PRESCHOOL NUTRITION AND SUBSEQUENT SCHOOLING ATTAINMENT: LONGITUDINAL EVIDENCE FROM TANZANIA

Harold Alderman
Johannes Hoogeveen
Mariacristina Rossi

Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania

Harold Alderman
The World Bank

Johannes Hoogeveen
The World Bank

Mariacristina Rossi
University of Rome II & CeRP CCA

Abstract

This study analyses how childhood health determines future academic performance in Kagera region in Tanzania. Academic outcomes considered are years of education and delay in enrollment, while the measure of childhood health is (relative to the median) height. The repercussions of malnutrition in childhood on subsequent learning and school performance are analyzed by using a unique longitudinal dataset. Results indicate the degree to which malnutrition leads to reduced lifetime earning capacity due to both delays in schooling and declines in total schooling.

1. Introduction

Severe hunger attracts public attention. In contrast, chronic hunger and malnutrition receive less policy interest although they are considerably more prevalent in developing countries. Malnutrition is neither a short term crisis nor a disease which brings immunity for survivors. Rather it is a life cycle process that affects productivity and resources devoted to health over a lifetime.

While decreasing the incidence of malnutrition is often justified on intrinsic grounds as well as in terms of reduced mortality, potential gains in productivity and reductions in economic costs contribute much of the total economic returns to investing in nutrition (Behrman et al., 2004). Such estimates are largely driven by the expected increase in future productivity and thus higher future wages that come from investing in maternal and child nutrition. If an adequately nourished child starts school earlier or repeats fewer grades at school, the child will either enter the labor force earlier than an undernourished comparable child or will complete more schooling or both. Even without a probable increase in cognitive abilities (Behrman et al., 2004) this child will receive higher earnings for his/her entire productive life.

A number of empirical studies confirm the underlying logic that nutrition programs targeted to children in their early childhood have appreciable influence on education. Many of these studies have controls for the fact that the factors that determine malnutrition also influence schooling (Maluccio et al., 2006, Alderman, Hoddinott and Kinsey, 2006, Glewwe and Jacoby, 1995, Glewwe, Jacoby and King, 2001, Behrman, 1996).[1]

The current paper looks at long-term impacts of nutritional status during childhood, concentrating on both delay in school entry and on the total years of schooling attained. Either of these could affect long-term economic prospects of children. Our aim is to detect

the impact that an improvement in nutritional status in childhood will have on the child and the future adult. In particular, we ascertain if and to what extent malnourished children are more likely to show lower academic grades ultimately achieved and if they achieve them later on time due to school entry delay. We ask if malnutrition affects delay in starting school only or both the time of entry in school and the final grades accomplished. We employ panel data from Kagera region in Tanzania for this task, and, thus, link a child's observed nutritional status with that child's subsequent school outcomes.

2. Analytic Approach

The most nutritionally vulnerable period of the life of a child is the first 18 months of life (Shrimpton et al. 2001). A child who is malnourished in the first stage of life is likely to remain stunted in his/her adulthood. Moreover, malnutrition early in life is negatively associated with cognitive development; malnourished children will most likely perform worse at school than their healthy siblings. Thus, the health consequences of poor nutrition may influence the amount and timing of school enrolment. A child who has fewer completed years of schooling or has learned less in his or her time in school or both will exhibit a lower productivity when in the labor market.

A few studies have separated the direct impact of nutrition on schooling from other factors that both determine schooling as well as nutritional status (See Behrman 1996 and Alderman, Hoddinott and Kinsey, 2006 and the references in these articles). This evidence could be in terms of enrolment delay, which postpones the age at which the final grade is achieved, or in the amount of grade repetition or comparatively early termination of studies. However, since many studies are based on children still in school, one needs to take this sample characteristic into account in order to determine the relation between ultimate grade achieved and malnutrition. Our study disentangles the impact of nutritional status both on

delayed enrolment and final grade achieved by accounting for right hand truncation of the outcome variable, without which the impact of malnutrition on schooling would be exaggerated (since students who enrolled late – the most malnourished - are more likely to still be in school).

The measure of nutritional status of the child that we use, percentage of median reference height, reflects parental decisions on investment in child's health as well as the child's genetic endowment. Child's height is thus the outcome of both household and community characteristics, including some transitory shocks. We consider two periods of life of an individual: early childhood and adolescence. Adolescence outcomes are observed in 2004 in terms of grades attained. These, as well as delays in enrollment, are related to childhood nutrition status in 1991-1994. Our analysis exploits the longitudinal dimension of the panel by linking information on individuals aged 10 or less in 1991-1994 with new data on the same individuals collected ten years later. To estimate the impact of children's nutrition on their future educational attainments we perform a regression of education achievements of respondents aged 10 or less in their last interview between 1991-1994, by using the height of the child at that time relative to that of the median of the reference population, for the same age, month and sex, as a regressor. As indicated below, we also used an instrumented version of this regressor in alternative specifications.

