

## **HIV/AIDS and Primary School Performance in Tanzania**

**Peter Wobst**, Trade and Macroeconomics Division, IFPRI; c/o Center for Development Research (ZEFb), Walter-Flex-Str. 3, 53113 Bonn, GERMANY; phone +49 (228) 73-1873; fax +49 (228) 73-1869; email [p.wobst@cgiar.org](mailto:p.wobst@cgiar.org) (correspondence address)

and

**Channing Arndt**, Department of Agricultural Economics, Purdue University W. Lafayette, IN, USA; phone +1 (765) 494-5837; email [carndt@purdue.edu](mailto:carndt@purdue.edu)

Contributed paper selected for presentation at the 25th International Conference of Agricultural Economists, August 16-22, 2003, Durban, South Africa

Copyright 2003 by P. Wobst and C. Arndt. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

Abstract: We examine the performance of the primary school education system in Tanzania over the 1990s—a decade characterized by substantial AIDS deaths. Given the relatively robust correlation between educational attainment and productivity established in the literature in both agricultural and non-agricultural sectors, human capital accumulation through education forms a major component of development strategy. At the same time, AIDS poses clear threats to the goal of human capital accumulation through education. To assess performance of the primary school system, we estimate non-stationary education transition matrices using a minimum cross entropy approach at the national, sub-national, and regional levels for girls, boys, and all students. Results indicate a deterioration in primary school performance using enrollments in grade 7, the final year of primary school, as a metric. This deterioration in performance occurred despite increased real resource allocations to the public education system and positive, if only tepid, overall economic growth trends. We conclude that the HIV/AIDS pandemic has quite likely slowed human capital accumulation in Tanzania.

Keywords: AIDS, human capital, cross entropy estimation, Tanzania

Acknowledgements: We acknowledge funding for this research by the US Department of Labor and would like to thank Yuki Yanoma for valuable research assistance.

## **1. Introduction**

We examine the performance of the primary school education system in Tanzania over the 1990s—a decade characterized by substantial AIDS deaths. The literature now points with reasonable confidence to relatively robust links between educational attainment and productivity in both agricultural and non-agricultural sectors. Given these links, economic performance in both agricultural and non-agricultural sectors is likely to suffer if the HIV/AIDS pandemic causes reduced rates of human capital accumulation (Arndt, 2002). In order to examine the performance of primary schools in Tanzania, we estimate education transition matrices using a minimum cross entropy estimation approach. Results indicate that performance of the primary school system has deteriorated over the 1990s. This deterioration in performance occurred despite increased real resource allocations to the public education system and positive, if only tepid, overall economic growth trends. We conclude that the HIV/AIDS pandemic has quite likely slowed human capital accumulation in Tanzania.

## **2. Education and Productivity**

Evidence continues to mount for a strong positive association between education levels and productivity growth in both agricultural and non-agricultural sectors in developing countries. An early review of 18 studies focused on agricultural productivity conducted by Lockheed, Jamison, and Lau (1980) concluded that education had a significant and positive impact on agricultural productivity. However, the authors qualified their results by stating that the effect of education was likely to be much stronger in “modernizing agricultural environments than in traditional ones” (p. 61). More than two decades later, this qualification still excludes much of Africa including Tanzania.

However, a series of recent microeconomic analyses tends to find a strong association between education and agricultural productivity even in traditional African contexts (including Tanzania). Positive returns were found by Weir and Knight (2000a and 2000b) for Ethiopia, Pinckney (1997) for Kenya and Tanzania, and Appleton and Balihuta (1996) for Uganda.

Furthermore, these studies point to significant education externalities. Put simply, the illiterate neighbors of more educated farmers benefit by simply watching and copying.

In non-agricultural sectors, Psacharopoulos (1994) provides a review of microeconomic studies. He finds positive returns to education with particularly high returns to primary schooling. Bloom, Canning, and Sevilla (2001) employ macroeconomic panel data to examine the issues of health, human capital and economic growth. Consistent with the microeconomic evidence, they find a positive association between schooling and aggregate economic output. Generally, the weight of evidence has placed significant investments in human capital as a centerpiece of a successful development strategy.

