheerens, J. 2000. *Improving School Effectiveness*, Paris: IIEP/UNESCO.

Somers, M.A., McEvan, P. and Willms, D. 2004. "How Effective Are Private Schools in Latin America?" in: *Comparative Education Review*, 48, (1), 48-70.

StataCorp. 2003. Survey Data, Stata Statistical Software, release 8, College Station: Stata Corporation

Tatto, M., Nielsen, D. and Cummings, W. 1991. *Comparing the Effects and Costs of Different Approaches for Educating Primary School Teachers: The Case of Sri Lanka*, Harvard: BRIDGES Report, 10.

Tekleselassie, A. 2005. "Teachers' Career Ladder Policy in Ethiopia: An Opportunity for Professional Growth or 'a Stick Disguised as a Carrot'" in: *International Journal of Educational Development*, 25, 618-636.

UNESCO. 2004. *Education for All* (EFA) Global Monitoring Report 2005, Education for All, The Quality Imperative, Paris: UNESCO. Available: http://portal.unesco.org

Vegas, E. and de Laat, J. 2003. "Do Differences in Teacher Contracts Affect Student Performance? Evidence from Togo" *Working Paper*, 26955, The World Bank.

Verspoor, A. M. 2003. *The Challenge of Learning: Improving the Quality of Basic Education in Sub-Saharan Africa*, Summary, Discussion Paper prepared for the ADEA Biennial Meeting 2003, Paris: ADEA.

Wößmann, L. 2004. "Institutional Comparisons in Educational Production" in*: Journal for Institutional Comparisons*, 2, (4), 3-6.

Wößmann, L. 2005. "Competition, Decentralization and Accountability: Lessons for Education Policies in Developing Countries from International Achievement Tests" *Nord-Süd Aktuell*, 19, (2), 142-153.