

The relationships between schooling inputs and outputs in South Africa: Methodologies and policy recommendations based on the 2000 SACMEQ dataset

Martin Gustafsson, South Africa

Abstract

The debates around production functions for schooling systems are summarised, and a few methodologies are briefly explained. The methodologies that receive attention are stepwise selection of variables (to reject variables), factor analysis (to combine several variables into one), the basic one-level regression model and the more complex two-level hierarchical linear model (HLM). With regard to the SACMEQ dataset, a process of initial variable selection and manipulation is described, as well as the use of a new reduced set of variables in some one-level and two-level modelling of how schooling inputs influence outputs, specifically reading and mathematics scores, at the Grade 6 level of the South African schooling system. This description will pay particular attention to the importance of the education policy system. In the section on policy recommendations, better management of learner repetition, and some basic improvements with respect to classroom methodology and teacher evaluation are put forward as possible interventions with a clear positive impact on performance and with a low or even negative cost. Investments that can enhance learner performance are, from lowest to highest cost, in-service training of teachers, improvements to the physical infrastructure of schools, and adult education targeted at the parents of learners. Ensuring a basic level of access to textbooks, and regular meals for learners, are other interventions with clear performance benefits. It is recommended that policies driving interventions, for example policy benchmarks for learner repetition, focus on the school as a whole.

This paper is a summary of a Masters thesis that contains more details on the analysis. The author can be contacted at martingust@worldonline.co.za.

Note on language: The South African practice of calling pupils 'learners' and teachers either 'teachers' or 'educators' is followed here.

Length of paper: 9,131 words

Acronyms

ABET	Adult Basic Education and Training
HA	historically advantaged
HD	historically disadvantaged
HLM	hierarchical linear model
ICT	information and communications technology
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais (National Institute of Educational Studies and Research)
LSM	learner support material
SAEB	Sistema Nacional de Avaliação da Educação Básica (National Basic Education Evaluation System)
SES	socio-economic status

Introduction

A production function is a model that explains how various inputs are converted into (usually one) output within a factory, a school, or some other organisation. Explaining production has been at the heart of much economic analysis. Adam Smith extolled the value of the profit motive as the guarantor of efficiency within the production processes of the firm. Karl Marx viewed the mode of production in the capitalist firm as a source of social tensions, the revolt of the working class and ultimately the fall of capitalism.

The formulation of production functions for school ‘production’ processes comes with considerable difficulties. Firstly, we need to deal with the contention that schools are not like industrial firms producing tangible goods with a clear price. Specifically, there is the question of what the output of schooling is, or what the Y of

$$Y_i = f(X_{i1} \dots X_{in}) + u_i \quad (1)$$

represents. The response is usually that learner performance in some standardised test should be taken as the output, with the proviso that this is but one output (albeit an extremely important one) schools are expected to achieve. Secondly, there is the common criticism that the dataset being used to formulate the production function is too limited. Particularly, there is often a lack of data on what actually happens in the classroom. And wherever cross sectional data is used, there is the criticism that we cannot identify the before and after situation, and hence cannot correctly disentangle the effects of the school, the socio-economic status (SES) of learners, and flow issues such as repetition. Ideally, we would want to have data on each learner’s performance at time t_1 , the schooling and home background inputs that were brought to bear on each learner’s education between time t_1 and time t_2 , and the learner’s improved learner performance at time t_2 . The ideal is however seldom a reality, and we need to make do with data on performance at time t_1 , and on the inputs that were applied during a preceding

period, often one year. Thirdly, and this is related to the previous criticism, there is so much that is left unexplained in any school production model that it is difficult to draw hard policy conclusions on the basis of the part of the production process that we can explain. Put differently, we end up with a situation where the residual u from equation (1) matters more than the production function f in (1) above (Crouch and Mabogoane, 1998). Analyses that use the learner as the basic unit of analysis tend to explain some 50 per cent of the production process with the available input indicators, even where the dataset is relatively comprehensive and reliable (see for instance Harbison and Hanushek, 1992, and Häkkinen, Kirjavainen and Uusitalo, 2003). Unless we can come up with a more robust dataset, the only way to deal with criticisms two and three is to pin numerous caveats to our observations and conclusions.

Despite the difficulties, production function analyses in education are receiving increasing attention. This is partly due to optimism around obtaining increasingly better datasets on schooling inputs and outputs in the future. It is also clear that even with limited data, the analysis can reveal important things about what matters more and what matters less in improving educational outputs in schooling systems. This is especially true in developing countries, where the provisioning of inputs still occurs at a rather basic level, and improving inputs has a demonstrable impact on learner performance (Pradhan, 1996: 75).

The analysis has been occurring at various levels. Data on sub-national pockets of the education system have been analysed, often in conjunction with a donor-funded intervention project (Harbison and Hanushek, 1992, in the case of Brazil, and Glewwe, Kremer and Moulin, 2000, in the case of Kenya). International datasets have been used to compare the production functions of several countries (Willms and Somers, 2003, in the case of Latin America's Laboratorio dataset). And national governments undertake data collection as part

of their education monitoring programmes, and this often leads to some form of production function analysis (the Systemic Evaluation of South Africa or Brazil's SAEB).

Variable selection and manipulation

The SACMEQ 2000 dataset has 169 variables derived from the learner, educator and school principal questionnaires. The number of X values we ought to have in (1) above (or the optimal value of n) is debatable, but for many reasons we need much fewer than 169. We require closer to 10 to 20 variables for our modelling purposes. We are thus faced with two challenges. Firstly, we need to reduce the number of variables by (a) discarding many of the original 169 variables and (b) by combining closely related variables into single variables. Secondly, we need some conceptual framework, or mental model, to guide us.

We begin with the challenge of the mental model. Given that the primary aim of this analysis is to inform education policymakers at the government level (as opposed to, for instance, school principals), it seems important to consider how government policies affecting the schooling system are organised. In South Africa (but even other countries), the following 22 policy areas (first column), organised into seven categories, and their corresponding schooling 'inputs' (second column) seem to offer one logical breakdown. (The use of the word 'input' should be qualified here because of our particularly broad use of the term. It is perhaps debatable whether level of parent involvement is an input in the way that quantity of learner support materials, or LSMs, is.) It should be noted that the inclusion of poverty relief as a policy area makes the scope of the framework wider than the education policy arena.