### **3. HIV/AIDS and Education Trends**

#### **3.1 General**

Tanzania has a mature HIV/AIDS pandemic of medium severity relative to other countries in the southern and eastern Africa region. Due to the maturity of the pandemic, sufficient time has passed for significant numbers of HIV positive individuals to develop AIDS. As a result, AIDS effects, not just high prevalence rates, can be observed. Also, the overall HIV prevalence rate, at about 8% of the adult population, is sufficiently high to warrant real concern about economic impacts. Moreover, effects on the education sector may be larger than the overall adult prevalence rate suggests. There is little reason to believe that HIV prevalence rates amongst teachers and other critical personnel in the education sector are any lower than the average prevalence rates in the adult population and considerable reason to fear that rates might be higher (Badcock-Walters and Whiteside, 2000).

Even in the absence of the pandemic, the dearth of qualified teachers, particularly teachers with experience, would be a real concern. AIDS has quite likely worsened the teacher supply constraint. Graph 1 shows that the primary school pupil/teacher ratio has deteriorated throughout the 1990s. In the late 1990s, the absolute number of primary school teachers declined. Due to the

roughly nine year lag between infection and death, these deaths could correspond with infection in the late 1980s and early 1990s—a time period during which HIV was rapidly becoming more generalized amongst the population according to the available evidence.

Concerns also exist on the demand side. Akabayashi and Psacharopoulos (1999) explore the trade-offs between child labor and human capital formation for Tanzania. Consistent with expectations, they find that hours of work by children are negatively correlated with reading and mathematical skills. As parents stricken with AIDS fall ill and die, the need for children to work clearly becomes more pressing. In addition, with the large volume of orphaning that is occurring, one can expect large numbers of children to be living with friends and relatives. An analysis of data from the Kagera region of Tanzania indicates that orphaning reduces the demand for education, particularly when the effected families are poor (World Bank, 1997).

While these concerns have been voiced for some time, we are not aware of any published analysis of primary education system trends over a sustained period characterized by AIDS deaths. The following analysis seeks to fill this gap.

### *3.2 Data and Methods*

Estimation focuses on school enrollment. Data was collected on primary school enrollment, by region, from 1991-2000 for boys and girls (MoE, various years). Primary school contains grades 1 to 7 (g1 to g7). We are particularly interested in the way that school children move through the educational system. For a child enrolled at a given education level, three possibilities for the following year are assumed with respect to enrollment:

- a) The child progresses to the next educational level,
- b) The child repeats the same educational level (far from uncommon), or
- c) The child exits the educational system (and presumably enters the labor force).

By assumption, children who exit the school system will not re-enroll at a later date.

The probabilities associated with each of these outcomes for each grade level are of interest. In order to estimate these probabilities, we postulate a transition matrix,  $T$ , giving probabilities associated with each possible outcome from period  $t$  to period  $t+1$ . All blank elements in the matrix are assumed to have value zero.

$$T = \begin{bmatrix} T_{g1g1} & T_{g1g2} & & & & & & T_{g1exit} \\ & T_{g2g2} & T_{g2g3} & & & & & T_{g2exit} \\ & & T_{g3g3} & T_{g3g4} & & & & T_{g3exit} \\ & & & T_{g4g4} & T_{g4g5} & & & T_{g4exit} \\ & & & & T_{g5g5} & T_{g5g6} & & T_{g5exit} \\ & & & & & T_{g6g6} & T_{g6g7} & T_{g6exit} \\ & & & & & & T_{g7g7} & T_{g7exit} \end{bmatrix}$$

The rows of the transition matrix  $T$  sum to one. This implies that the entire enrollment at each scholastic level must be accounted for. Consider row 2, which corresponds to children enrolled in grade 2. According to row 2, children at this scholastic level could repeat the same level next year ( $T_{g2g2}$ ), enroll in grade 3 ( $T_{g2g3}$ ), or exit the school system with no probability of return ( $T_{g2exit}$ ). By assumption, children enrolled in grade 2 cannot leap to grade 4 or higher and will not fall back to grade 1. The remaining rows can be similarly interpreted.