Delay could reflect a filter imposed by schools if administrators use height as a sign of school readiness or it could reflect parental assessment of whether their child is ready to start the learning process or not. Glewwe, Jacoby and King, (2001) verify that late school entry may improve a child's ability to perform school work. Nevertheless, even if, contrary to a plausible and common assumption, a child's capacity to learn is uncorrelated or positively correlated with entry date late entry may lead to a decline in total school attained since it likely raises the opportunity cost of each year of school. Moreover, even if entry delay did not influence the total years of schooling it would still be costly as the child will enter the

labor market with a delay, *ceteris paribus*. Such a postponement reduces life-time earnings. Glewwe and Jacoby (1994) calculate the cost of an average delay of two years of school to be about 6% of the individual life-time wealth, assuming the interest rate to be equal to 3% and an infinite horizon[2].

3. Econometric Model

As mentioned, we look at the age at which a child starts school as well as completed years. Since decisions on school entry as well as on desired grade completion both reflect ability which is observed by the household but not by the researcher, we test for a negative correlation between delaying entrance and grade achieved. That may come about, for example, if children entering school earlier are more likely to be more able to learn as well.

The reduced form of the two jointly estimated outcomes can be written as follows:

$$D_k = \alpha_D f(S_{1k}) + \beta_D Z_k + u_k \qquad (1)$$

$$A_{2k} = \alpha_A f(S_{1k}) + \beta_A Z_k + v_k \qquad (2)$$

Where subscript 1 and 2 refer to period 1994 and 2004, respectively; $f(S_{1k})$ is a function of S_{1k}, which represents the nutritional status of respondents when interviewed ten years ago and allows for a more flexible impact of health status on educational outcome; D_k and A_{2k} stands for years of delay in enrollment and school attainment of the child k in the second period of observation. D_k can assume positive discrete values as does A_{2k} for which the maximum value is, in general, the maximum grade attainable[3]. The terms v_k, u_k are individual specific error terms correlated via the correlation coefficient ρ, to be estimated (Glewwe and Jacoby, 1994). Z is a vector of household, community and individual characteristics.

Both dependent variables are latent variables observed as discrete and ordered variables. The actual number of individual years of schooling is determined by the value of a

latent (unobserved) variable which indicates the (latent) propensity for schooling as is the propensity to go to school on time, of which we observe the counterpart D, the number of years of delay. The dependent variables, D_k and A_{2k}, estimated are thus redefined as unobserved latent variable D_k^* and A_{2k}^*, corresponding to the observed values of years of delay in school enrollment and school attainments. An individual will have *m* years of delay and *n* years of schooling if the latent variable D_k^*, A_{2k}^* falls between two threshold parameters or cut-off points corresponding to years of delay *m* and *m+1* and grades *n* and *n+1* as follows: $\text{cut}_m < A_{2k}^* < \text{cut}_{m+1}$, $\text{cut}_n < A_{2k}^* < \text{cut}_{n+1}$ where the threshold parameters (cut_m, cut_n) have to be estimated by the model along with the set of coefficients β and α.

Assuming normality of the error terms in equation (1) and (2) we can write the probabilities associated to each of the possible state of delay in enrollment and years obtained. For example the likelihood term corresponding to *n* completed years of schooling with delay is as follows:

$$PR(years = n \ \& \ Delay = 1) = \Phi(years < n+1, Delay = 0, \rho) - \Phi(years \leq n, Delay = 0, \rho).$$

where $\Phi(Y<y, X<x, \rho)$ stands for the joint cumulative bivariate normal distribution of Y and X, with correlation coefficient equal to ρ.

We allow malnutrition to have an impact on both educational outcomes. Malnourished children are less ready to go to school and also less able to learn when in school. In addition to influencing the timing of schooling as well as overall ability, malnutrition could have an indirect effect on the level of education obtained if a delay in school entry shifts school entry to a later point in time. Although the gap in educational years could still be recovered over time if delaying school has no effect on the actual grades obtained, children who delay are usually associated with higher drop-out of school and worse learning when in school.

The dependent variable, years of schooling, indicate a cumulative investment in an individual's education. Both delay on school entry and grades attained are the observed

outcome of the latent propensity for schooling. Moreover, it is the outcome of a series of ordered discrete choices (see for similar approach Glewwe and Jacoby 1994 and Glick and Sahn 2000) and not a continuous variable of grade attainment. Going to next grade is conditional on the previous grade attainment. An additional feature of the sample is censoring of the variable of interest. Years of education are right censored for those who are still in school; years attained do not necessarily indicate the total years of education to be achieved by the respondent if (s)he is still going to school. As the respondent's age is between 10 and 20, most of children are still enrolled in school and expected to complete additional years of education.

One way to eliminate this problem of right censoring would be to consider only a sample of respondent old enough to have already completed their educational attainments by the time of the interview. However, by restricting the sample to older respondents we would loose the information on their nutritional status in their childhood, given the characteristics of the data. As the focus of this study is to detect how malnutrition in early childhood affects outcomes as adolescent, we select those respondents who were children in the first waves of the survey and control for the right censoring of the dependent variable. Consequently we employ two ordered probit model equations to estimate both the determinants of years of schooling and delayed years of school entry, extended to control for right censoring of the years of school. Given the measure of the school outputs, it would be unwise to use an OLS technique, which implies a cardinal ranking between grades completed, while grades completed can only be ordered.