Table I: Policy-oriented mental model				
	<i>Policy area</i>	<i>Input</i>	<i>Variable meaning</i>	<i>Variable name</i>
Educators	Teacher training (pre-service)	Quantity/quality of pre-service teacher training	Years pre-service training	yrs_preserv_math/read (E)
	Teacher training (in-service)	Quantity/quality of in-service teacher training	Days of in-service training	day_inserv_math/read (E) (y/xinservd)
	Teacher conditions of service	Educator salary and fringe benefits	Teacher SES	teacher_ses_math/read (E)
	Evaluation and rewards for teachers	Incentives for educators to perform	Evaluation intensity	teacher_eval_math/read (E) (y/xshadv)
	Teacher supply/distribution	Learner/educator ratio		class_size2_math/read (E) (y/xclsizv)
Curriculum	Curriculum	Relevance/clarity of the curriculum	Class methodology value	class_meth_math/read (E)
	School year/day	Contact time	Teacher hours in a year	hrs_year_math/read (E)
	Grade repetition	Level of learner repetition	Number of years repeated	repetition (L) (prepeat)
	School admissions and streaming policy	Level of stratification		
LSMs	Materials development	Quality of LSMs		
	Materials provisioning	Quantity of LSMs	Textbooks per learner	textbooks_math/read (L) (ptextm/r)
	ICT	Quantity of cutting edge LSMs		
Infra-structure	School construction/equipping	Quality of school buildings and equipment	Level of school infrastructure	school_infra (S)
Management	Management training	Management capacity of school principal	Principal's years of pre-service	yrs_preserv_prin (S)
	School principal conditions of service	School principal salary and fringe benefits	Principal's teaching load	prin_teach_load (S)
	Governance training	Level of community involvement	Level of parent involvement	par_involve_math/read (E)
	Provincial/district support	Quantity/quality of district support	Intensity of district support	dist_support (S)
Access issues	Scholar transport	Transport for remote learners	Proximity to urban facilities	ruralness (S) (slocat)
	School nutrition	Health of learners	Average number of meals per day	daily_meals (L)
Households	ABET	Educational support from parents	Years of schooling of parents	parent_educ (L)
	Poverty relief	Socio-economic welfare of household	Learner SES	learner_ses (L)
	Sports and culture	Level of non-school education and culture facilities		
General			Learner's gender	learner_gender (L) (psex)
			Degree of teacher latecoming	teacher_disc (S) (stchpr01)
			Learner's age in years and months	learner_age (L)

One key function of the above schema would be to assist in ensuring that we do not ignore important areas of policy intervention and that we do not over-specify single policy areas. Put differently, we should perhaps aim to have one, and only one, input variable for each policy area.

The steps followed in order to reduce the 169 SACMEQ variables to a new reduced set of variables were as follows:

(a) Establish the bivariate association of all variables. A Stata computer program was designed to gauge the association between each of the 169 variables, on the one hand, and the reading and mathematics scores, on the other (variables *ratotp* and *matotp* were used). The program ran bivariate regression analyses for all ratio variables. In the case of ordinal and nominal variables, the various codes were used to establish dummy variables, and these dummy variables were regressed against the scores. The highest R^2 obtained for each of the 169 variables was noted.

(b) Identify variables with the strongest net association. This step was also performed programmatically. After the dummy variables had been created in step (a), altogether 831 variables were obtained from the original 169 questionnaire variables. Of the 831 variables, around 200 linked to the best R^2 values obtained in step (a) were put through a backward selection process (a variant of stepwise selection) so that eventually around 25 variables with the best net associations with the scores could be identified. We can speak of a ‘net association’ because the stepwise selection approach involves gauging the significance of one explanatory variable whilst controlling for the simultaneous associations of the other explanatory variables in the model. The process was repeated for the reading scores and the mathematics scores.

It should be pointed out that the stepwise selection approach is not supported by many analysts, largely due to the fact that the method has been used irresponsibly, or whilst ignoring the social and economic dynamics of the system at hand, in our case the schooling system (Baker, 2000: 82). The use of the policy framework in the selection of SACMEQ variables helped to prevent this problem.

(c) Select the ‘best variables’ using the policy framework. An attempt was made to find variables with either a relatively high R^2 value from step (a) or a strong net association emerging from step (b) that could be linked to each of the 22 policy areas. This seemed possible for all but four of the policy areas. However, some of the variables linked to policy areas were not ideal, and the link was quite tenuous. For example, where one would ideally want a variable on the principal’s salary, or at least SES, it was decided to use the available data on the principal’s teaching load, given that this is an aspect of the school manager’s conditions of service. The fourth column in the above schema gives the names of the new variables, with L, E and S indicating whether the variable is descriptive at the level of the learner, the educator or the school. Where one original variable was used (possibly with some weighting of the coding system in the case of nominal and ordinal variables) as the new variable, the names of both the new variable and the original variable (in brackets) appear in the last column of the above schema. The third column explains the meaning of the new variable.

Three variables were included that were not clearly linkable to any one policy area, but were deemed important due to their importance in other studies and due to the fact that they displayed a high level of significance in the bivariate analysis of step (a). These three variables relate to the learner’s age and gender, and to the level of discipline and commitment amongst teachers.

(d) Combine several variables into one factor. Where it was clear that several closely related original variables were linkable to one policy area, factor analysis was used to extract a single variable, or ‘factor’, that synthesised the values of the several variables. This statistical method is commonly used in education input-output analyses (Willms and Somers, 2001: 415; Hungi, 2005: 2; Barbosa, Fernandes and Dos Santos, 2000). The new learner SES, teacher SES and school infrastructure variables were derived using factor analysis. To take an example, the variable *learner_ses* was derived from six original variables relating to the physical condition of the learner’s home and the presence of the three household items that emerged as significant from step (a). Of the six original variables, the condition of the floor yielded the highest R^2 value when regressed against the reading score – the R^2 value was 0.27 (the codes in the floor variable were weighted). When all six variables were regressed against the reading score in a multivariate model, an R^2 value of 0.41 was obtained. The single factor variable obtained from the original six variables yielded an R^2 value of 0.39. The factor thus yielded more prediction than any of the original variables, though not as much as we would have obtained had we retained all six variables.

One-level modelling of the SACMEQ data

Typically, production functions for schooling systems involve the use of the regression model without any hierarchical attributes that separate out, for instance, school level and learner level effects. Such a regression model will be referred to as a one-level regression model here, to differentiate it from the hierarchical linear model (HLM) discussed in the next section. The simple one-level regression model structures equation (1) as follows:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_n X_{ni} + u_i \quad (2)$$

Critically, the above regression model provides us with a slope coefficient for each explanatory input variable, for instance β_l for the variable X_{li} . This slope coefficient tells us

the magnitude of the change we can expect in Y_i , given a change in the value of the explanatory variable, for instance X_{ji} . The association between the input and the output is the *net* association, after we have taken into account, or ‘controlled for’, the effects of the other explanatory variables in the model (Gujarati, 2003: 205).

The new reduced set of variables described in the previous section were regressed against the mathematics and reading scores – the results are provided in the first two tables of Appendix A. Variables were excluded if they failed the 2- t rule of thumb. Moreover, school-level means of the learner-level input variables were constructed and took precedence if their significance as measured by t was greater than their learner-level counterparts. The R^2 statistics for the mathematics and reading models were 0.55 and 0.63 respectively. The SACMEQ dataset is thus capable of explaining a large portion the performance scores relative to other, similar datasets.

The third and fourth tables in Appendix A show the results of a segmentation of the model by historical disadvantage. In these model outputs, coefficient of variation (c.v.) is provided instead of the standardised beta coefficient. All the variables from the unsegmented models were used. The historically disadvantaged (HD) segment was considered to be whole schools covering the least advantaged 80 per cent of weighted learners, where the school mean of *learner_ses* was used to measure disadvantage. Segmenting the data in this manner allows us to take into account the apartheid legacy of the South African schooling system, which understandably was still a prominent feature of the schooling system just six years after the end of apartheid in 1994. A histogram of the reading scores illustrates how prominent the legacy of a divided system was in 2000.