Once the transition matrix has been estimated, the evolution of enrollments proceeds as follows. Let  $E_t$  be a seven element column vector (seven grades in primary school) with enrollment in grade 1 in year  $t$  in the first element and zeros elsewhere. The number of individuals at each scholastic level in period  $t+1$  is then  $S_{t+1} = T'S_t + E_t$ . Note that, since enrollments in grade 1 (rather than new entrants to grade 1) are supplied exogenously, the transition probability  $T_{g1g1}$  must be set equal to zero.

T matrices were estimated for each region, for the nation as a whole, and for various groupings of regions. Given the relatively short time series, a minimum cross entropy approach, similar to the approach of Karantininis (2002), was employed. The approach relies upon information theory and is well-suited to problems where data series are short and/or incomplete (Golan, Judge, and Miller, 1996). The philosophy of the minimum cross entropy approach is to use all available information, but not assume any further information. Consistent with this philosophy, the minimum cross entropy approach also permits use of prior information on transition probabilities. While the approach permits hypothesis testing (see, for example, Golan and Vogel 2000 or Karantininis 2002), the power of these tests is weak and/or their small sample properties are unknown. The accent in the minimum cross entropy approach is placed on robust parameter estimation.

Following Karantininis (2002), non-stationary transition matrices (transition probabilities are permitted to evolve through time) were estimated. Individual transition probabilities were permitted a linear trend through time. For example, a given element of the matrix T may start at a high (low) value in 1991 and then gradually decline (increase), in equal increments per year, towards a lower (higher) value by the year 2000. The implied transition matrices for each year are constrained to meet all the requirements of a transition matrix (e.g., rows sum to one and all elements fall within the [0,1] interval).

### 3.3 Results

The fit of projected enrollments to the data is relatively good. Pseudo r-squared measures are frequently above 0.9. The estimated non-stationary transition matrices (not shown due to space considerations) exhibit a marked tendency for reduced repetition rates through time and a consistent tendency for the probability of moving to the next level to rise. These trends are present at the national level for both girls and boys at almost every scholastic level. The reductions in repeat probabilities and increases in higher level transition probabilities imply that students are moving more quickly through the system. Since the row sums of the transition matrix must be equal to one,

the difference between the change in the repetition rate and the change in the next level transition probability determines the change in the exit probability. This change in the exit probability determines whether students move more quickly to higher levels or out of the educational system.

Whether the trend of up (desirable) or out (undesirable) predominates is not clear from simple examination of the estimated non-stationary transition matrix. For example, at the national level, exit probabilities are declining for grades 2 and 3 making trends at these grade levels uniformly positive. However, exit probabilities are increasing with time, sometimes very substantially, for grades 4 to 7. Given that the early primary exit probabilities are decreasing and that probabilities of transitioning to higher grade levels are increasing, the exact implications for enrollments at, for example, grade 7 are not entirely clear from simple inspection of the estimated matrices.

Table 1 illustrates the net effect of these trends in transition probabilities. The Table shows *ratios* of projected enrollments for girls, boys, and all students at national, regional, and sub-regional levels in the year 2000 using alternative transition matrices. In the numerator, enrollments in 2000 are projected assuming that the 1991 estimated transition matrix had prevailed over the entire period. In the denominator, enrollments in 2000 are projected assuming that the 2000 estimated transition matrix had prevailed over the entire period. Actual grade 1 enrollments are used for both projections. If the ratio is greater than one, then the transition probabilities estimated for the beginning of the decade resulted in higher enrollments in grade 7 than the transition probabilities estimated for the end of the decade.

Transition probability trends are revealed to be strongly negative for primary school using grade 7 enrollments as a metric. At the national level, enrollments would have been 12% higher had the transition probabilities for 1991 prevailed throughout the decade. Further analysis of population and labor force data by Arndt and Wobst (2002) support the conclusion of declining educational performance and tend to point in the direction of HIV/AIDS as a likely cause. The number of children enrolled in primary school as a share of the population aged 5-14 declined from 50% in

1991 to 46% in 2000. At the same time, comparison of the 1990/91 and 2000/01 labor force indicates that the share of 10-14 years olds in the labor force increased dramatically over the decade passing from 6% to 13%. The surveys also point to high death rates among adults aged 25-40 over the same period (see Arndt and Wobst for complete comparisons and caveats).