Several functional forms and indicators were tried to identify the type of relation between health nutritional status and educational outcome. With any indicator used the impact of nutrition on educational output is clearly non-linear. We thus used a quadratic form of height (as a percentage of the median of the reference population with the same age and sex) as regressor to control for the non-linearity. This variable maps to z-scores that are

commonly used in studies of nutrition, but has the advantage of always being a positive number. We verified that this functional form did not force the curvature by comparing the results with that using spline functions with a node at z scores for height of -2 and with nodes at 85 and 95 percent of median heights.

4. Data Set

The study uses a panel data set that combine the 2004 Kagera Health and Development Survey (KHDS) with all individuals who were household members in any round of the KHDS 1991-1994.[4] All household members in 2004 belonging to the same household of respondents in 1994 were administered the full household questionnaire which collected information on a wide range of topics, including: housing amenities, consumption, income, assets, time allocation of individuals, business activities, remittances, support from organizations, education, and health, including anthropometric measures, which are the key variables of our study. The community questionnaire collected data on the physical, economic and social infrastructure of the baseline communities.

While the questionnaires for the KHDS 2004 were revised to take into account the 10-year retrospective, where possible, comparability is maintained with the baseline KHDS survey instruments. Revisions to the questionnaires were made in order to reflect changes in the region since 1994 (such as, the presence of new community organizations or recent health campaigns). Moreover, the revised questionnaire was redesigned in an effort to capture key transitions that have occurred since the previous interview. These revisions included expanded questions on the circumstances of deaths. For individuals who no longer reside together, there is information on remittances, loans, bride price payments, social communication and labor transfers between previous members. For all panel respondents, there is a module on the incidence of economic shocks (both positive and negative) in the last 10 years. For respondents who re-located since their interview in 1991-1994, there is

information on migration. Lastly, a special module on informal insurance groups was included.

A considerable effort was made to track surviving respondents to their current location, be it in the same village, a nearby village, within the region, outside the region, or even outside the country. The success of panel surveys is often measured in terms of re-contact of households, rather than individuals therein. By this measure, excluding households in which all previous members are deceased (17 households with 31 people), the KHDS 2004 survey re-contacted 93% of the baseline households (835 out of 895 households; 8.4% of these were interviewed outside Kagera). At the individual level, 78% of children who were under 10 years old in the first wave were reinterviewed, 7% died and 15% were not found. This attrition rate compares favorably with other panel surveys in both low-income countries and high-income countries.[5] All households where at least one member was ever interviewed in 1991-1994 were re-interviewed in 2004. Over the decade, it was found that many of the original household members now formed a new, separate household. Typically individuals noted as dependent children in the early survey lived independently in the later one; others got divorced, orphaned or went to live with family. The members of the original 835 households were associated to over 2,700 households in 2004.

Following the methods set out in Fitzgerald, Gottschalk and Moffitt (1998) and Alderman et al. (2001b), we estimated a probit to determine whether there was attrition based on observable variables. This involves regressing a dependent variable equal to 1 if the school attainment is observed in 2004 and 0 otherwise on initial height-for-age and a set of child and family characteristics. As there was no statistically significant relationship between initial height-for-age and attrition – the coefficient was -.005 with a standard error of .027 – there is no evidence of attrition bias.

5. Descriptive Statistics

Our analysis focuses on the impact of health in childhood on academic achievements. These two outcomes of investment in human capital are measured at different stages in life. School achievement in year t is a cumulative result of regular investments including those in nutrition in early stages of life. The health status of an individual reflects various interacting factors such as the quality of environment, genetic factors and available resources, which are other inputs in the academic achievement production function[6]. As mentioned, the children in the sample are aged between 10 and 20 years and some of them may have not completed their educational attainments. Years of education are, therefore, right censored for those children who are still in school.

Figure 1 plots achieved years of education against age of the child both for the sub-sample of children who were stunted in the first wave of data collection and those who were not. If the main difference between malnourished children and others is in the date of school enrollment, we would expect a constant difference in years of school for those who are stunted when children and those who are well nourished in their childhood. Conversely, if stunted children drop out of school earlier or repeat grades more frequently, the gap may increase over time. The highest gap in attainment in figure 1 is observed in the age category 15-16. The gap shrinks slightly between 17 and 20 years suggesting that a fraction of stunted children may have delayed their school completion relative to non-stunted children. However, the gap does not close appreciably indicating that stunted children have lower academic attainments overall[7].