Insert here: Figure 1: Histogram of reading score in South Africa

The distribution of scores is clearly bimodal if we graph the reading scores. The distribution is less obviously bimodal if we graph the mathematics score or the mean of the two scores, but the general pattern remains. South Africa is not the only SACMEQ country with this pattern, but it is particularly evident in the case of South Africa. What the graph indicates is that in some senses we are dealing with two schooling systems within one: a historically disadvantaged (HD) one and a historically advantaged (HA) one. Of note is the size of HA segment. The performance scores graph above, but even the histogram of *learner_ses*, suggest that the HA segment comprises some 20 per cent of the whole in terms of weighted learners. Given that white learners made up some 6 per cent of Grade 6 learners in 2000 (Annual Survey of Schools), the remaining 14 per cent would be learners from groups which had been discriminated against under apartheid. Stats SA data indicates that some two-thirds of this 14 per cent would be African learners, whilst one-third would be either coloured or Indian learners (calculated from Statistics South Africa, 2002 and 2004). The great majority of learners in the HD segment would be African.

In general, the model outputs were analysed as follows: Attention was paid to the slope coefficients wherever these were associated with t statistics greater than or equal to 2. Attention was also paid to how the variable value was derived, both in terms of the original questionnaire questions and the construction of the new variable. A change in the variable value as a result of some policy intervention was hypothesised, and modelled. The expected change in the performance scores, in terms of a percentage increase in the mean scores, was calculated. Given that the reading and mathematics scores are of a different magnitude (means of 39 and 23 respectively in the case of South Africa), it was important to indicate the improvement in percentage terms. In many instances, the policy intervention resulted in similar percentage improvements for the reading and mathematics means, confirming the validity of the input-output relationship depicted in the models. In the calculations, it was

assumed that slope coefficients that were similar for the same variable across the various models in Appendix A could be regarded as reliable indicators of the net relationship between the input and the output. In some instances, the modelling was possible using just the mean values for the whole system, or for just the HD and HA 'sub-systems', whilst in other instances values had to be manipulated at the level of individual records in the dataset. The hypothetical policy interventions and the expected performance score improvements are captured in Table II appearing at the end of this section. As part of the analysis, possible improvements to the SACMEQ questions, in the interests of better production functions, were considered.

We turn first to the policy area of the pre-service training of educators. In terms of quantity of pre-service training, the pattern is rather different for the HD and HA segments. The mean for the *yrs_preserv* variables (which capture total years of pre-service schooling and training for individual mathematics and reading educators, as well as the corresponding mean for all educators in the school) was 14.9 for HD schools and 15.7 for HA schools. The unsegmented models indicated that the overall slope coefficient was between 2.7 and 3.0, meaning that the addition of one year of pre-service training would translate into a performance improvement of around 8 per cent (reading) to 12 per cent (mathematics). The utility of the regression model should be noted here. Had we examined the relationship between pre-service training and the performance scores on their own, it might have seemed as if each additional year of pre-service training were associated with a huge improvement in the performance scores of over 45 per cent. However, when we control for the effects of other variables, the association is considerably less dramatic. What is striking is that the slope coefficients for the HD and HA models are lower than those for the general unsegmented models. Some knowledge of the apartheid educator training system should tell us that this is largely due to the fact that teacher training was unequal not just in quantitative terms, but also

qualitative terms. The higher slope coefficients of the general models are capturing both the quantitative and qualitative inequalities, whilst the slope coefficients of the separate HD and HA segments capture mainly quantitative differences. Using the various slope coefficients, we can simulate what the effects would be on performance if we upgraded educators in the HD segment to the pre-service training levels of educators in HA segment. Scores in HD schools would increase by a whole 25 per cent. However, this policy intervention is the equivalent of wiping out the entire apartheid human capital backlog with regard to educators, clearly a very ambitious task. A more realistic intervention would be to upgrade educators in the half of the system with the lowest pre-service education and training values with the equivalent of one year's training of the type received by educators in the HA segment. This intervention would result in an increase in the scores of around 3 per cent for the system as a whole, and of 5 per cent for HD schools (these figures are captured in Table II). Importantly, the analysis indicated that so-called compositional or peer effects were strong. It is not just the pre-service training of the individual educator that is important, but also the mean of the training level of all educators within a school. (In this respect, the inclusion of pre-service training questions in the SACMEQ school principal questionnaire relating to all educators in the school is useful.) The simulation discussed here involves the upgrading of all educators within a school, not just the individual mathematics and reading educators.

Given that the upgrading of educators in terms of their training offers substantial opportunities for improving performance, the effectiveness of various in-service training solutions should obviously be a key concern. Data on the quantity of in-service training received seemed to be the best data available in the SACMEQ dataset. The variable *day_inserv* is retained in the models, but the direction of the effect is ambiguous. This should not surprise us, if we take into account that in-service training offered by the state would be targeted towards worse performing schools, so to some extent we would expect exposure to

in-service training to be associated with lower scores. At the same time, we would expect greater exposure to effective in-service training to be associated with better scores within a group of schools that started off with the same baseline. It is impossible to disentangle the two effects on the basis of the SACMEQ data, so it is not possible to arrive at a function which says that x additional days of in-service training may improve performance by y per cent. However, the SACMEQ data does allow us to make some more general observations about the effectiveness of the in-service training system. Around 30 per cent of learners in HD schools were taught by educators who said they had received no in-service training during the previous three years, and the scores of these learners were lower than the HD average. This suggests that the state's targeting of in-service training was still inadequate in 2000. Moreover, it is noteworthy that around half of the educators classified the in-service training they received as 'reasonably effective', whilst some 20 per cent regarded it as 'very effective', with the latter group being associated with slightly lower scores. To some extent, we would expect educators achieving lower learner scores to value the training more. However, the data suggests that either the training is set at too low a level for a great number of educators, or that the training is of an inadequate quality (educators with better learner scores may be better equipped to evaluate the quality of the training). It seems important to expand the treatment of in-service training in future SACMEQ questionnaires in a number of ways. Firstly, differentiating between training provided by the state (or NGOs working with the state) from training initiated by schools or educators would assist in separating the selection effects from the training effects in the analysis. Secondly, data on the educators' evaluation of training received should ideally differentiate between satisfaction with the level of training (relative to the educator's needs) and the educator's view of the general soundness of the training being offered.

Data relating to the level of pedagogic advice received by educators from the school principal emerged as a relatively strong explanatory variable associated with better scores, and this data was used to construct the new variable *teacher_eval*. Less frequent meetings in this regard ('once a year' or 'once a term') were associated with better scores than more frequent meeting ('once or more a month'), probably suggesting that more structured encounters forming part of an evaluation cycle are more effective. A simulation revealed that if evaluation practices found in HA schools were introduced to all HD schools, then scores in the latter would improve by around 4 per cent.

