No. 191 March 1997 Discussion Papers

Statistics Norway
Research Department

Liv Belsby and Bjørn K. Wold

Primary Schooling in Zambia Squeezed at Community and Household Level

Discussion Papers No. 191 • Statistics Norway, March 1997

Liv Belsby and Bjørn K. Wold

Primary Schooling in Zambia Squeezed at Community and Household Level

Abstract:
A part of the current economic reform program in Zambia, is to increase standard, efficiency and equity in the primary school sector. This paper studies primary school attendance. A logistic regression analysis is used to show that community level and household level variables affect the likelihood of attending primary school for children between seven and thirteen years of age. A number of community level variables have a significant effect and are also interesting from a policy point of view. Household level variables have even stronger effects. A policy conclusion: There is a need to introduce a widespread scholarship program to ensure equity and efficiency. A scientific conclusion: A multilevel analysis contributes additional insights.

A final policy conclusion: There is a need to introduce a widespread scholarship program to ensure equity and efficiency. A final scientific conclusion: A a dedicated multilevel analysis might yield additional insights.

Keywords: Primary school attendance, Zambia, logistic regression model, community level, household level.

JEL classification: C35, I21, O15.

Acknowledgement: We are grateful to Nils Henrik Risebro and Jan Bjørnstad for useful comments.

Address: Liv Belsby, Statistics Norway, P.O.Box 8131 Dep., N-0033 Oslo. E-mail: lbe@ssb.no
Bjørn K. Wold, Statistics Norway, P.O.Box 8131 Dep., N-0033 Oslo. E-mail: bkw@ssb.no

1. Introduction

1.1. Background

As other post-colonial African countries Zambia gave priority to education, and the education system was substantially expanded during the 1960s and the 1970s. By independence in 1964, there were 378 417 pupils enrolled in Primary schools. Ten years later this number had increased to 858 191 (Kelly, 1991). After the fall in copper prices and the increase in oil prices around and after the mid 1970s, Zambia experienced a dramatic decline in the economy. Assuming the economic decline was conjunction, the Zambian Government implemented an expansive economic policy. From 1978 it was obvious that the economic decline was structural. The government still continued with an expansive economic policy and tried to retain a welfare state, now financed by borrowed money. The Zambian government never really admitted the severe situation. Official welfare provisions and means were retained on paper but deteriorated in practice.

The resource allocation to education was reduced yearly throughout the 1980s. The process started by cutting investment, no new schools were built with government funding, no extensions, and then no maintenance nor rehabilitation. Recurrent expenditures were protected for some years, but the funding for any type of supplies, books and teaching materials vanished over the years. In 1991, 97 per cent of the education sector budget was spent on salaries and allowances, leaving very little for teaching materials and books. Due to a well organised teachers union, with considerable political support and power, salaries were better protected, but in the following years inflation reduced salaries as well. The economic hardship, the rapid population growth, and the young population, put a heavy burden on the education system, resulting in overcrowded classes. Due to migration from provinces to district centres, provincial capitals and cities in the early 1980s (Silanda, 1988), urban areas in particular have a too low capacity. Due to reduced migration and heavy population increase in the last decade, overcrowding is now a nation-wide problem. This is partially "solved" by not admitting the full cohort of 7 year old children. Obviously such a "solution" creates a cumulative problem in the consecutive years.

Since 1985, the total budget allocation to the education sector has decreased, in absolute terms, as a percentage of GNP and as a percentage of the public budget. The MMD Government (voted into power in 1991 and reelected in 1996) has however introduced two policy changes. First, they have shifted the priority within the education sector from tertiary and secondary to primary schools. Second, they have introduced a limited cost recovery and privatisation of education. There is still a shortage of teaching materials, textbooks and notebooks. As other employees (in the public sector), teachers are now experiencing a sharp decline in their real wages (World Bank, 1993). To compensate for their reduced income teachers are often giving private tuition, leaving less time for their regular job. The poor quality of primary schools is by itself worrying since it might reduce the quality of education provided, the expected return to parents and pupils, and hence their motivation. The study by (Booth et al, 1995) of four poor communities in Zambia and the comparative study by (Moser, 1996) from poor urban communities in four countries including Lusaka in Zambia, warns against the fact that a growing numbers of households may not find the necessary resources to keep their children in school. Especially if the school fees are further increased.

Generally access to school enrolment has been and is better in urban areas (CSO, 1994). Also economic, socio-economic, and cultural factors, all of which affect attendance rate, do vary between urban and rural areas. The line-of-rail areas along the old colonial railway from South-Africa, through Livingstone and Lusaka to the Copperbelt areas and cities like Ndola and Kitwe, have for decades been the winners. These areas have been in the forefront of both resource allocation and of economic, social and cultural development. Rural areas in the line-of-rail provinces have gained by being located in the hinterland of these urban areas and served the markets to mutual benefit. The real losers have

been remote areas, including remote district centres, their hinterland and the vast remote country-side. They have managed to retain their primary school teachers (but often a lower rate of fully educated teachers), but short of maintenance, equipment and material, they face a severe situation. The rift between central and remote areas grew larger due to the structural adjustment policy. Hence small and medium scale farmers switched back to a subsistence-like farming strategy (Wold et al., 1996). They can no longer afford school-fees for all their children, nor afford to restrain from using the labor of their children in the fields and at home.

The socio-economic condition of the family, often expressed by the household income, the household expenditure or the education of the parents, has been known to influence the school attendance of the children in the household. Households where one or more adults are educated are more likely to send their children to school (Behrman, 1993).

1.2. Model approach

In the present paper we model the probability that a child attends school as a function of factors at different levels which according to other studies and theoretical outlines might affect school attendance. We have studied the relation between the school attendance rate and centrality, average cash income in the community (proxy for economic opportunities), school expenses, distance to school, some quality indicators for the school, the distance to roads of different types, and distance to public service institution, as well as other variables measuring the economic condition of the household, the education of the mother and father of the child(ren), the age of the head of the household, and the age and the sex of the child. The data build a hierarchical structure, with three levels; the community, the household and the child. Some of the factors are found simultaneously and significantly to influence the probability of attending school. They are kept in the model. We present the estimated model, and the attendance rate for various levels for the significant factors.