Table 1. Years of delayed enrollment for respondents aged 10 or less in 1991-1994 who eventually attended school and percentage never went to school

	Female		Male	
	Not stunted (1)	Stunted (2)	Not stunted (3)	Stunted (4)
Years of delayed enrolment (if attended school)	1.58	2.19	1.84	2.36
Never attended school	7.94	14.49	7.35	8.43

Source: KDHS Survey 1991-2004

Table 1 reports the average years of delay in school entry for children that were malnourished during their childhood and those that were not with a breakdown by gender. Enrollment is considered delayed when a child has not entered school at the age of 7. The majority of children in Kagera (80%) do not begin school at the official starting age of 7; the average delay of those who do not start on time is two years. Nevertheless, both boy and girls are delayed in their school entry when stunted more than their healthy counterparts (about half a year). In addition, consistent with Alderman et al. (2001a), malnourished girls appear to have an increased probability of never attending school. The increased delay attendant to malnutrition could reflect a parent's assessment of school readiness or rationing of school spaces by administrators. Additionally, it could be motivated by other factors such as isolation. Households who are closer to school are probably less adverse to send their children to school at the right age. However, as all villages in the sample had a primary school this is unlikely to be a factor for school delay in this study.

6. Estimation Results

Summary statistics are provided in Table 2 and estimation results are shown in Table 3.

Column 1 in Table 3 reports coefficient estimates for equations estimated with nutrition taken

as exogenous while column 2 repeats the estimation with an instrumented variable for nutrition. Delay in enrolment is reported in 1(a) and (2a) while years of education, estimated by using an ordered probit model which takes into account the censoring of the variable are reported in 1(b) and 2(b). We tested whether the correlation coefficient of the errors in the two estimates, ρ, was significant by first estimating the two equations jointly with maximum likelihood using with the Berndt-Hall-Hall-Hausman technique. As the estimated value of ρ was -0.089 (s.e.= 0.22), we do not impose further cross equation restrictions in the estimates reported.

One potential problem with our set of estimates in columns 1 relates to endogeneity of the childhood nutritional status, by generating a correlation between nutritional status and the error component u_k and v_k. Endogeneity could stem from unobservable household preferences towards children investments. Households with higher propensity to children's investments would feed them better as well as enroll them in school earlier and keep them in school longer. Moreover, a child specific component (ability) that is known to parents but not the researcher might lead them to invest differently in each child. If children who show higher abilities to school are also better fed by parents, as suggested by Glewwe and Jacoby, (1995) the effect of height would be overestimated.[8]

Thus, we need to find instruments that are correlated with nutritional status in childhood and not correlated with accomplished academic years. We take into account the potential endogeneity of nutritional status by instrumenting nutrition status using crop loss reported by the household as instruments as well as recent weather shocks at the community level interacted with the age (at the time of the interview)[9]. The instrumenting equations use community fixed effects to rule out the possibility of bias from unobserved community level effects including infrastructure that influence both nutritional status and schooling. This rules out the inclusion of covariate drought and similar shocks as instruments, although it is still

possible to include interactions of these community level shocks and individual characteristics.

As drought and other agricultural shocks are time varying and been shown to affect nutritional status of young children[10] (Hoddinott and Kinsey, 2001) these are conceptually valid if the shocks are not themselves long lasting[11]. Similar shock variables were used by Alderman Hoddinott and Kinsey (2006) who compared such instruments with a preferred maternal fixed effects instrumented approach and found little difference in the magnitude of the coefficients (table 4 page 464) though both approaches to instrumenting resulted in larger coefficients than if the nutrition variable was not instrumented.

We also perform a set of standard tests to verify the choice of instruments. For example, as instrumental variables methods are potentially biased if the instruments are weak, that is, if they do not explain the endogenous regressor or if they influence the variable of interest (in this case, schooling outcomes) directly rather than only through the endogenous variable we include specification tests. These are discussed in conjunction with the ancillary equation in appendix 1. The instruments also pass the over-identification test, for the linear regression of years of education run on exogenous and instrumented endogenous variables. The Hansen[12] J-statistics for over identification restrictions are equal to 2.2 and 11 for the models of delay in school entry and years of school accomplished, respectively. Thus, the null hypothesis of absence of correlation between instruments and the error term cannot be rejected. Moreover, the F-test on the instrument set is equal to 8.3, which is above the critical value at 10% of error (see Stock and Yogo, 2004) indicating that the instruments have explanatory power. As the instrumented equations are conceptually preferred, we will confine the remainder of the discussion to the results in column 2.

These findings confirm that malnutrition in childhood affects the age of school entry and are in line with those of Glewwe and Jacoby (1995). The results also show that shorter stature in childhood has an impact on the ultimate grade achieved. However, the impact of

better nutritional status on school attainments declines as nutritional status improves.[13] As mentioned, this pattern was confirmed with spline functions.

Other significant determinants of education are the logarithm of household expenditure and the maximum grade of education within the family (excluding the individual). The same variables act in opposite way to determine enrolment delay; children belonging to wealthy households being more likely to be sent to school on time.   School quality – as indicated by both teacher pupil ratios and by boards per class – also influence schooling. However, gender affects neither school entry nor completed years of school.[14]

Low stature in childhood is thus responsible for a direct effect on both delay in enrolment and years of school accomplished. As malnutrition is responsible for school enrolment delay, and delay affects the ultimate grade achieved, it follows that malnutrition also affects the final achievements, through its impact on enrolment delay.