These results at the national level support concerns about the implications of HIV/AIDS for the education sector. While strongly indicative, the exact role of AIDS in these national trends is unclear. Education transition probabilities could evolve positively or negatively for lots of reasons with AIDS being only one. Unfortunately, untangling the exact implications of AIDS is impossible. The true counter-factual, a parallel world without the AIDS pandemic, will never be observed.

In an effort to gain some comparative traction, transition matrices for primary school enrollments were also estimated at the regional level. The primary purpose of these estimations was to attempt to take advantage of variations in prevalence rates across the 20 administrative regions comprising Tanzania. Evidence for variation in prevalence rates across regions during the early 1990s exists. For example, the prevalence rate in 1992 among ante-natal women in Kilimanjaro (from a rural observation post) was 6.4% while the observed rate from a rural observation post in Rukwa was 11.3%. Published data on HIV infection among blood donors also show strong regional variation (MoH 2000 and 2001).

Based on these data, five regions were labeled as highly afflicted in the early 1990s. If highly afflicted by the early 1990s, significant AIDS deaths should have occurred before 2000—the final year of the estimation period. The regions labeled as highly afflicted were Dar Es Salaam, Iringa, Kagera, Mbeya, and Rukwa. Enrollments in these regions were aggregated to form a group labeled HIGHAFF. All other regions were aggregated into a group labeled LESSAFF. In addition, a final group, labeled HIGHXM, comprised of Dar Es Salaam, Kagera, and Rukwa was created.

The sub-regional and regional estimates reveal that the negative trends in transition probabilities for primary school appear to be almost completely generalized across regions. Nearly

all of the ratios are greater than one indicating that the 1991 transition probability matrix is preferred in terms of enrollments in grade 7. However, the attempt to gain insight from the variation in prevalence rates across regions gives counter-intuitive results. As indicated in Table 1, the highly afflicted group (HIGHAFF) actually performs better than the less afflicted group (LESSAFF).

Closer examination of the regional results indicates that Iringa and Mbeya, both members of the highly afflicted group, are performing relatively well (members of the highly afflicted group are *italicized* in the Table). Since prevalence rates in Iringa were measured in urban zones and at the roadside, it is possible that the published prevalence rates substantially overstate the actual prevalence rate within the region. This is also possible for Mbeya though the evidence for a fairly generalized pandemic in that region by 1991 is far more complete. With Iringa and Mbeya dropped from the highly afflicted group as in HIGHXM, the results deteriorate dramatically conforming to the expectation of more severe effects in highly afflicted regions. However, it is always the case that the average becomes smaller if positive innovations are removed.

There are a series of potential explanations for the failure to pick up the influence of HIV/AIDS across regions. These include:

- 1) The implications of HIV/AIDS for the education system are, in fact, not that strong. Other factors caused the deteriorating performance of primary schools over the 1990s.
- 2) The regional groupings into highly afflicted and less afflicted are inaccurate. A more accurate grouping would indicate a stronger influence of HIV/AIDS.
- 3) Trends in primary school transition probabilities may have been quite different at the beginning of the estimation period. One must look at the change in the trend rather than the trend itself while controlling for other factors that might influence trends.
- 4) Migration across regions biases the regional matrices making performance in regions with out-migration appear worse (students exit due to out-migration) and performance in regions with in-migration appear better (students appear due to in-migration). Since migration is predominantly

rural to urban and urban areas tended to have higher prevalence rates especially in the early 1990s, the migration effect would tend to offset AIDS effects.

- 5) Supply side effects, driven primarily by the aggregate supply of teachers at the national level, are much stronger than demand side effects. Educational attainment in regions with low prevalence rates is negatively affected due to the paucity of teachers, particularly experienced teachers, nationwide. Also, prevalence rates amongst teachers might be only weakly correlated with the overall prevalence rate in the region where they are located.

#### **4. Conclusions and Suggestions for Future Research**

The failure to find a strong influence of HIV/AIDS in cross section does not provide substantial enough support to alleviate concerns about the implications of HIV/AIDS for educational attainment and human capital accumulation. The hypothesis is that HIV/AIDS tends to negatively affect trends in the education system. In the results obtained, there are no regions that exhibit a substantial positive trend and many that exhibit fairly dramatic declines despite substantial increases in overall resources allocated to the educational sector. In addition, economic performance over the period, while far from good, was not particularly poor and exhibited a positive trend. HIV/AIDS still looms as a likely major contributing factor.