The new variable dealing with class size, *class_size2*, is excluded from all the models due to its low net association with performance (class size was squared to enhance prediction and take into account the increasing marginal effect of class size). This seems surprising, especially given that the SACMEQ data indicates that large class sizes were still prevalent in 2000 (19 per cent of learners were in classes with more than 50 learners). The variable is strongly correlated with a number of other variables, for instance the powerful *teacher_disc* variable relating to the problem of teacher latecoming (as reported by the school principal). It is possible that the effects of large classes are being manifested through other variables relating to, for example, teacher motivation. The lack of any hard evidence for the performance benefits of decreasing class sizes (as opposed to other policy interventions) is in keeping with the findings of a number of other studies on South African schooling, for instance Crouch and Perry (2002). Another variable not retained in any of the models is *teacher_ses*, which we can regard as a reflection of the relative equality of income of South African educators, which in turn is linked to the central bargaining processes applicable to educator salaries.

If we turn to the curriculum variables, meaning those variables dealing with what happens in the classroom, we find two variables that are retained as significant and apparently strong predictors of performance: the variable *class_meth* reflecting type of teaching methodology, and *repetition*, reflecting degree of learner repetition. Turning to the first of these variables, analysis of the data from several teacher and learner questionnaire items indicated that for reading, promoting listening skills and having parents sign for homework done was associated with better scores, whilst for mathematics allowing learners to work on their own, interacting on a one-to-one basis with individual learners, assigning homework, and getting parents to sign homework books appeared to be valuable practices. On the basis of these findings, weightings for good classroom practice were created. The slope coefficients obtained, and some simulations, suggest that inserting the kinds of classroom methodology practices found in the HA segment of the system into the HD segment of the system (in other words making the HD mean equal to the HA mean) could improve reading scores by 2 per cent and mathematics scores by 7 per cent across HD schools. It should be remembered that this is net of the effects of other variables. In other words, the models are indicating that even without a dramatic improvement in the training levels of educators, we could improve the scores, in particular the mathematics scores. These mechanics should obviously not be taken too literally. Clearly, changes in classroom methodology do require some training. However, there are sufficient examples of schools which are disadvantaged in terms of, for instance, pre-service training levels, but which nevertheless achieve relatively good mathematics scores, for us to say that there are performance improvements we could expect to obtain even before we succeed in bringing about major changes to the formal training profile of educators.

The variable *repetition* has a relationship with learner performance that is more significant than that of any other explanatory variable. This is confirmed in both of the general models and in the models dealing with the HD segment. The association is always

negative, meaning more repetition is associated with lower learner performance. Moreover, the variable is more significant when the school-level average is used, than when the learner-level value is used. In other words, compositional effects appear stronger than individual effects. The slope coefficient is around 5, so if on average all learners repeat an extra year, the mean score drops by 5 points. If half of the learners repeat an extra year, then the score drops by 2.5, and so on. Simulations based on the model findings indicate that if the average years repeated at any time in the past per Grade 6 learner in HD schools (0.75 years) were reduced to the level found in HA schools (0.17 years), then scores in HD schools would improve by 12 per cent. A more modest intervention, whereby no school would have an average greater than 0.5 years, would improve the scores by 6 per cent in HD schools, and by 4 per cent in the system as a whole. If having high levels of repetition in a school is so clearly associated with lower performance, after we have controlled for other variables, the obvious question is why school principals and teachers allow such high levels of repetition. Are educators misunderstanding the dynamics of learning, is what the models indicate here misleading, or is there some other explanation? We can probably not expect programmes such as SACMEQ to deliver the answers. The issue warrants more focussed research into the matter, especially given the apparent magnitude of the effects.

The variable dealing with contact time, *hrs_year*, is not retained in all models, and was thus not regarded as a significant explanatory variable. There is thus no hard evidence from the SACMEQ data that poorer performance is caused by the shortening of the school year in certain schools.

The variable *textbooks* dealing with learner access to textbooks emerged as being significant as an explanatory variable only if it was capped at 0.5 textbooks per learner, in other words if the influence of differences below that level were taken into account. Around a

half of learners stated that they had access to their own textbook, and a ratio of 0.5 or fewer textbooks per learner applied to some 33 per cent (reading) or 40 per cent (mathematics) of weighted learners. Raising access to textbooks so that each learner shared a textbook with no more than one other learner would raise the scores by between 1 per cent and 2 per cent.

Better school infrastructure, as measured by the variable *school_infra*, is strongly associated with better learner performance. Even if we use the more conservative slope coefficients from the mathematics models, the simulation indicates that raising the quality of school infrastructure in all HD schools to that of the average HA school would improve the scores by around 14 per cent in the HD schools. Here we should bear in mind, however, that *school_infra* is strongly correlated with several other variables, in fact the strongest case of multicollinearity amongst the new variables is that between *school_infra* and the variable *ruralness*. It is thus very possible that the data on the physical infrastructure is to a large extent masking other factors relating to the location of a school in a more rural environment, for example the longer distance between learners' homes and the school, and greater levels of unemployment in the community. In fact, if we calibrate *school_infra* like *ruralness*, in other words if we give it a value of 1, 2 or 3, we find that the school infrastructure variable diminishes in importance, whilst the ruralness variable increases in importance. In fact, in the unsegmented mathematics model, both the *t* statistic and the slope coefficient for *ruralness* become greater than the corresponding statistics for *school_infra*. Clearly, the calibration of variables has an important influence. There are good reasons to believe that factors linked to the physical infrastructure of schools and their ruralness are highly important determinants of learner performance. Whilst some disentangling of the factors may be possible on the basis of the SACMEQ data, further analysis would seem necessary to clarify the dynamics. The current emphasis on the specialness of schooling in rural areas, as manifested in the

government's recent *A new vision for rural schooling* (South Africa, 2005), and in a major study by the Nelson Mandela Foundation (2005) seems justified by the SACMEQ data.

It was difficult to find data to match the four school management policy areas identified in the framework. The management capacity variable captures the school principal's years of pre-service training, and is thus clearly not an optimal indicator of training in management. In the absence of data on the principal's remuneration, the school principal's teaching load was considered an important condition of service factor. There was no data dealing with parent involvement in school governance, so instead a variable dealing with contact between the teacher and the parent was used. For the gauging of district support, there was better data available, and so the variable *dist_support* reflects the number of departmental visits to the school. None of the school management variables are retained in either of the general models, and in no case do we obtain slope coefficients from these variables that translate into a feasible improvement in the scores of more than about 1 per cent. The tenuousness of the SACMEQ variables dealing with school management (particularly issues such as the principal's assessment of district support, and the frequency and nature of management meetings between the principal, educators and parents), rather than the unimportance of school management, should be the preferred explanation. We should bear in mind Crouch and Mabogoane's (1998) argument that to a large extent it is management that accounts for the unexplained portion of the input-output function (this portion equals 37 per cent in the case of the reading scores model).

The SACMEQ data on lunches eaten by learners indicate that the proportion of learners receiving lunch on all days is 65 per cent and 84 per cent for the HD and HA schools respectively, whilst the corresponding proportions of learners receiving no lunch on any day are 8 per cent and 2 per cent. The *daily_meals* variable is retained in both the mathematics

and reading models, and the slope coefficients allow us to estimate that if all learners were to eat three meals a day, we might expect a performance improvement of around 2 per cent.