Including factors not only at household and person levels, but also at community level, is important for several reasons. Firstly, school attendance varies from community to community. This is likely to be due to measurable characteristic of the community, for example the school expenses, the quality level the community manage to keep the school on, how far a considerable number of the pupils live from the school, or on the maintenance of the school building. Second, the adjustment policy may influence the economy of the community, and thereby also these factors. Hence detailed knowledge of the mechanism linking these factors with school attendance is important, both from a policy standpoint and for targeting[1]

In section two we present the data and the variables included in the study, and in section three we describe the statistical model and the estimation. In section four we present the estimated model with the factors found to be significantly linked with school attendance.

2. The CS and the PSII data

The data sources are the Community Survey - CS (CSO forthcoming) and the Priority Survey II - PSII (CSO 1994), both data sets were collected 1993. Both the CS and PSII data sets were surveyed by The Central Statistical office in Lusaka, Zambia, during April to June 1993. The hierarchical structure of

[1] Another reason to include variables on both the community and household levels is to reduce confounding. Confounding means that the variable we regard, which apparently causes the obsereved pattern, is not the true or the only cause. The effect may be caused partly by a variable correlated both with this variable and school attendance. An example from our data may be confounding of quality variables for the school and the socioeconomic condition of the households in the community. These variables are likely to be correlated. Discarding the quality variables for the school may lead to overinterpretation the meaning of socioeconomic condition.

the combination CS-PSII data provides an opportunity to simultaneously analyse how factors on community, household and child level are linked with school attendance.

Concerning the quality of the data; the CS data unfortunately contains a considerable amount of missing observations in some of the variables. This effects all variables on the community level, except the centrality variable, which is recorded by the PSII. Consequently, most of the variables in our model are recorded in the PSII survey. The age group in the study is from seven to thirteen years, i.e., pupils who attend primary school if they have had normal study progression. Approximately 10000 observations clustered within 631 communities are included in the study. The explanatory variable are described below.

2.1. Variables at community level

Cash income may be a major determinant of school attendance. As a indicator for this we use the average wage for men and women in agricultural and non-agricultural activity in the community.

School expenses have risen the last years, and it is a subject of debate whether high school fees may reduce school attendance. However, it is also claimed that the increased income for the community gives better schools, which again may in the long run have an positive effect on education. In the short run the higher quality of a school may motivate the parents to give high priority to school attendance. School expenses is deflated by dividing by the average total income of the households in the community. The average income is calculated from the households included in the PSII survey. The majority of the children walk to school, and long distance to the school may influence a child's access to school. Thus distance (in kilometers) is included in the model.

As quality measures for the school we have used separate / not separate classes, whether the teachers are working extra shifts, the age of the school, maintenance of roof, building, and furniture condition.

Further we have included whether text and notebooks are available at the local market.

As mentioned above; geographical centrality is linked with important characteristics of a community as economic opportunities and public services. Centrality may through these (and other factors) influence the school attendance. The area along the railway from Livingston to the Copperbelt; the line-of-rail, is an example. The centrality variable is defined according to distance to cities, towns, province capitals and the Line of rail. The following classification is used: Lusaka, Ndola and KitweRural areas less than 50 km Lusaka; Ndola and Kitwe, other cities and towns along the line-of-rail, areas less than 50 km from other cities and towns along the Line of rail, towns, other province capitals, areas less than 50 km from other towns and provincial capitals, district centres and other rural areas.

Access to transport is measured with distance to the nearest dry, all season or tarred road, distance to the nearest of public or private road transport, railway stop or water way transport. We also include distance to the nearest health clinic or hospital.

2.2. Variables at household level

The economic condition in the household is measured by the total household expenditure and the total household expenditure pr. adult equivalent.

The human assets are included by the variables mother's and father's education, if mother and father father present, and number of dependants between seven and thirteen years.

The age of the head of the household is included in the analysis, since this variable may indicate whether there is a cohort effect, possibly mixed with the effect of the age of the head of the household.

The physical assets are measured by number of rooms per person and source of drinking water.

2.3. Variables for the child

As a measures for the opportunity cost we use whether the child is active in agricultural activity or not. The age and sex are also included.

3. Modelling school attendance

3.1. The model

School attendance is coded with a binary variable which takes the values 1 when the child attends school and 0 otherwise. The probability p_{chp} that child p in household h in community c will attend school is modelled as a logistic function of the variables described in the chapter above, i.e.,

$$p_{chp} = \frac{1}{1 + \exp(-\beta' X_{chp} - \theta' Y_{ch} - \gamma' Z_c)}, \qquad (0)$$

where Z_c is a vector of community level variables, such as school expenses and centrality. And where Y_{ch} and X_{chp} are vectors representing the household and person specific variables, such as total household expenditure, the age of the head of the household and the age and sex of the child (Behrman 1993). Most of the variables in the model are categorical. They are implemented by help of indicator variables taking only the values zero and one.

The PSII and the CS data build a hierarchical structure with the observations clustered within 631 communities. Consequently the observations will most likely be correlated due to cluster effects within the community. Or more precisely, the variation between communities which is not accounted for by the explanatory variables, will give correlation between observations within the community. Moreover, children from the same family are also likely to be correlated, given the explanatory variables on household level. Among others (Aitkin, 1985) discusses this cluster effect, and proposes to model it by random components. He also points out that ignoring the existing cluster effects leads to too low p-values.

Including random components and thereby modelling the possible correlation within communities may give a spin-off effect. A strong correlation may indicate that there are factors associated with the community, which give additional information on the variation in school attendance. And, such effects should have been covered by the model.

3.2. Estimation

The logistic model used here ignores the cluster effect caused by the random components that are implicit in the hierarchical structure of the data. We estimate the maximum likelihood parameters by a standard method for logistic models using SAS. Ignoring cluster effects often leads to too low p-values. The unbiasedness of the parameter estimates themselves are normally minor, see e.g. (Anderson & Aitkin, 1985). The variable selection was based on the p-values according to a goodness of fit test. Variables with p-values higher than approximately 1 per cent were, with one exception, removed from the model. This low level was chosen to compensate for ignoring the cluster effect.