The set of coefficients obtained by estimating equations (1) (2) and (3) do not represent the marginal effect on the dependent variable associated to each regressor. Moreover, even the sign of the coefficients, for equation (2) does not necessarily correspond to the sign of the marginal effects (see Long 1997). As this makes the estimation results less apparent, we provide a simulation based on the predicted values in Table 4 using the instrumented results. As our model is non-linear, the impact of nutrition improvement depends on the value at which we set other regressors. Our predictions are calculated as the average years of school of the observations in our sample if nutritional status of malnourished respondents were improved holding the other variables at their observed level rather than at predicted value at the mean of these variables. For example, if the height of a boy is increased from 80% to 95% of the median, this will increase schooling by 0.93 years, while the increase will be 0.85 years for boys whose height is equal to 85% of the median.[15] The impact is slightly less for girls at the respective percentage of reference heights. Children

who are severely malnourished face almost twice the risk of never attending school than well nourished children.

6. Discussion of Results on Schooling

The analysis of the Kagera panel shows that the children who are malnourished have lower schooling and delay their school entry. What does this mean in terms of lost earnings? To address this question we regressed wage earnings from the same overall 2004 Kagera sample on schooling, controlling for sample selection (appendix 2). Each additional year of schooling leads to an increase in the logarithm of earnings of 0.083. An alternative way of measuring returns to schooling is to use per capita consumption instead of earnings. Per capita consumption reflects the resources available to the household and is observable for all households, while earnings only for those individuals working as wage employees. The return to schooling measured by using this alternative method is similar in magnitude to that found using the earnings equation despite the potential bias from correlations of unobserved assets and schooling.

We then used the estimated return to schooling in the simulations on the benefits of the additional salary that are presented in table 5. We calculate the corresponding annual benefit derived from the additional earning associated with an improvement in nutrition, through the channel of additional years of education, by making the assumption that the supply of hours worked does not change. Since bringing a child from 85 to 95 percent of median height will increase schooling by approximately three quarter of a year on average, this leads to 30,000 TZ shillings (calculated as 8% of the average annual salary, including boys and girls)[16]. The stream of benefits perceived during life time due to better nutrition (LTB) are calculated as follows:

LTB: $\left(\dfrac{1}{1+r}\right)^{E-1} \sum_{i=1}^{60-E} \dfrac{Y+B}{(1+r)^i}$ (3)

where E is the age of entrance into the labor market, estimated as the sum of starting school age (7), estimated number of years of delay and estimated number of completed school years. B, is the additional benefit in monetary value due to additional 0.75 school years. The benefits are reported in terms of the present value of the annual benefits for a number of years equal to 60 (assumed retirement age) minus the completion of schooling age (as in Table 4), discounted to age 1. Discounting the stream of benefits to the first year of life of the child allows comparison of benefits with the costs of a nutritional intervention; as mentioned, the most appropriate year for such an intervention is the first year of life when a child is most vulnerable to malnutrition. The overall impact of an improvement in nutrition, through its effect on delay in school entry and additional years accomplished, is summarized in the age at which the child would finish school.[17]

The impact of an increase in nutrition is always positive. For an interest rate of 5%, an investment on nutrition targeting children whose height is 85% and reached 90% of the median height after intervention lead to a monetary benefit of about 200,000 TZ Shillings.

By considering an average discount rate at 5% and concentrating only on the impact of nutrition on schooling, we can thus conclude that any program successfully able to moving a child from 85% of median height to 95% of median that cost less than 276,963 TZ Shillings would be justified in economic terms (1,000 TZ Sh are equal to 0.82 $). .

This estimate may even be conservative since, in order to focus on the impact on education, we do not include the impact of adult height on wages in the comparison of expected benefits and possible program costs. Height is clearly significant in the wage regression as it is in many other studies, from both urban and rural setting and from developed as well as low income countries (Thomas and Strauss, 1998). However, it is not certain that in addition to reflecting work capacity, the height coefficient does not also pick up differences in cognitive ability associated with avoiding nutritional shocks in childhood. Indeed, it would

be desirable to have an estimate of the additional impact of improved cognitive ability; in the absence of any measures of cognitive ability or of test scores we underestimate the total influence of nutrition on earnings under the plausible assumption that nutritional status influences the amount of learn conditional on the years of schooling. If height is included in our calculations the benefits increase so that a program would have a favorable benefit cost ratio if it costs less than 658,671 TZ Shillings (calculated as the sum of the difference in earnings in column (5) and column (2), over an individual's lifetime at 5% discount rate), or equivalently, 658 $.

In a recent paper, Alderman, Hoogeveen, and Rossi (2006) show a strong effect of program interventions on children's nutritional status. If children's nutritional status contributes to their academic achievements, and early childhood nutrition outcomes persist over time, such successful early childhood nutrition interventions can lead to improved education attainments for adolescents.