The level of education of parents is a prominent explanatory variable in all the models in Appendix A. This agrees with the findings from many other studies. If we simulated an improvement such that the level of the 20th percentile of the variable *parent_educ* was made the minimum (in other words all parents below this level would be brought up to this level), we would obtain an overall improvement in both the mathematics and reading scores of around 1 per cent. If we used the 40th percentile instead of the 20th percentile as our parent education standard, the improvement in scores would be around 3 per cent.

The variable *learner_ses*, dealing with the socio-economic level of the household apart from the parents' level of education, is also retained in all the models. If we perform a simulation similar to the one we performed for *parent_educ*, we obtain an overall improvement to the scores of less than 1 per cent when we use the 20th percentile as our standard, and of around 1 per cent when we use the 40th percentile as the standard.

A variable with a very strong association with learner performance in nearly all the models is *teacher_disc*, a variable indicating whether the school principal regards latecoming amongst educators as a problem. The problem is identified in 96 per cent of the HD segment of the system, and in 36 per cent of the HA segment. Across most models, more perceived latecoming is associated with lower scores. It is in fact remarkable that this single variable, one of the original 169 SACMEQ variables, should have such a high net association with performance. Removing the latecoming problem from the entire system would increase scores by around 20 per cent for HD schools, and 15 per cent for the system as a whole according to a simple simulation. Multivariate regressions for the other SACMEQ countries revealed that for close to half of the countries, *teacher_disc* was not a significant explanatory variable.

Moreover, the significance of the variable in the case of South Africa is substantially higher than for any other country. The possibility of some problem with this variable across the entire dataset was thus ruled out. The issue seems to warrant some closer examination. In particular, it would be important to establish the degree to which the principal's perceptions are correct, the degree to which teacher latecoming is a system-wide symptom of a lacking culture of teaching and learning, whether latecoming is having a substantial impact on classroom contact time, and to what degree we are dealing with management capacity problems on the part of individual school principals. The fact that the problem is widespread across both rural and non-rural schools would probably rule out transport problems and long distances as the key determinant of the educator latecoming problem.

Table II: Hypothetical policy interventions and expected performance change			
<i>Variable</i>	<i>Hypothetical change</i>	<i>Approx. net effect on HD scores</i>	<i>Approx. net effect on overall scores</i>
yrs_preserv	Raise the training level of educators in the half of the system with the greatest deficit by the equivalent of one year of pre-service training.	+5%	+3%
	Raise educator training of HD part of system in quantitative and qualitative terms to that of HA part of system.	+25%	+18%
teacher_eval	Raise the level effectiveness of teacher evaluations by the principal in HD schools to that in HA schools.	+4%	+3%
class_meth_math	Raise the average classroom methodology indicator in HD schools to that of the HA schools <i>with respect to mathematics</i> .	+7%	+5%
class_meth_read	Raise the average classroom methodology indicator in HD schools to that of the HA schools <i>with respect to reading</i> .	+2%	+1%
repetition	Decrease the average learner years of repetition in the 61% of the system where schools exceed the 0.5 level, to 0.5.	+6%	+4%
	Decrease the average learner years of repetition in the 89% of the system where schools exceed the average level for HA schools (0.27), to this HA level.	+12%	+8%
textbooks_math	Raise the average number of <i>mathematics</i> textbooks per learner so that each learner enjoys a ratio of at least 0.5 per learner.	+3%	+2%
textbooks_read	Raise the average number of <i>reading</i> textbooks per learner so that each learner enjoys a ratio of at least 0.5 per learner.	+1%	+1%
school_infra (N.B. closely correlated to ruralness)	Raise the level of physical infrastructure of all schools to the present average for HA schools.	+14%	+10%
daily_meals	Raise the intake of daily meals so that all learners receive all their daily meals (currently some 51% of learners do).	+3%	+2%
parent_educ	Raise the level of education of the least educated 20% of parents to the level of the 20 th percentile.	+1%	+1%
	Raise the level of education of the least educated 40% of parents to the level of the 40 th percentile.	+4%	+3%
learner_ses	Raise the SES of the least advantaged 40% of learners to the level of the 40 th percentile.	+2%	+1%
teacher_disc	Remove the problem of perceived indiscipline of educators from all schools.	+20%	+15%

Two-level modelling of the SACMEQ data

The hierarchical linear model (HLM) as put forward by, for instance, Bryk and Raudenbush (1992), is a type of linear regression model that pays special attention to the dynamics of the different levels of a system. With respect to schooling, we can think of the level of the district, the school, the class and the learner. Here only two levels, that of the school and the learner, will receive attention.

It is illustrative to begin a treatment of a two-level HLM by considering the ‘null model’, or the model devoid of any explanatory variables. The model can be represented as follows:

$$Y_{ij} = (\alpha_0 + \varepsilon_{0j}) + u_{ij} \quad (3)$$

Here the test score of learner i in school j can be thought of as some intercept, α_0 , applicable to the whole system (this would be roughly the mean score for the system), plus an additional value ε_{0j} , applicable only to school j (this is roughly the difference between the system mean and the school mean), plus the learner’s own error term, u_{ij} . The variances of u_{ij} and ε_{0j} are known as the level 1 (within-school) and level 2 (between-school) variances respectively, and are shown in Appendix B to equal 88.9 and 180.8. The proportion of the total variance existing at level 2 is known as the intra-class correlation coefficient, and can be regarded as the degree of inequality existing between schools, relative to the overall inequality of the schooling system. We can think of a zero intra-class correlation coefficient as being an ideal, insofar as this would reflect a system where there was no inequality between schools. All the inequality would exist within schools, meaning no learner would be disadvantaged on the basis of the school he found himself in.

Two things stand out about the 67 per cent intra-class correlation coefficient obtained from the SACMEQ data (using the reading score model). Firstly, the value for the statistic is high. Willms and Somers (2001: 417) regard a normal value for a developing country to be around 30 per cent (developed countries, which are generally more internally equal, would tend to have a value of around 15 per cent). Amongst the other SACMEQ countries, Uganda and Namibia have intra-class correlation coefficients that are lower than South Africa’s, yet above the Willms-Somers threshold, whilst Botswana, with 22 per cent, and Mauritius, with 25 per cent, are well below that threshold. The second thing that should be noted is that the

intra-class correlation coefficient for performance in South Africa is higher than the corresponding statistic for SES levels. Whilst the performance statistic is 67 per cent, the SES statistic is 63 per cent. What this means is that the inter-school inequalities, relative to overall inequalities, are greater with regard to performance than they are with regard to socio-economic status. Willms and Somers argue that it is important for this to be the other way round. Schools should have an equalising effect on society, so a higher intra-class correlation coefficient for performance than for socio-economic status is something one should try and reverse in South Africa.

A proper HLM is produced if we add explanatory variables to the model in equation (3) to obtain a model such as the following:

$$Y_{ij} = (\alpha_0 + X_{1j} + \varepsilon_{0j}) + \beta_2 X_{2ij} + u_{ij} \quad (4)$$

Here the school-level variable X_1 allows us, in effect, to create a separate intercept for each school. A more complex HLM would also have a separate slope coefficient, β_2 , for each school.