The model presented in the next section contains variables found to be significant. However, it should be kept in mind that the Pearson Chi-square test for overall fit of the model is 8945 with 8676 degrees of freedom, yielding a *p*-value of 2.15 per cent, indicating that the model fit is not the best, in spite a thorough estimation procedure, seeking the most important variables, the optimal grouping and transformation of explanatory variables. Most likely, other factors, not measured and maybe difficult to measure, influence school attendance. The model estimation is based on unweighted data.

To reduce the influence of outliers on the estimates observations, with values larger than the 95 per cent percentile on the variable total household expenditure pr. adult equivalent, or number of rooms pr. household size were excluded from the estimation. Also observations with values bigger than the 99 per cent percentile on the variables number of dependants aged between seven and thirteen and school expenses deflated by the average total household income in the community were removed. Furthermore, observations with twenty or more kilometers to school were removed.

4. Estimated model for school attendance

4.1. Interpretation

The estimating procedure described in the above chapter leads to a model presented in the table below. The interpretation of the value of the coefficients is that the bigger, the more positive impact on school attendance. Most of the variables are categorical. They are implemented in the model such that element which corresponds to explanatory variable *k* in the term $\beta' X_{chp}$ equals β, and analogous for the terms $\theta' Y_{ch}$ and γZ_c in equation (0). For the continuos variables this means that the probability of school attendance increases with increased value of the variable. The probability of school attendance can be calculated using equation (0). The unweighted estimated probability of school attendance tends to give a higher estimate of school attendance than estimates based on weighted data. This may be explained by some overrepresentation of urban households.

Odds ratios, $\exp(\beta)$, are presented in the table. When the events occur seldom, odds ratio for, say, *X=1,* approximates the relative risk, *Pr(school attendance | X=1)/Pr(school attendance | X=0)* $=p_1/p_0$. However, probability of school attendance mostly varies from sixty to ninety, and consequently odds ratios will depend on the level of the probabilities. Therefore we chose to compute the estimated probability when all variables except the variable of interest are kept unchanged. The reference level, p_0, is the level which corresponds to $\beta = 0$. The right column contains the weighted relative frequency of school attendance for various levels of the explanatory variables.

We are not only interested in *which* variables influence the probability of school attendance, but also how many children are likely to be effected by the explanatory variables at various levels. The relative frequencies displayed in the table provide rough estimates for the proportions of the children to be affected.

Table 1. Estimated coefficients in the model(0), odds ratios, p-values, estimated probability of school attendance and weighted attendance rates for various levels of the explanatory variables, and estimated proportion of children to be affected. The estimates are based on the PSII and CS survey

Variable	Coeffi-cient	Odds ratio	Esti-mated proba-bility, %	p-value in %	Esti-mated propor-tion %	Weighted atten-dance rate
Intercept	-3.95			0.01		
Community level						
Distance to school						
0-2 km	-0.09	1.09	73.35	0.00	75.1	76.42
3-5 km	-0.08	0.92	69.91	60.93	11.8	70.00
6-10 km	-0.46	0.63	61.40	66.54	9.6	56.00
10 km <	-0.00		71.63	1.91	3.5	72.02
Separate classes						
all separate	0.37	1.45	77.19	4.18	87.3	73.81
two or more combined in grade one to three	0.54	1.72	80.06	7.17	10.4	71.85
above third grade also two or more combined	0.00		70.01	1.59	2.2	64.07
Roof maintenance						
good	0.21	1.23	78.70	0.19	23.1	79.00
average	-0.06	0.94	72.70	1.18	43.1	72.52
poor	0.00		75.02	41.05	33.7	71.23
Centrality						
Lusaka, Ndola and Kitwe	-0.23	0.79	60.81	0.01	20.5	78.59
Rural Areas <50 km from Lusaka, Ndola and Kitwe	0.79	2.19	79.06	6.46	12.0	90.22
Other cities and towns along the Line of rail	0.66	1.93	76.89	0.01	7.5	86.66
Areas < 50 km fr. ot. cit. & towns along Line of rail	0.26	1.29	68.98	0.01	7.6	69.35
Towns, other provincial capitals	0.15	1.16	66.67	8.43	27.7	67.83
Areas < 50 km from other towns & province cap.	0.26	1.29	68.98	17.45	5.6	64.66
District centres	0.13	1.14	66.28	11.03	7.7	66.32
Other rural areas	0.00		63.29	37.40	11.3	63.46

Table 1. cont.

Variable	Coeffi-cient	Odds ratio	Esti-mated proba-bility, %	p-value in %	Esti-mated propor-tion %	Weighted atten-dance rate
Household level						
Square root of household expenditure / adult equiv.	0.013			0.01		
Less or equal to 2743=Q_1			65.54		25	64.09
Greater than 0.33 and less or equal to 9194= Q_2			80.02		50	77.90
Greater than 9194=Q_3	*		87.68		25	85.90
Number of rooms pr. person	0.91			0.01		
Less or equal to 0.33=Q_1			70.77		25	69.84
Greater than 0.33 and less or equal to 0.6= Q_2			78.58		50	79.50
Greater than 0.6=Q_3	*		83.37		25	83.48
Mother's education						
8 or more years	1.27	3.56	86.16	0.01	15.0	91.95
7 years	0.91	2.47	81.20	0.01	21.9	82.74
1-6 years	0.50	1.64	74.15	0.01	30.7	72.42
No education	0.00		63.62	0.01	32.4	59.94
Father's education						
8 or more years	0.65	1.91	76.87	0.01	28.5	88.13
7 years	0.33	1.39	70.75	0.01	17.4	73.83
1-6 years	0.13	1.14	66.49	0.03	27.1	68.85
No education	0.00		63.51	9.20	26.9	62.65
Number of dependants between 7 and 13 years	0.13			0.01		
1			73.12		19.2	71.97
2			76.50		33.6	77.47
3			79.31		30.7	80.10
4			81.14		12.5	82.63
5	*		82.05		4.0	77.95
Age of head of household						
50 -	0.82	2.27	82.04	0.01	33.9	70.83
30 - 49	0.59	1.80	78.37	0.01	61.5	75.79
- 29	0.00		66.81	0.01	4.6	63.31
Age of the child						
11-13	2.27	9.68	90.60	0.01	41.4	84.11
10	2.01	7.46	88.13	0.01	16.5	79.00
9	1.72	5.82	85.37	0.01	13.7	75.44
8	1.03	2.80	73.60	0.01	14.4	62.74
7	0.00		49.89	0.01	14.1	44.82

* Only one coefficient since the variable is used as a continuos variable in model (0).