The agenda for future work is quite full. For example, a more complete explanation of differences in regional trends in educational attainment would be valuable both in terms of understanding the implications of HIV/AIDS more completely and in developing policy alternatives for countering negative impacts. More generally, a focus on constructive policy options is now merited.

## 5. References

Akabayash, Hideo and George Psacharopoulos (1999). "The Trade-off Between Child Labour and Human Capital Formation: A Tanzanian Case Study." *The Journal of Development Studies*, Vol. 35, No. 5, June 1999, pp. 120-140. Appleton, Simon and Arsene Balihuta (1996). "Education and Agricultural Productivity: Evidence from Uganda." *Journal of International Development*. Vol. 8, No. 3, pp. 415-444.

Arndt, Channing (2002). "HIV/AIDS and Macroeconomic Prospects for Mozambique: An Initial Assessment." The World Bank, Africa Region Discussing Paper, January 2001.

Arndt, Channing and Peter Wobst (2002) "HIV/AIDS and the Labor Market in Tanzania". Trade and Macroeconomics Discussion Paper Series 102. International Food Policy Research Institute: Washington, D.C.

Badcock-Walters, Peter and Alan Whiteside (2000). "HIV/AIDS and Development in the Education Sector." Unpublished paper. Health Economics and HIV/AIDS Research Division, University of Natal.

Bloom, David E., David Canning, and Jaypee Sevilla (2001). "Health, Human Capital, and Economic Growth." Commission on Macroeconomics and Health. World Health Organization. *Working paper* no. WG1: 8.

Golan, Amos, George Judge and Douglas Miller (1996). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York: John Wiley & Sons.

Golan, Amos and Stephen J. Vogel (2000). Estimation of Non-Stationary Social Accounting Matrix Coefficients with Supply-Side Information. *Economic Systems Research*. December 2000, pp. 447-471.

Karantininis, Kostas (2002). "Information-based Estimators for the Non-Stationary Transition Probability Matrix: An Application to the Danish Pork Industry." *Journal of Econometrics*, Vol. 107, pp. 275-290.

Lockheed, Marlaine E., Dean T. Jamison, and Lawrence J. Lau (1980). "Farmer Education and Farm Efficiency: A Survey." *Economic Development and Cultural Change*. pp. 37-76.

MoE (various years). Basic Statistics in Education: Regional Data. Various issues, Ministry of Education: Dar es Salaam.

MoH (2000 and 2001). HIV/AIDS/STI Surveillance: Reports Nos.14 and 15. National AIDS Control Programme. Ministry of Health: Dar es Salaam.

Psacharopoulos, George (1994). "Returns to Investment in Education: A Global Update." *World Development*. Vol. 22, No. 9, pp. 1325-1343.

Pinckney, Thomas C. (1997). "Does Education Increase Agricultural Productivity in Africa?" in *Proceedings of the International Association of Agricultural Economists Meeting: 1997*. pp. 345-353.