The ‘model with SES’ in Appendix B is like equation (4) above, but with two learner-level explanatory variables where X_2 is, and no X_1 variable. The inclusion of two household background variables allows us explain some 32 per cent of the variance in reading scores existing at the between-school level, and some 4 per cent of the variance in reading scores existing within schools (see the last two rows of Table 5). Thus despite the fact that the learner SES and parents’ level of education variables describe individual learners, they tell us more about the performance differences between schools, than about the differences between learners in schools. The compositional or peer effects of these variables are clearly important. It is not just the learner’s own home background that influences that learner’s performance, but also the home background of the other learners in the same school.

If we add another six variables that appeared as strong predictors of performance in the one-level model, we obtain the ‘full model’ of Appendix B. Four of these variables are at the school level, in other words at the position of variable X_i in (4) above. In the full model, we find that most (73 per cent) of the between-school variance has been explained, whilst the model has only succeeded in explaining some 9 per cent of the within-school variance. With the full model, then, we are left with a situation that is the reverse of the null model – we now have more unexplained residual variance at level 1 (81.2) than at level 2 (48.4). The basic pattern is the same as the one found by Barbosa and Fernandes (2001) in their analysis of Brazil’s SAEB data.

The slope coefficients obtained using the two-level model are roughly equal to those obtained in the one-level models. There appear to be no differences large enough to substantially change the discussion of the input-output relationships in the previous section. The new information provided by the two-level model relates to how performance inequalities are structured in the schooling system, and what the likely impact would be of policy interventions on these inequalities. This receives attention in the next section.

Policy implications

We now turn to how the foregoing analysis translates into recommendations for what government can do to improve performance, and the equality of performance, in a sustainable manner, given its limited human and financial resources. It would be important to make recommendations about the optimal utilisation of some increases to the public expenditure envelope, but clearly recommendations based on unrealistic budgetary growth would be of little practical value.

Some caveats are in order. Firstly, the SACMEQ data describes schooling at the Grade 6 level in 2000. The South African schooling system is undergoing continual and rather

fundamental change, and the situation in 2005 would clearly not be the same as in 2000.

However, it is also likely that the production function, or the dynamics whereby certain inputs influence performance more than others, would not have changed substantially since 2000.

Secondly, no dataset can claim to offer a definitive picture of how education works. There are a number of datasets in South Africa that lend themselves to production function analysis, and policy changes should ideally be based on salient findings that are repeated across several analyses. What we are dealing with here, in other words, is an input into a wider debate.

Thirdly, it should be remembered that SACMEQ covered 168 schools and 3,163 learners, or 0.3 per cent of the Grade 6 population. This is a small sample, but it is adequate to render statistically significant results, on condition, of course, that the sampling methodology was sound, which, judging from the SACMEQ documentation, it was.

In economics terms, technical efficiency is attained when it is impossible to produce more outputs with the given bundle of inputs. It is useful to first consider those parts of the models that imply performance improvements with no change, or a very minimal change, to the bundle of inputs.

The SACMEQ data suggest that the management of learner repetition is crucial for improving performance scores. The repetition variable emerges as the single most significant variable explaining performance, and the expected impact of any policy or management change in this regard is large. South Africa has a repeater policy which states that no learner should repeat a phase more than once. Around 15 per cent of learners were exceeding this level of repetition in 2000 according to the SACMEQ data. But the matter is more complex than eliminating this contravention of the policy threshold. Even if no learners exceed the threshold, there are costs and benefits associated with different levels of repetition. For example, a mean of 0.1 repeated years per learner in a school needs to be differentiated from a

mean of 0.2 repeated years. The SACMEQ data indicates quite clearly that the levels of repetition prevailing in 2000 were not good for the learner performance of the repeaters themselves, but also the peers of the repeaters. A better management of repetition is a cost-saving intervention that allows for better resourcing of other aspects of schooling. This should make the intervention a particularly attractive one in any drive to improve learner performance scores. The reduction of repetition is perhaps best translated into smaller class sizes. Large classes, in particular classes exceeding 50 learners, though not clearly associated with worse learner performance in net terms, are still arguably undesirable.

Technical efficiency enhancements also seem possible with respect to teaching methodology. This is particularly so in the case of mathematics. The SACMEQ data suggests that allowing learners to work more on their own, more one-to-one interaction between teachers and learners, more homework assignment, and greater insistence that parents sign for homework done, are particularly effective levers to improve the mathematics scores. Some direct training of educators would be required, but advocacy, for instance through mass media programmes aimed at educators, could be regarded as an important implementation tool. The data also indicate that more effective professional encounters between the school principal and teachers can improve performance noticeably. In this regard, the finding that fewer rather than more encounters have a greater impact is worth taking into account. It would seem that there is a need for such encounters to be linked to clearly structured educator appraisal cycles. In South African policy terms, this would largely mean ensuring that the Integrated Quality Management System is correctly implemented.

Turning to interventions involving additional budgets, we should differentiate between more once-off investments, which in turn can be divided into physical and human capital investments, and increases to the current expenditure envelope. It is moreover important to

gain some sense of the relationship between the costs and benefits. Very rough calculations, the details of which are not provided here, were performed to estimate what it would cost per learner to achieve a 1 per cent increase in the performance scores (for example an increase in the mean reading score from 39.0 to 39.4), using interventions relating to the different inputs.

The calculations suggested that the most cost-effective capital investment was investment in the knowledge and skills of educators, through in-service training programmes. Roughly, a 1 per cent improvement in the scores could be ‘purchased’ through a once-off investment equalling 20 rand per learner. Obviously, the calculation is heavily dependent on a number of assumptions, in addition to the findings of the models discussed above. One key assumption is that the in-service training be of an adequate quality. Arguably, in 2000 all in-service training was not of an adequate standard. Only some 20 per cent of educators regarded the training they received as ‘very effective’. The fact that in 2000, some 28 per cent of educators, who were associated with below average learner scores, had not received any in-service training during the previous three years, underlines the opportunities offered by in-service training as a lever to improve performance.

Improving the physical infrastructure of schools to achieve a 1 per cent improvement in scores was estimated to cost 100 rand per learner, thus making this intervention substantially more costly, relative to benefits, than in-service training. Importantly, the correlation between the physical infrastructure of schools and their degree of ruralness is particularly high, so it is difficult to disentangle the two effects. Where schools are rural and isolated, making schools more accessible to learners through scholar transport, the provision of bicycles or a better transport infrastructure could be at least as important as improving the physical infrastructure of the school.

The educational level of parents has a clear impact on the school performance of the learner. This raises the possibility of improving learner performance by targeting the parents of learners through ABET programmes. This intervention appears to be particularly costly, with a 1 per cent improvement in the learner scores costing some 250 rand per learner.

The most cost-effective current expenditure intervention appears to be to ensure that all learners have at least a minimally acceptable level of access to textbooks. The SACMEQ data appears to confirm the finding (Glewwe et al, 2000) that there is no substantial performance difference associated with having a textbook/learner ratio of 1:1, as opposed to a ratio of 1:2. However, below the 1:2 level, there is clearly a detrimental effect on learner performance. In 2000, at least one-third of learners were below this critical level. The rough calculations that were performed indicated that a 1 per cent improvement in the scores could be brought about by increasing average per learner expenditure on textbooks by just 5 rand per year.