4.2. Community level

Generally the community level variables have less degree of explanation than the household and child level variables.

When the school is situated in the community where the child live, the distance is coded as 0 km. For about fifty percent this is the case. Initially this variable was used as a continuos variable in the model. However, plotting the logit function of the mean for intervals indicated that it was far from being linear, and we did not succeed in finding a good transformation. Thus we preferred to use the classes given in the table. The estimates indicate that a distance to school between 5 and 10 kilometers gives the most negative influence. Distances which are less than six kilometers or more than ten kilometre do not seem to influence school attendance differently. Six kilometers is a fairly long distance to walk especially for the younger ones and transportation considered too expensive or not accessible. Only around two percent of the observations have school distance of more than 10 km. It is difficult to explain their higher attendance rate. A possible explanation might be that the need of transportation is more accepted and provided, but it is difficult to understand how parents could afford this.

The of teaching more than one grade in a classroom is thought to be clearly negative. This is however not the case in Zambia. Or rather, the marginal effect is negative, but when we apply this model and control for other factors, the effect is not only not significant at a one percent level, but is in fact positive for teaching more than one grade in a class at the three lowest grades. Only from grade four and above, teaching more than one grade in one classroom negatively affect attendance rate. Fortunately less than two percent of the children are offered this. One explanation might be that teaching more than one grade in one classroom is compensated by higher community involvement in small schools. Whatever the explanation is, this important finding is an opening for building smaller schools without jeopardizing attendance.

Maintenance of the roof may be considered as an indicator for how well the school is maintained. Not surprisingly roof maintenance seems to be positive correlated with school attendance rate. We would naturally expect quality in itself to affect school attendance, but it could also be the case that active involvement from community or parents both led to increased public resources/ or communal work on the school premises as well as to higher attendance rate. Around thirty percent of the schools are provided with a well maintained roof. And an average maintenance does not seem to be significantly differently from a poor one.

According to the estimates, geographical centrality gives rise to different rates of schooling, the coefficients vary from 0.79 to -0.23. The most positive impact on school attendance have centrality two and three, being rural areas less than fifty kilometres from Lusaka, Ndola and Kitwe, and from other cities along the Line of rail. The p-values indicate that centrality one - Ndola, Lusaka and Kitwe, and centrality four to seven do not differ significantly from centrality eight at a five percent level. Our interpretation is that low school attendance is both an «inner city», or rather «squatter inner city» problem, and a problem in less central and remote areas. The differences are especially large when looking at the marginal school attendance, but remains very high even when controlling for a number of other factors in our model.

The youngest school children in the distant areas often face reduced access to the school both because of long walking distances and lack of school places. Only around 25 per cent of the seven years old in centrality eight attend school, while the average weighted attendance rate for the seven years old is 44.82 per cent. This low attendance rate indicates that children living in distant areas often postpone their school entrance. The weighted attendance rate in per cent for the seven years olds for the centrality one to seven in percent are; 53, 73, 70, 44, 40, 33 and 45 respectively.

4.3. Household level

It is well documented that both capital and human resources might affect school attendance. Poor households may not afford to send their children to school, either because they cannot spare the expenses or because they need the children at home. We have used the household's total expenditure pr. adult equivalent as a measure for economic welfare, since this is considered both more valid and more reliable than the household's income. Taking the square root improved the fit. This indicate that the more well off a household is, the less a further increase in expenditure improve the attendance rate. The school attendance is much higher among the highest twenty-five percent household pr. adult equivalent expenditure than among the lowest twenty-five percent, 87.68 and 65.54 per cent respectively.

The number of rooms pr. person in the household is positively linked with school attendance. We assume this is both reflecting the asset value of the house and that overcrowding makes homework more difficult, passing exams less likely and school attendance less attractive. The estimated probability for the lower twenty-five percent , 70.77 per cent, is well below that for the upper twenty-five percent, 83.77 per cent.

As one would expect, mother's education has a strong positive effect on school attendance. The probability of school attendance will increase from 61.13 per cent for a child with a non-educated mother to 72.06 per cent for a child whose mother have between one and six years of education. Eight or more years of education increases the estimated probability to 84.85 per cent. The positive effect of the mother's education is widely acknowledged. The father's education also clearly influences the attendance rate positively. However, not as strongly as the mother's education. As we expected the marginal effects of either mother's or father's education decreases somewhat when controlling for other factors. This is likely to be due to the correlation between mother's and father's education and their somewhat reduced combined effect. For children whose both parents have no education we found that the weighted attendance rate was only 56.75 per cent.

The number of dependants is a continuos variable in the model. However, we still display the estimated probabilities, based on the model, and the estimated frequencies for dependants between one and five. (Observations above five were not included in the model due to the outlier analysis.) We had expected that the lower economic welfare (total consumption per adult equivalent) in households with many children 7 -13 years would reduce the probability of school attendance. Surprisingly, this is not the case, the number of children 7 - 13 years is positively correlated with schooling; the more dependants, the higher the probability for school attendance. An explanation may be lower opportunity costs, i.e., that household with more children are more likely to have covered their needs for help in the domestic work and farming. Another possibility is that households with many members are more likely to be well established.

The youngest parents are least likely to send their children to school. This may both be a cohort or age effect. Cohort effect means that the characteristic is linked with the generation and not with the age. If it is a cohort effect among young parents, it may indicate a rather alarming trend reflecting lower perceived opportunities after completing school. The model estimate indicates that there is a positive marginal effect for the head aged fifty and above. The marginal effect when the head is in the middle age group is less positive. However, this group has the highest attendance rate. This is probably because this group often has other characteristics which are positively linked with school attendance. The youngest parents often have younger children, and the younger children attend school less than the older. To control if this was the case, we examined the attendance rate for the seven years olds separately. Also for these children the attendance rate was lowest for the group with parents younger than thirty years.