Figure 1. Completed number of school years by age categories.
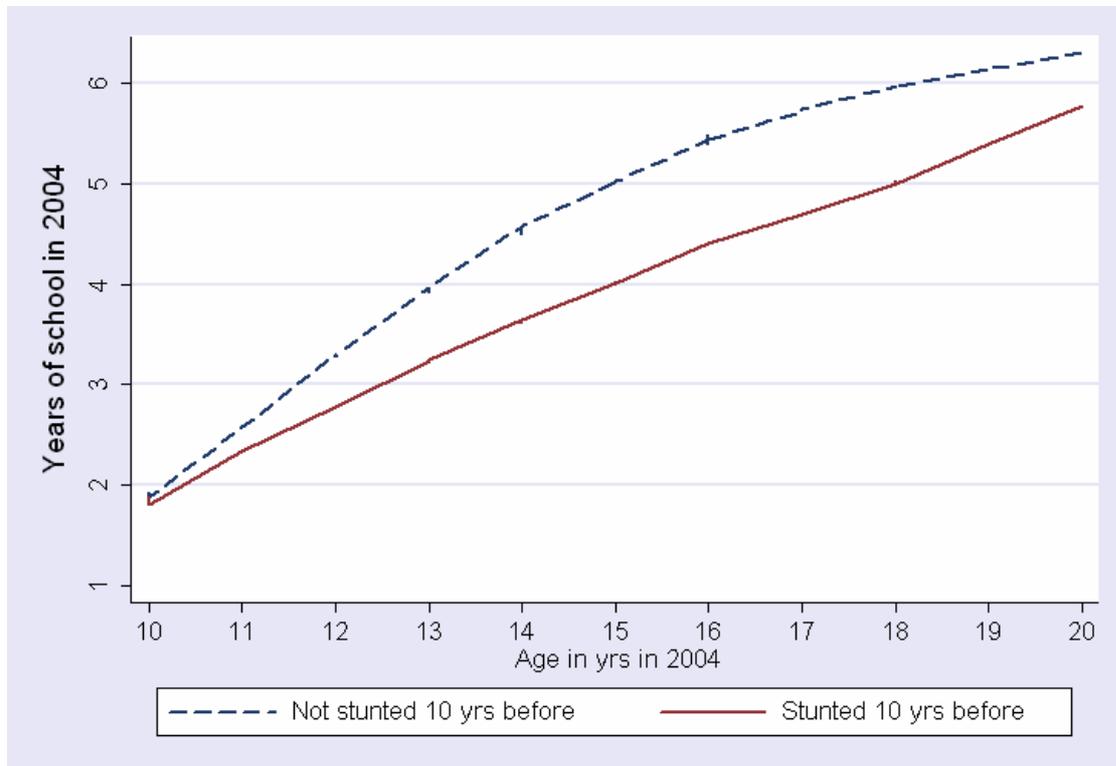
Figure 2. Years of education of panel respondents aged 10-20 in 2004
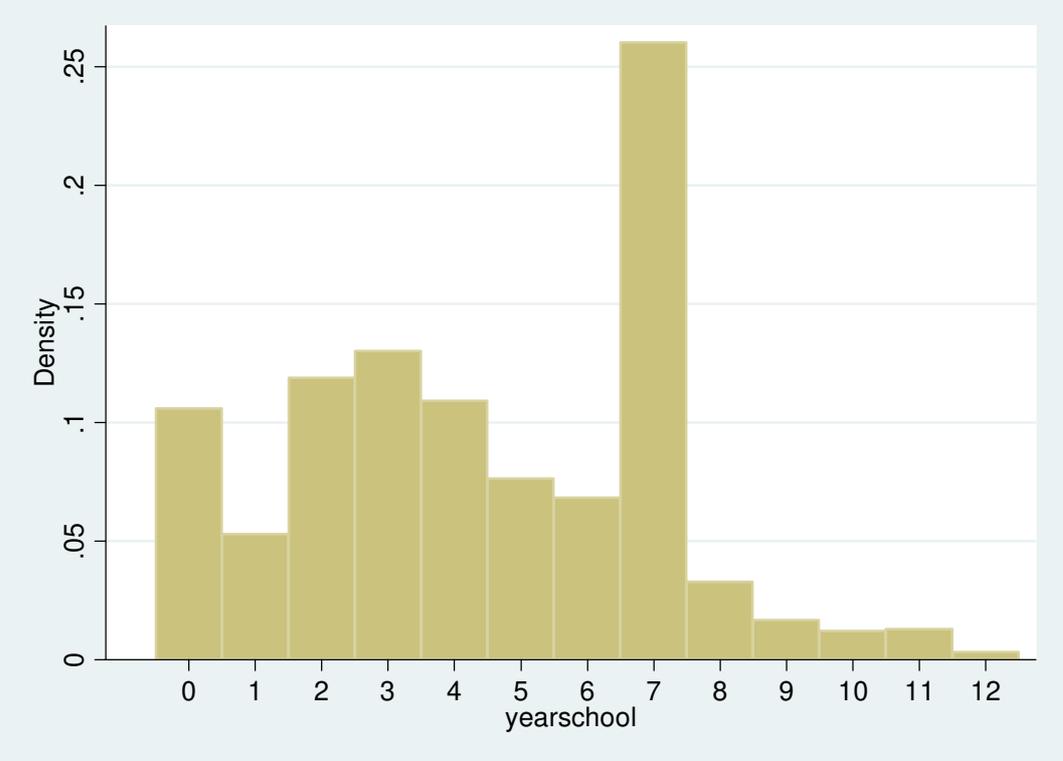
Table 2. Summary Statistics