On the basis of the association between daily meals and learner performance found in the SACMEQ data, it was estimated that at an additional cost of 150 rand per learner per year, a performance improvement of 1 per cent could be attained. Obviously, there are arguments that go beyond the learner performance argument favouring the elimination of all learner under-nourishment. The SACMEQ data indicates that in 2000 as many as 35 per cent of learners in historically disadvantaged schools were not eating a lunch on every day of the week.

Across 85 per cent of the SACMEQ sample, school principals reported that they experienced problems with teachers arriving late at school. The net effect of this variable on performance was found to be particularly strong. Removing the problem was associated with a whole 15 per cent improvement in learner scores. It is not possible to tell from the data what

intervention would be required to tackle this problem, or even what the precise nature of the problem is. If the fundamental problem is one of poor management, then the solution would be management training, which carries a relatively low cost. If the fundamental problem is a lack of motivation amongst educators associated with matters such as large class sizes and poor physical infrastructure, then clearly the implied costs are much greater.

A two-level analysis of the SACMEQ data indicates that between-school performance inequality is particularly high in South Africa. Moreover, the schooling inputs that have received attention in this analysis have a larger impact on this between-school inequality, than on within-school inequality. Even changing the values of variables specific to individual learners, such as years repeated, has a greater impact on the inequality existing between schools than the inequality existing within schools. This emphasises the importance of interventions that treat the school holistically. There is a need for policies and benchmarks specifying acceptable and permissible levels of repetition for whole schools, or whole classes. Similarly, in-service training needs to focus strongly on the upgrading of educator skills in whole schools, given the powerful educator peer effects that influence learner performance. This is, in fact, the premise of the Department of Education's Whole School Evaluation programme.

A final caveat is needed. Whilst it is useful to decompose the school production process according to its various inputs, and assign net impact values to each input (to a large degree this is what this analysis has been concerned with), the inter-dependencies of the various inputs cannot be ignored. The models presented here are too rudimentary to describe, for instance, the degree to which the impact of input X_1 depends on maintaining input X_2 at a critical minimum level. Such dynamics can and should be modelled. These dynamics should

also caution us against over-emphasising single inputs, to the exclusion of others. Not everything is equally important, but at the same time there is no single magic ingredient.

Appendix A

Table 1: The 'best' reading model (Stata output)

$R^2=0.631$	$n=3045$	$F=323$	$level=0.95$
dependent var: read_score			
	<i>coefficient</i>	<i>beta coeff.</i>	<i>t stat</i>
learner vars			
learner_ses	0.66	0.131	8.0
learner_age	-1.12	-0.104	-8.8
daily_meals	1.11	0.047	4.1
textbooks_read	3.63	0.038	3.2
educator vars			
yrs_preserv_read	3.07	0.089	5.9
hrs_year_read	0.00	-0.071	-5.4
teacher_eval_read	0.35	0.065	5.1
class_meth_read	0.32	0.041	3.3
par_involve_read	0.21	0.034	2.7
day_inserv_read	0.01	0.023	2.1
school vars			
school_infra	1.23	0.212	9.1
teacher_disc	-5.04	-0.106	-7.0
ruralness	1.07	0.053	2.6
learner vars (school mean used)			
repetition	-6.21	-0.151	-10.3
parent_educ	0.58	0.112	6.5
learner_gender	14.20	0.091	7.9
_cons	-15.34		-1.9

Excluded: teacher_ses_read, class_size2_read, yrs_preserv_prin, prin_teach_load, par_involve_read, dist_support.

Table 2: The 'best' mathematics model (Stata output)

$R^2=0.550$	$n=3005$	$F=244$	$level=0.95$
dependent var: math_score			
	<i>coefficient</i>	<i>beta coeff.</i>	<i>t stat</i>
learner vars			
parent_educ	0.18	0.091	6.1
learner_ses	0.23	0.079	4.4
learner_age	-0.27	-0.042	-3.2
educator vars			
yrs_preserv_math	2.69	0.132	7.5
class_meth_math	0.50	0.104	6.5
teacher_eval_math	0.24	0.069	4.9
day_inserv_math	-0.01	-0.029	-2.3
school vars			
teacher_disc	-5.12	-0.184	-10.2
school_infra	0.51	0.150	7.5
dist_support	-0.05	-0.076	-5.1
prin_teach_load	0.09	0.060	4.4
learner vars (school mean used)			
repetition	-4.24	-0.175	-11.0
daily_meals	1.96	0.076	5.2
learner_gender	4.20	0.045	3.6
textbooks_math	2.51	0.036	2.7
_cons	-23.76		-4.2

Excluded: teacher_ses_math, class_size2_math, hrs_year_math, yrs_preserv_prin, prin_teach_load, par_involve_math, ruralness.

Table 3: Reading score model by historical disadvantage (Stata output)

	Historically disadvantaged			Historically advantaged		
level=0.95	R ² =0.325	n=2514	F=75	R ² =0.483	n=531	F=30
dependent var: read_score	coefficient	c.v.	t stat	coefficient	c.v.	t stat
learner vars						
learner_ses	0.54	0.54	6.3	1.52	0.03	5.4
learner_age	-1.07	0.02	-8.3	-2.93	0.00	-4.9
daily_meals	0.91	0.10	3.4	3.79	0.02	3.3
textbooks_read	4.86	0.24	4.1	-0.46	0.08	-0.1
educator vars						
yrs_preserv_read	1.96	0.00	3.5	2.80	0.00	1.0
hrs_year_read	0.00	0.27	-5.8	0.00	0.07	-0.5
teacher_eval_read	0.26	2.88	2.8	0.43	0.50	2.8
class_meth_read	0.55	0.16	5.0	-1.07	0.07	-3.6
par_involve_read	0.17	0.86	2.0	-0.33	0.18	-1.1
day_inserv_read	0.01	3.12	2.1	0.01	1.60	0.5
school vars						
school_infra	1.17	0.62	8.0	-0.01	0.03	0.0
teacher_disc	-1.22	0.02	-0.8	-6.97	1.80	-4.5
ruralness	1.54	0.22	3.3	1.36	0.02	1.0
learner vars (school mean used)						
repetition	-5.22	0.24	-8.5	-2.61	1.09	-0.5
parent_educ	0.33	0.06	3.3	2.17	0.01	5.6
learner_gender	11.33	0.04	5.7	39.12	0.04	5.7
_cons	-0.25		0.0	-24.40		-0.5

Excluded: teacher_ses_read, class_size2_read, yrs_preserv_prin, prin_teach_load, par_involve_read, dist_support.