4.4. Child level

Coefficients for the age factor vary from zero to 2.29. That is the biggest variation for the variables in the present model. The much lower attendance rate for the seven and eight years old expresses probably mainly the lack of places at school. When there are too few places the youngest children have to wait to be offered a place in school. But as (Silanda, 1988) states, an explanation is also that in sparsely populated areas children often have to walk long distances to school. Hence seven years old might be considered too young to walk to school and school start is postponed.

Note that we did not find any gender bias, i.e., there was no difference in how often boys and girls in this age group attended school.

4.5. Conclusion and policy recommendations

4.5.1. Conclusion

This multilevel analysis shows that school attendance is affected by a complex set of factors, at community level, at household level and to a lesser degree at child level.

The community level. Low school attendance is in many cases a community level problem. The following factors yield low school attendance for the whole community:

- Children living in remote villages (remote and «less central») and the poor big city areas (the squatter areas in the three biggest cities) attend school less frequently than children from other cities, towns, centres and villages located in the hinterland of cities and towns.
- If more than one of the first three grades have to share a teacher, that rather *increase* the probability of school attendance. We assume that is caused by smaller classes, but data does not allow us to test this.
- Whether the school is in, adjacent or close to the living area does not matter, unless the school is more than 6 kilometers away, in that case the attendance rate drops. As expected, younger children are more severely affected by remoteness and long walking distances than older ones.
- While we have not been able to measure that an active participation by the community increase the probability of school attendance, there are some results that indicate this being the case.

The household level. Within the communities, but across households there are a number of factors which are significant and strong. The effects of capital and human resources at household level are all as expected, significant and strong, as follows:

- Children from households with capital resources measured as economic welfare and as number of rooms in the home have a substantially higher probability of attending school.
- Children from households with human resources measured by mother's and father's education also have a substantially higher probability of attending school.

This strong effects of household level variables show the need to focus on both community and household level factors if the education sector policy aims at securing equity not only in opportunities, but also in outcome.

The child level. At child level the main findings are two:

- We did not find an expected gender bias, girls are as likely as boys to attend primary school.
- As expected the school attendance rates for the recommended age of schooling, seven years old children are quite low both for girls and boys.

Outside the scope of this paper, but documented in (CSO 1994) is that girls stop schooling before boys. At secondary level a substantially higher rate of boys continue. One reason is a larger burden of homework, but another reason is late school start has a different impact on girls and boys. When girls are reaching puberty, quite some will drop out, either due to pregnancy or risk of pregnancy. Hence the equally late school start for girls and boys is still in fact a gender problem.

4.5.2. Policy Recommendations

This study is not designed to provide *policy recommendations.* In order to do that, one would need to look into the economics of internal budget allocation as well as the budget share for the educational sector. But the findings from this study highlight certain issues that deserve to be considered for a redesign of an educational policy aiming at efficiency and equity in formal education. The main message of the study is that an educational sector policy needs to address access to primary schools at community level, at household level and at child level in order to achieve equity and efficiency.

Community level

The two geographical areas which are worse of than the others need special attention, as follows:

- For remote and less central villages and areas and efficient and relatively cheap policy recommendation is to build smaller and simpler schools for grade 1 to 3 closer to where the children are living.
- For inner city squatter areas, the study has revealed no special policy recommendation. In these areas it is a matter of more schools, classes and teachers.

Household level

It was well known that children with parents with more economic and/ or human resources than average, were more likely to attend school. This study tells us however also that this bias is not only a geographical one, even within the same communities, households resources is, may be not the destiny of children, but a strong trend. To break the spiral across generations, a real scholarship program is needed. Elements to consider are:

- School fee exemption;
- Public support to PTA (Parents Teacher Associations) fees and various allowance programs; and
- Public support to private tuition.

Combining community and household level targeting would increase both efficiency and equity and policies to consider are the following ones. It should be considered to limit public support for poor families to remote/ less central areas and inner city squatter areas in Lusaka, Ndola and Kitwe.

Child level

At child level, the policy should address both the factors behind and the impact of late schooling start, as follows:

- As addressed at the community level, it should be considered to build small and simple schools for grade 1-3 in remote and less central areas, close to where children are living.
- To avoid a future gender bias, girls should be given preference for a early start of schooling when the capacity is limited such as in the squatter areas.

References

Alwang, J. and Siegel P.B. (1994): Rural Poverty in Zambia: An Analysis of Causes and Policy Recommendations, World Bank report, Draft. Washington DC: World Bank.

Anderson, D. A. and Aitkin, M. (1985): Variance component models with binary response: interviewer variability, *Journal of Royal Statistical Society B* **47**, 203-210.

Behrman J. (1993): Analyzing human resource effects: Education: in Demery, Ferroni, and Grootaert (1993): *Understanding the Social Effects of Policy Reform,* Washington DC: The World Bank, 114-136.

Booth, D. et al. (1994): Coping with Cost Recovery, Stockholm: Stockholm University.

Central Statistical Office (1994): Priority Survey II, 1993, Tabulation Report. Lusaka.

Central Statistical Office (forthcoming): Community Survey, Tabulation Report. Lusaka

Demery L., Ferroni M. and Grootaert C. (eds) (1993): *Understanding the Social Effects of Policy Reform,* Washington DC: The World Bank.

Goldstein, H.I. (1986): Multilevel mixed linear model analysis using iterative generalized least squares, *Biometrika* **73**, 43-56.

Goldstein, H.I. (1991): Non-linear multilevel models, with an application to discrete response data, *Biometrika* **78**, 45-51.

Kelly, J.M. (1991): *Education in a declining economy, the case of Zambia 1975-1985*, World Bank number 8. Washington DC: The World Bank.

Moser, C.O.N. (1996): *Confronting Crisis, A Comparative Study of Household Responses to Poverty and Vulnerability in Four Poor Urban Communities,* ESD No 8. Washington DC: The World Bank.

Rodriguez, G. and Goldman, N. (1995): An assessment of estimation Procedures for Multilevel Models with Binary Responses, *Journal of Royal Statistical Society A* **158**, 73-89.

Silanda, E.M. (1988): Education and regional differences, A study of the nature and educational disparities in Zambia, Institute of international education, Stockholm: University of Stockholm.