	Mean	Standard Deviation
% of median of reference height	93.17	4.95
% of median squared	8704.56	926.88
Parent had Secondary schooling	0.08	0.26
# teacher per class	1.66	0.77
# board per class	0.84	0.20
Male	0.51	0.50
Age of child in years	15.69	3.25
Age squared	256.56	105.17
Mother's height	158.392	5.441
Father's height	168.599	5.393
Urban	0.17	0.38
Maximum education in household	6.80	2.03
(Log)per capita hh expenditure 2004	11.92	0.64
Electricity	0.63	0.48

Community regressors are at baseline.

Table 3. Years of Education and Delayed School Entry.

School Entry Years of Delay	Coefficients for Uninstrumented model	Coefficients for Instrumented model
Regressors	(1a)	(2a)
% of median	-0.485	-0.591
	(5.05)***	(3.42)***
Residuals % median		0.047
		(0.58)
% of median^2	0.002	0.003
	(4.96)***	(3.39)***
Residuals % of median^2		-0.000
		(0.79)
Female	-0.103	-0.062
	(1.63)	(0.87)
Age in years	0.053	0.037
	(5.17)***	(2.96)***
Mother's height/10	-0.001	0.025
	(0.02)	(0.33)
Father's height/10	0.034	0.014
	(0.59)	(0.21)
Secondary school	0.022	0.119
	(0.17)	(0.82)
# teachers/class	0.003	0.006
	(0.06)	(0.12)
# boards/class	-0.499	-0.532
	(3.02)***	(2.89)***
Urban	-0.178	-0.162
	(1.63)	(1.39)
Max years of school in the household	-0.066	-0.077
	(4.24)***	(4.51)***
(Log)per capita expenditure	-0.271	-0.289
	(5.15)***	(4.82)***
Electricity in the household	0.122	0.100
	(1.67)*	(1.20)
Completed Years of School	(1b)	(2b)
% of median height	0.380	0.575
	(2.74)***	(2.18)**
Residuals % median		0.000
		(0.25)
% of median^2	-0.002	-0.003
	(2.94)***	(2.41)**
Residuals % of median^2		0.018
		(0.16)
Female	0.005	0.017
	(0.06)	(0.17)

Age in years	-0.044	0.009
	(2.91)***	(0.39)
Mother's height/10	0.129	0.233
	(1.50)	(2.06)**
Father's height/10	-0.005	-0.014
	(0.06)	(0.14)
Secondary school	0.297	0.235
	(1.55)	(0.98)
# teacher/class	-0.081	-0.083
	(1.31)	(1.11)
# boards/class	0.760	0.951
	(3.41)***	(3.44)***
Urban	0.069	0.232
	(0.49)	(1.45)
Max years of school in the household	0.116	0.126
	(5.40)***	(4.96)***
(Log)per capita expenditure	0.274	0.312
	(4.10)***	(3.89)***
Electricity in the house	0.052	0.064
	(0.54)	(0.58)

Observations: 1147 panel respondents aged 10-20 in their last interview in 1991-1994. Right censored observations (children attending school): 739.

Absolute value of z statistics in parentheses.

* Significant at 10%; ** significant at 5%; *** significant at 1%

All standard errors are adjusted for cluster effects.

Vector b and g of equation (1) and (2) are not the marginal effects.

The constant threshold parameters of the ordered probit estimates are not reported

Table 4. Simulations based on regression predictions.

Child Height	(1) If 80% of median height	(2) If 85% of median height	(3) if 90% of median height	(4) If 95% of median height
1a. Expected years of education. Male	5.82	5.90	6.64	6.75
1b. Expected years of education Female	6.02	6.07	6.62	6.77
2a. Expected years of delay. Male	3.16	2.82	2.22	1.98
2b.Expected years of delay. Female	3.24	2.68	2.15	1.85
3a. Probability of never going to school. Male	0.13	0.12	0.08	0.07
3b. Probability of never going to school. Female	0.10	0.11	0.07	0.06

The simulations are based on the prediction of expected number of years E(Years). Simulations are based on our preferred model, which instruments past nutritional status. The set of coefficients used for predictions is that in column (2a) and (2b), which allow for endogeneity of nutritional status. The expected number of years of schooling for a child who enter school on time, for example, is calculated as: $E(Years) = \sum_{i=1..12} i * PR(years = i)$

and the probability PR(i) is derived as:

$$PR(years = i) = [\Phi(years < i+1) - \Phi(years \le i)]$$

Where: $\Phi(.)$ is the cumulative normal distribution function. Simulations are calculated as the average of predictions rather than using a prediction at average values. For example, column 2 is derived as the average predicted years completed by increasing to 85% of the median the height of all individuals whose height is below that level.

Table 5. Total life-time benefits from an additional year of education and earlier school entry using different discount rates.