Table 4: Mathematics score model by historical disadvantage (Stata output)

	Historically disadvantaged			Historically advantaged		
level=0.95	R ² =0.163	n=2479	F=32	R ² =0.431	n=526	F=26
dependent var: math_score	coefficient	c.v.	t stat	coefficient	c.v.	t stat
learner vars						
parent_educ	0.09	0.18	3.4	0.42	0.03	3.4
learner_ses	0.08	0.54	1.6	1.41	0.03	5.6
learner_age	-0.20	0.02	-2.8	-1.44	0.00	-2.8
educator vars						
yrs_preserv_math	1.03	0.00	3.1	4.44	0.00	1.9
class_meth_math	0.35	0.17	5.0	0.32	0.06	1.0
teacher_eval_math	0.25	4.07	4.2	0.08	0.98	0.6
day_inserv_math	-0.01	4.01	-1.5	0.00	1.32	-0.1
school vars						
teacher_disc	2.10	0.02	2.6	-5.61	1.80	-4.5
school_infra	0.44	0.62	7.3	-0.38	0.03	-0.8
dist_support	-0.02	1.08	-1.8	-0.03	0.57	-1.2
prin_teach_load	0.00	0.61	0.2	0.47	1.19	5.6
learner vars (school mean used)						
repetition	-3.17	0.24	-9.2	0.49	1.09	0.1
daily_meals	1.12	0.03	3.5	24.71	0.00	4.5
learner_gender	3.26	0.04	2.9	-7.51	0.04	-1.5
textbooks_math	3.52	0.18	3.8	3.36	0.11	1.2
_cons	-2.84		-0.5	-103.30		-3.0

Excluded: teacher_ses_math, class_size2_math, hrs_year_math, yrs_preserv_prin, prin_teach_load, par_involve_math, ruralness.

Appendix B

Table 5: Two-level model (HLM output)¹

dependent var: read_sc	Level 1 units		3045
	Level 2 units		162
	<i>null model</i>	<i>model with SES</i>	<i>full model</i>
Fixed effect	<i>coefficient (t stat)</i>		
For intercept β_0			
intercept α_0	39.02 (25.3)	30.87 (27.4)	-42.22 (-1.9)
slope <i>s_infra</i>			1.99 (7.0)
slope <i>t_disc</i>			-6.19 (-2.9)
slope <i>preser_r</i>			4.57 (3.3)
slope <i>t_eval_r</i>			0.58 (3.2)
For slope <i>repetit</i>			
intercept			-2.27 (-9.4)
For slope <i>meals</i>			
intercept			1.37 (4.9)
For slope <i>p_educ</i>			
intercept		0.30 (6.9)	0.25 (6.4)
For slope <i>l_ses</i>			
intercept		0.84 (7.0)	0.64 (5.7)
Random effect	<i>variance (p value)</i>		
For intercept β_0	180.8 (0.000)	123.0 (0.000)	48.4 (0.000)
Level-1	88.9	85.1	81.2
Intra-class correlation coeff.	67.0%	59.1%	37.4%
Level 1 explained	0.0%	4.3%	8.7%
Level 2 explained	0.0%	32.0%	73.2%

¹ The software used for this modelling of the data is the student version of HLM 6.0, available on the Internet at <http://www.assess.com/Software/HLM.htm#demo> (accessed 20 December 2004). Three limitations should be noted. Firstly, the HLM software requires variable names to be no longer than eight characters, so the variable names used in the one-level models had to be shortened. Secondly, the student version of HLM limits the number of variables that may be inserted into the model. Hence only eight explanatory variables appear in the 'full model'. Thirdly, the student version provides the random effect statistics only for unweighted data. Hence the variance statistics provided here are those for unweighted SACMEQ records. This would not affect the findings substantially. An examination of the residual files produced by the software (they are produced whether one weights or does not weight records) indicates that the differences would not have been great. The intercepts and slope coefficients are based on weighted records.

References

- Baker, B.D. 2001. Can flexible non-linear modeling tell us anything new about educational productivity? *Economics of Education Review*, 20: 81-92.
- Barbosa, M.E.F. & Fernandes, C. 2001. *A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em Matemática dos alunos da 4a série*. [Internet] Curitiba. Available from: <<http://www.mec.es/cide/rieme/documentos/barbosa/Barbosa.pdf>> [Accessed 9 July 2004].
- Barbosa, M.E.F., Fernandes, C., Dos Santos, D. et al. 2000. *Modelagem do SAEB-99: Relatório técnico: Redução da dimensão da base de dados do SAEB99*. Obtained through communication with INEP.
- Bryk, A.S. & Raudenbush, S.W. 1992. *Hierarchical linear models*. Newbury Park: Sage.
- Crouch, L. & Mabogoane, T. 1998. When the residuals matter more than the coefficients: An educational perspective. *Journal of Studies in Economics and Econometrics*, 22(2).
- Crouch, L. & Perry, H. 2002 *Determinants of learning and learning change in DDSP schools' Grade 3 numeracy and literacy assessment results*. District Development Support Programme, Pretoria. Unpublished mimeograph.
- Glewwe, P., Kremer, M. & Moulin, S. 2000. *Textbooks and test scores: Evidence from a prospective evaluation in Kenya*. [Internet] Washington: World Bank. Available from: <http://post.economics.harvard.edu/faculty/kremer/webpapers/Textbooks_Test_Scores.pdf> [Accessed 5 June 2004].
- Gujarati, D.N. 2003. *Basic econometrics*. New York: McGraw-Hill.
- Häkkinen, I., Kirjavainen, T. & Uusitalo, R. 2003. School resources and student achievement revisited: New evidence from panel data. *Economics of Education Review*, 22: 329-335.
- Harbison, R.W. & Hanushek, E.A. 1992. *Educational performance of the poor: Lessons from rural Northeast Brazil*. Oxford: Oxford University Press.
- Hungi, N. 2005. *Factors influencing Standard 6 pupils achievement in Kenya: A multilevel analysis*. [Internet] Harare, SACMEQ. Available from: <<http://www.sacmeq.org>> [Accessed 15 April 2005].
- Nelson Mandela Foundation. 2005. *Emerging voices: A report on education in South African rural communities*. [Internet] Pretoria: HSRC. Available from <<http://www.hsrcpublishers.ac.za>> [Accessed 24 February 2005].
- Pradhan, S. 1996. *Evaluating public spending: A framework for public expenditure reviews*. [Internet] Washington: World Bank. Available from: <<http://www.worldbank.org/participation/s&pa/spathematic.htm>> [Accessed 5 October 2004].
- South Africa: Department of Education. 2005. *Report of the Ministerial Committee on Rural Education: A new vision for rural schooling*. Pretoria. Available from: <<http://www.education.gov.za/content/documents/677.pdf>> [Accessed 15 June 2005].
- South Africa: Statistics South Africa. 2002. *Earning and spending in South Africa: Selected findings and comparisons from the income and expenditure surveys of October 1995 and October 2000*. Pretoria. Available from: <<http://www.statssa.gov.za>> [Accessed August 2005].
- South Africa: Statistics South Africa. 2004. *Census 2001: Primary tables South Africa*. Pretoria. Available from: <<http://www.statssa.gov.za>> [Accessed March 2005].
- Willms, J.D. & Somers, M. 2001. Family, classroom and school effects on children's educational outcomes in Latin America. *School effectiveness and school improvement*, 12(4): 401-445.

Items to be included within body of the paper

Figure 1: Histogram of learner reading score in South Africa