Wold B. K. et al. (1996): The Sub-Sahara African Women Farmer under Structural Adjustment - Zambia. Supply response in a Gender-Perspective - the case of Zambia. Report (preliminary edition). Oslo, Ås, Lusaka, Chilanga: Statistics Norway, Noragric, Central Statistical Office, Farming Systems Association of Zambia.

World Bank (1993): Zambia, Prospects for Sustainable and Equitable Growth, Report No. 11570-ZA. Washington DC.

Zeger, S. L. and Karim, R.M. (1991): Generalized linear models with random effects: a Gibbs sampler approach, *Journal of American Statistical Association* **86**, 79-86.

Recent publications in the series Discussion Papers

45 O. Bjerkholt, E. Gjelsvik and Ø. Olsen (1989): Gas Trade and Demand in Northwest Europe: Regulation, Bargaining and Competition.

46 L.S. Stambøl and K.Ø. Sørensen (1989): Migration Analysis and Regional Population Projections.

47 V. Christiansen (1990): A Note on the Short Run Versus Long Run Welfare Gain from a Tax Reform.

48 S. Glomsrød, H. Vennemo and T. Johnsen (1990): Stabilization of Emissions of CO_2: A Computable General Equilibrium Assessment.

49 J. Aasness (1990): Properties of Demand Functions for Linear Consumption Aggregates.

50 J.G. de Leon (1990): Empirical EDA Models to Fit and Project Time Series of Age-Specific Mortality Rates.

51 J.G. de Leon (1990): Recent Developments in Parity Progression Intensities in Norway. An Analysis Based on Population Register Data

52 R. Aaberge and T. Wennemo (1990): Non-Stationary Inflow and Duration of Unemployment

53 R. Aaberge, J.K. Dagsvik and S. Strøm (1990): Labor Supply, Income Distribution and Excess Burden of Personal Income Taxation in Sweden

54 R. Aaberge, J.K. Dagsvik and S. Strøm (1990): Labor Supply, Income Distribution and Excess Burden of Personal Income Taxation in Norway

55 H. Vennemo (1990): Optimal Taxation in Applied General Equilibrium Models Adopting the Armington Assumption

56 N.M. Stølen (1990): Is there a NAIRU in Norway?

57 Å. Cappelen (1991): Macroeconomic Modelling: The Norwegian Experience

58 J.K. Dagsvik and R. Aaberge (1991): Household Production, Consumption and Time Allocation in Peru

59 R. Aaberge and J.K. Dagsvik (1991): Inequality in Distribution of Hours of Work and Consumption in Peru

60 T.J. Klette (1991): On the Importance of R&D and Ownership for Productivity Growth. Evidence from Norwegian Micro-Data 1976-85

61 K.H. Alfsen (1991): Use of Macroeconomic Models in Analysis of Environmental Problems in Norway and Consequences for Environmental Statistics

62 H. Vennemo (1991): An Applied General Equilibrium Assessment of the Marginal Cost of Public Funds in Norway

63 H. Vennemo (1991): The Marginal Cost of Public Funds: A Comment on the Literature

64 A. Brendemoen and H. Vennemo (1991): A climate convention and the Norwegian economy: A CGE assessment

65 K.A. Brekke (1991): Net National Product as a Welfare Indicator

66 E. Bowitz and E. Storm (1991): Will Restrictive Demand Policy Improve Public Sector Balance?

67 Å. Cappelen (1991): MODAG. A Medium Term Macroeconomic Model of the Norwegian Economy

68 B. Bye (1992): Modelling Consumers' Energy Demand

69 K.H. Alfsen, A. Brendemoen and S. Glomsrød (1992): Benefits of Climate Policies: Some Tentative Calculations

70 R. Aaberge, Xiaojie Chen, Jing Li and Xuezeng Li (1992): The Structure of Economic Inequality among Households Living in Urban Sichuan and Liaoning, 1990

71 K.H. Alfsen, K.A. Brekke, F. Brunvoll, H. Lurås, K. Nyborg and H.W. Sæbø (1992): Environmental Indicators

72 B. Bye and E. Holmøy (1992): Dynamic Equilibrium Adjustments to a Terms of Trade Disturbance

73 O. Aukrust (1992): The Scandinavian Contribution to National Accounting

74 J. Aasness, E. Eide and T. Skjerpen (1992): A Criminometric Study Using Panel Data and Latent Variables

75 R. Aaberge and Xuezeng Li (1992): The Trend in Income Inequality in Urban Sichuan and Liaoning, 1986-1990

76 J.K. Dagsvik and S. Strøm (1992): Labor Supply with Non-convex Budget Sets, Hours Restriction and Non-pecuniary Job-attributes

77 J.K. Dagsvik (1992): Intertemporal Discrete Choice, Random Tastes and Functional Form

78 H. Vennemo (1993): Tax Reforms when Utility is Composed of Additive Functions

79 J.K. Dagsvik (1993): Discrete and Continuous Choice, Max-stable Processes and Independence from Irrelevant Attributes

80 J.K. Dagsvik (1993): How Large is the Class of Generalized Extreme Value Random Utility Models?

81 H. Birkelund, E. Gjelsvik, M. Aaserud (1993): Carbon/energy Taxes and the Energy Market in Western Europe

82 E. Bowitz (1993): Unemployment and the Growth in the Number of Recipients of Disability Benefits in Norway

83 L. Andreassen (1993): Theoretical and Econometric Modeling of Disequilibrium

84 K.A. Brekke (1993): Do Cost-Benefit Analyses favour Environmentalists?

85 L. Andreassen (1993): Demographic Forecasting with a Dynamic Stochastic Microsimulation Model

86 G.B. Asheim and K.A. Brekke (1993): Sustainability when Resource Management has Stochastic Consequences

87 O. Bjerkholt and Yu Zhu (1993): Living Conditions of Urban Chinese Households around 1990

88 R. Aaberge (1993): Theoretical Foundations of Lorenz Curve Orderings

89 J. Aasness, E. Biørn and T. Skjerpen (1993): Engel Functions, Panel Data, and Latent Variables - with Detailed Results

90 I. Svendsen (1993): Testing the Rational Expectations Hypothesis Using Norwegian Microeconomic Data Testing the REH. Using Norwegian Microeconomic Data