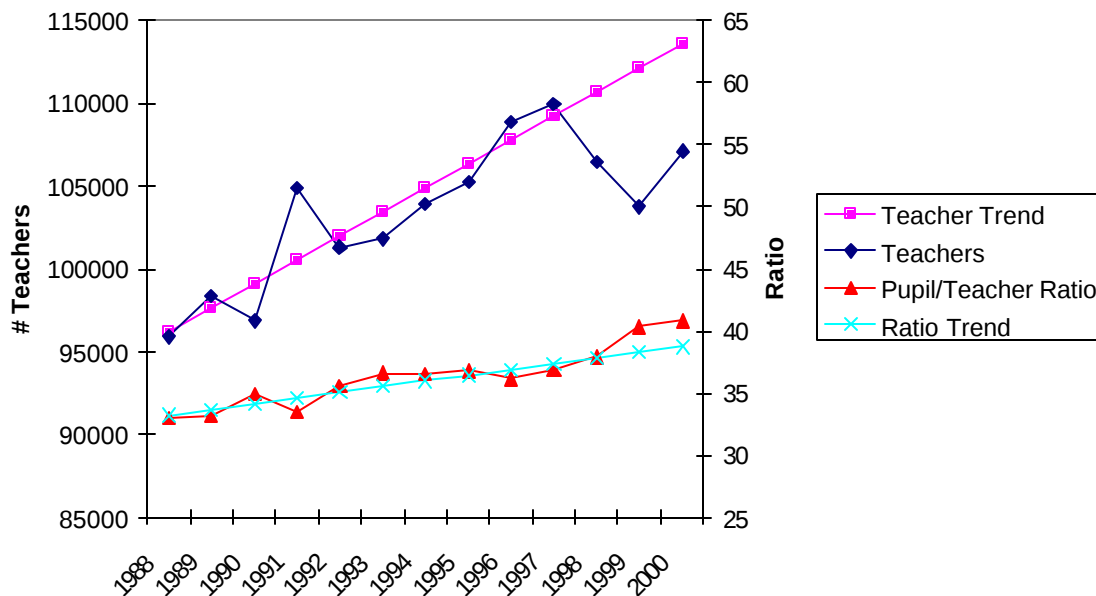
Weir, Sharada and John Knight (2001a). "Adoption and Diffusion of Agricultural Innovations in Ethiopia: The Role of Education." Center for the Study of African Economies, WPS2000-5, Oxford.

Weir, Sharada and John Knight (2001b). "Education Externalities in Rural Ethiopia: Evidence from Average and Stochastic Frontier Production Functions." Center for the Study of African Economies, WPS2000-4, Oxford.

World Bank (1997). *Confronting AIDS: Public Priorities in a Global Pandemic*. New York: Oxford University Press.

## 6. Tables and Graphs

Graph 1: Primary school teachers and pupil/teacher ratio (1988 – 2000)



Note: Teacher and ratio trends are calculated for the period 1988 to 1997 and projected for the remaining periods.

Table 1: Ratios of projected enrollments in grade 7 in 2000

| Province       | Girls | Boys  | All Students |
|----------------|-------|-------|--------------|
| NATION         | 1.086 | 1.151 | 1.120        |
| HIGHAFF        | 1.078 | 1.119 | 1.098        |
| HIGHXM         | 1.239 | 1.254 | 1.247        |
| LESSAFF        | 1.084 | 1.169 | 1.126        |
| ARUSHA         | 1.038 | 1.132 | 1.092        |
| <i>DSALAAM</i> | 1.129 | 1.318 | 1.215        |
| DODOMA         | 1.060 | 1.315 | 1.195        |
| <i>IRINGA</i>  | 1.053 | 0.947 | 0.993        |
| <i>KAGERA</i>  | 1.298 | 1.215 | 1.265        |
| KIGOMA         | 1.089 | 1.148 | 1.120        |
| KNJARO         | 1.031 | 1.052 | 1.042        |
| LINDI          | NA    | NA    | NA           |
| MARA           | 1.157 | 1.258 | 1.214        |
| <i>MBEYA</i>   | 0.960 | 1.078 | 1.019        |
| MOROGORO       | 0.944 | 1.059 | 1.004        |
| MTWARA         | 1.290 | 1.428 | 1.353        |
| MWANZA         | 1.067 | 1.101 | 1.082        |
| PWANI          | 1.021 | 1.162 | 1.091        |
| <i>RUKWA</i>   | 1.310 | 1.348 | 1.292        |
| RUVUMA         | 0.932 | 1.082 | 1.005        |
| SNYANGA        | 1.133 | 1.048 | 1.037        |
| SINGIDA        | 0.971 | 1.068 | 1.017        |
| TABORA         | 1.255 | 1.140 | 1.201        |
| TANGA          | 1.187 | 1.212 | 1.178        |

Notes on Table 1:

- 1) The figures are the ratios of projected enrollments in grade 7 in 2000 based on the 1991 and 2000 transition matrices estimated by the non-stationary approach. The ratio is 1991/2000 so a number greater than one indicates a negative trend in enrollment transition probabilities.
- 2) All enrollment figures are in thousands of students.
- 3) The estimator for Lindi region failed to converge.
- 4) Regions characterized as high afflicted (HIGHAFF) are in italics.