91 E. Bowitz, A. Rødseth and E. Storm (1993): Fiscal Expansion, the Budget Deficit and the Economy: Norway 1988-91

92 R. Aaberge, U. Colombino and S. Strøm (1993): Labor Supply in Italy

93 T.J. Klette (1993): Is Price Equal to Marginal Costs? An Integrated Study of Price-Cost Margins and Scale Economies among Norwegian Manufacturing Establishments 1975-90

94 J.K. Dagsvik (1993): Choice Probabilities and Equilibrium Conditions in a Matching Market with Flexible Contracts

95 T. Kornstad (1993): Empirical Approaches for Analysing Consumption and Labour Supply in a Life Cycle Perspective

96 T. Kornstad (1993): An Empirical Life Cycle Model of Savings, Labour Supply and Consumption without Intertemporal Separability

97 S. Kverndokk (1993): Coalitions and Side Payments in International CO_2 Treaties

98 T. Eika (1993): Wage Equations in Macro Models. Phillips Curve versus Error Correction Model Determination of Wages in Large-Scale UK Macro Models

99 A. Brendemoen and H. Vennemo (1993): The Marginal Cost of Funds in the Presence of External Effects

100 K.-G. Lindquist (1993): Empirical Modelling of Norwegian Exports: A Disaggregated Approach

101 A.S. Jore, T. Skjerpen and A. Rygh Swensen (1993): Testing for Purchasing Power Parity and Interest Rate Parities on Norwegian Data

102 R. Nesbakken and S. Strøm (1993): The Choice of Space Heating System and Energy Consumption in Norwegian Households (Will be issued later)

103 A. Aaheim and K. Nyborg (1993): "Green National Product": Good Intentions, Poor Device?

104 K.H. Alfsen, H. Birkelund and M. Aaserud (1993): Secondary benefits of the EC Carbon/ Energy Tax

105 J. Aasness and B. Holtsmark (1993): Consumer Demand in a General Equilibrium Model for Environmental Analysis

106 K.-G. Lindquist (1993): The Existence of Factor Substitution in the Primary Aluminium Industry: A Multivariate Error Correction Approach on Norwegian Panel Data

107 S. Kverndokk (1994): Depletion of Fossil Fuels and the Impacts of Global Warming

108 K.A. Magnussen (1994): Precautionary Saving and Old-Age Pensions

109 F. Johansen (1994): Investment and Financial Constraints: An Empirical Analysis of Norwegian Firms

110 K.A. Brekke and P. Børing (1994): The Volatility of Oil Wealth under Uncertainty about Parameter Values

111 M.J. Simpson (1994): Foreign Control and Norwegian Manufacturing Performance

112 Y. Willassen and T.J. Klette (1994): Correlated Measurement Errors, Bound on Parameters, and a Model of Producer Behavior

113 D. Wetterwald (1994): Car ownership and private car use. A microeconometric analysis based on Norwegian data

114 K.E. Rosendahl (1994): Does Improved Environmental Policy Enhance Economic Growth? Endogenous Growth Theory Applied to Developing Countries

115 L. Andreassen, D. Fredriksen and O. Ljones (1994): The Future Burden of Public Pension Benefits. A Microsimulation Study

116 A. Brendemoen (1994): Car Ownership Decisions in Norwegian Households.

117 A. Langørgen (1994): A Macromodel of Local Government Spending Behaviour in Norway

118 K.A. Brekke (1994): Utilitarism, Equivalence Scales and Logarithmic Utility

119 K.A. Brekke, H. Lurås and K. Nyborg (1994): Sufficient Welfare Indicators: Allowing Disagreement in Evaluations of Social Welfare

120 T.J. Klette (1994): R&D, Scope Economies and Company Structure: A "Not-so-Fixed Effect" Model of Plant Performance

121 Y. Willassen (1994): A Generalization of Hall's Specification of the Consumption function

122 E. Holmøy, T. Hægeland and Ø. Olsen (1994): Effective Rates of Assistance for Norwegian Industries

123 K. Mohn (1994): On Equity and Public Pricing in Developing Countries

124 J. Aasness, E. Eide and T. Skjerpen (1994): Criminometrics, Latent Variables, Panel Data, and Different Types of Crime

125 E. Biørn and T.J. Klette (1994): Errors in Variables and Panel Data: The Labour Demand Response to Permanent Changes in Output

126 I. Svendsen (1994): Do Norwegian Firms Form Extrapolative Expectations?

127 T.J. Klette and Z. Griliches (1994): The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous

128 K.E. Rosendahl (1994): Carbon Taxes and the Petroleum Wealth

129 S. Johansen and A. Rygh Swensen (1994): Testing Rational Expectations in Vector Autoregressive Models

130 T.J. Klette (1994): Estimating Price-Cost Margins and Scale Economies from a Panel of Microdata

131 L. A. Grünfeld (1994): Monetary Aspects of Business Cycles in Norway: An Exploratory Study Based on Historical Data

132 K.-G. Lindquist (1994): Testing for Market Power in the Norwegian Primary Aluminium Industry

133 T. J. Klette (1994): R&D, Spillovers and Performance among Heterogenous Firms. An Empirical Study Using Microdata

134 K.A. Brekke and H.A. Gravningsmyhr (1994): Adjusting NNP for instrumental or defensive expenditures. An analytical approach

135 T.O. Thoresen (1995): Distributional and Behavioural Effects of Child Care Subsidies

136 T. J. Klette and A. Mathiassen (1995): Job Creation, Job Destruction and Plant Turnover in Norwegian Manufacturing

137 K. Nyborg (1995): Project Evaluations and Decision Processes

138 L. Andreassen (1995): A Framework for Estimating Disequilibrium Models with Many Markets

139 L. Andreassen (1995): Aggregation when Markets do not Clear

140 T. Skjerpen (1995): Is there a Business Cycle Component in Norwegian Macroeconomic Quarterly Time Series?

141 J.K. Dagsvik (1995): Probabilistic Choice Models for Uncertain Outcomes

142 M. Rønsen (1995): Maternal employment in Norway, A Parity-Specific Analysis of the Return to Full-Time and Part-Time Work after Birth

143 A. Bruvoll, S. Glomsrød and H. Vennemo (1995): The Environmental Drag on Long- Term Economic Performance: Evidence from Norway

144 T. Bye and T. A. Johnsen (1995): Prospects for a Common, Deregulated Nordic Electricity Market

145 B. Bye (1995): A Dynamic Equilibrium Analysis of a Carbon Tax

146 T. O. Thoresen (1995): The Distributional Impact of the Norwegian Tax Reform Measured by Disproportionality

147 E. Holmøy and T. Hægeland (1995): Effective Rates of Assistance for Norwegian Industries

148 J. Aasness, T. Bye and H.T. Mysen (1995): Welfare Effects of Emission Taxes in Norway

149 J. Aasness, E. Biørn and Terje Skjerpen (1995): Distribution of Preferences and Measurement Errors in a Disaggregated Expenditure System

150 E. Bowitz, T. Fæhn, L. A. Grünfeld and K. Moum (1995): Transitory Adjustment Costs and Long Term Welfare Effects of an EU-membership – The Norwegian Case

151 I. Svendsen (1995): Dynamic Modelling of Domestic Prices with Time-varying Elasticities and Rational Expectations

152 I. Svendsen (1995): Forward- and Backward Looking Models for Norwegian Export Prices

153 A. Langørgen (1995): On the Simultaneous Determination of Current Expenditure, Real Capital, Fee Income, and Public Debt in Norwegian Local Government

154 A. Katz and T. Bye(1995): Returns to Publicly Owned Transport Infrastructure Investment. A Cost Function/ Cost Share Approach for Norway, 1971-1991

155 K. O. Aarbu (1995): Some Issues About the Norwegian Capital Income Imputation Model

156 P. Boug, K. A. Mork and T. Tjemsland (1995): Financial Deregulation and Consumer Behavior: the Norwegian Experience

157 B.E. Naug and R. Nymoen (1995): Import Price Formation and Pricing to Market: A Test on Norwegian Data

158 R. Aaberge (1995): Choosing Measures of Inequality for Empirical Applications

159 T.J. Klette and S.E. Førre (1995): Innovation and Job Creation in a Small Open Economy: Evidence from Norwegian Manufacturing Plants 1982-92

160 S. Holden, D. Kolsrud and B. Vikøren (1995): Noisy Signals in Target Zone Regimes: Theory and Monte Carlo Experiments

161 T. Hægeland (1996): Monopolistic Competition, Resource Allocation and the Effects of Industrial Policy

162 S. Grepperud (1996): Poverty, Land Degradation and Climatic Uncertainty

163 S. Grepperud (1996): Soil Conservation as an Investment in Land

164 K.A. Brekke, V. Iversen and J. Aune (1996): Soil Wealth in Tanzania

165 J.K. Dagsvik, D.G. Wetterwald and R. Aaberge (1996): Potential Demand for Alternative Fuel Vehicles

166 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part I: Theory

167 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part II: Inference

168 R. Aaberge, A. Björklund, M. Jäntti, M. Palme, P. J. Pedersen, N. Smith and T. Wennemo (1996): Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States

169 K. Nyborg (1996): Some Norwegian Politicians' Use of Cost-Benefit Analysis

170 E. Berg, S. Kverndokk and K. E. Rosendahl (1996): Market Power, International CO_2 Taxation and Petroleum Wealth

171 R. Aaberge, U. Colombino and S. Strøm (1996): Welfare Effects of Proportional Taxation: Empirical Evidence from Italy, Norway and Sweden

172 J.K. Dagsvik (1996): Dynamic Choice, Multistate Duration Models and Stochastic Structure

173 J.K. Dagsvik (1996): Aggregation in Matching Markets

174 H.C. Bjørnland (1996): The Dynamic Effects of Aggregate Demand, Supply and Oil Price Shocks

175 A. Bruvoll and K. Ibenholt (1996): Future Waste Generation. Forecasts Based on a Macroeconomic Model

176 T. Fæhn and L. A. Grünfeld (1996) Recent Leaps Towards Free Trade. The Impact on Norwegian Industry and Trade Patterns

177 R. Barrell and K. A. Magnussen (1996): Counterfactual Analyses of Oil price Shocks using a World Model

178 E. Bowitz and S. I. Hove (1996): Business cycles and fiscal policy: Norway 1973-93

179 H.C. Bjørnland (1996): Sources of Business Cycles in Energy Producing Economies: The case of Norway and United Kingdom

180 K. Nyborg (1996): The Political Man and Contingent Valuation: Motives Do Count

181 E. Berg, S. Kverndokk and K.E. Rosendahl (1996): Gains from Cartelisation in the Oil Market

182 R. Aaberge and I. Aslaksen (1996): Decomposition of the Gini Coefficient by Income Components: Various Types of Applications and Interpretations

183 B. Bye (1996): Taxation, Unemployment and Growth: Dynamic Welfare Effects of "Green" Policies

184 T.J. Klette and F. Johansen (1996): Accumulation of R&D Capital and Dynamic Firm Performance: A Not-so-fixed Effect Model

185 B. Bye (1996): Environmental Tax Reform and Producer Foresight: An Intertemporal Computable General Equilibrium Analysis

186 S. Grepperud (1997): Soil Depletion Choices under Production and Price Uncertainty

187 N.M. Stølen and T. Åvitsland (1997): Has Growth in Supply of Educated Persons Been Important for the Composition of Employment?

188 T.J. Klette and Z. Griliches (1997): Empirical Patterns of Firm Growth and R&D Investment: a Quality Ladder Model Interpretation

189 J. Aune, S. Glomsrød, V. Iversen and H. Wiig (1997): Structural Adjustment and Soil Degradation in Tanzania. A CGE-model Approach with Endogenous Soil Productivity

190 E. Biørn and T.J. Klette (1997): Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation

191 L. Belsby and B.K. Wold (1997): Primary Schooling in Zambia Squeezed at Community and Household Level

Discussion Papers

B

Returadresse:
Statistisk sentralbyrå
Postboks 8131 Dep.
N-0033 Oslo

Statistics Norway
Research Department
P.O.B. 8131 Dep.
N-0033 Oslo

Tel.: + 47 - 22 86 45 00
Fax: + 47 - 22 11 12 38

ISSN 0803-074X

Statistics Norway
Research Department