Discount rate	Total lifetime earnings (baseline)	Total lifetime earnings if 90%	Total lifetime earnings if 95%	Present discounted value of benefits including height	Change in benefits from 85% to 90%	Change in benefits from 85% to 95%
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1%	15,085,656	15,828,642	16,047,823	17,407,808	**742,987**	**962,167**
3%	9,061,757	9,501,055	9,645,048	10,462,425	**439,299**	**583,292**
5%	4,227,190	4,425,738	4,504,154	4,885,861	**198,548**	**276,963**
10%	1,186,097	1,237,684	1,266,967	1,374,338	**51,587**	**80,870**

Monetary values are expressed in 2004 TZ Sh.
Annual benefit: 17,259 TZSh.
Present values are calculated by discounting, using the interest rate in column 1, the stream of earnings perceived until the age of 60, starting when the school finishes (calculated as 7 + the number of delayed enrolment years + the number of years at school,) and discounted to year 1.
Column (2) is calculated as the present value of the stream of earnings over the working life-time, supposing earnings are constant and equal to the average individual earnings (490,000 TZ Sh.).
Column (3), (4) illustrates the present value of additional school if height is increase to 90% and 95% of the median, respectively.
Column (5) illustrates the present value of additional 44,000TSh (9% of annual salary) due to additional 5cm of height.
Column (6) is calculated as the difference between (3) and (2), column (7) as the difference between (4) and (2).

Appendix 1.

Table A1. **Ancillary regression used to instrument endogenous variable of nutritional status in childhood.**

Dependent variable: height as a percentage of the median of the reference population[1]	Coefficients	Variable Mean
Crop Loss & age less or equal 3	-0.958	0.365
	(1.68)*	
Crop Loss & age 4-5	-1.602	0.177
	(2.92)***	
Crop Loss & Female	0.304	0.399
	(0.49)	
Flood & age less or equal 3	-1.005	0.210
	(1.66)*	
Flood & age 4-5	0.241	0.102
	(0.34)	
Flood & Female	-0.286	0.230
	(0.51)	
Female	0.663	0.484
	(1.04)	
Age in Years	-0.425	4.441
	(4.99)***	
Mother's years of Education	-0.007	5.176
	(0.12)	
Father's years of Education	0.140	6.110
	(2.42)**	
Mother's Height	0.113	158.392
	(4.11)***	
Father's Height	0.133	168.599
	(4.66)***	
Log per capita expenditures	0.674	10.301
	(2.03)**	
Log per capita land	0.150	10.702
	(2.08)**	
Constant	45.512	
	(6.84)***	

Absolute value of t statistics in parentheses. Community dummies are also included in the regression.

[1]Height refers to the first observation of the individual in the first rounds of the survey (1991-1994). Natural shocks that have affected height in childhood refer to 5-6 years prior to the start of the initial KHDS rounds.

* significant at 10%; ** significant at 5%; *** significant at 1%. R squared equal to 0.12, observations used: 1118.

All regressors refer to last interview during 1991-1994.

Appendix 2. Earnings regression

Nutritional programs are often advocated as instruments to implement educational attainments. Our results show that improved health standards during childhood are effective in increasing the likelihood of completing more educational grades. As education exhibits significantly positive returns, higher attainments imply higher future wages. In order to illustrate how much future wages might increase we estimate an earnings function for wage earners using the same data set as used for school attainment- albeit not the same individuals. Since only a small fraction of the population receives a wage - the majority work in the agricultural sector as a self-employed - this estimate account for sample selection. . We control for working sample selection using a standard Heckman approach.

The main interest in this regression is the coefficient of schooling used in the simulations. However, it is also noteworthy that adult height is also statistically significant. It is also interesting that men earn more as they age and women do not. However, young women seem to have higher earnings than young men. The cross over in this particular specification is around age 40.

Table A2a. Summary Statistics

Regressors	Means	St. Deviations
Lnsalary	5.31	1.14
yearschool	5.75	3.14
Female	0.56	0.50
Age	30.37	10.10
Height	162.28	8.34
yearschool	5.75	3.14
Kids	1.16	1.04
hhmember	5.42	3.06
Old	0.16	0.42

Table A2b. Earning regression.

	(1)	(2)
	(Log)salary	Selection equation
Height	0.018	-0.004
	(3.09)***	(1.02)
Years school	0.083	0.037
	(6.23)***	(4.60)***
Female	0.980	-1.378
	(3.13)***	(8.08)***
Age	0.023	-0.005
	(4.02)***	(1.46)
Female*age	-0.024	0.023
	(2.66)***	(4.56)***
Work in Dar es Salaam	0.519	-0.089
	(1.69)*	(0.46)
# children		-0.011
		(0.38)
Female*# children		-0.038
		(0.90)
# hh members		-0.058
		(5.92)***
# old hh members		-0.176
		(3.07)***
Constant	7.248	0.305
	(7.53)***	(0.52)

Absolute value of z statistics in parentheses

* Significant at 10%; ** Significant at 5%; *** Significant at 1%. Ro estimated equal to -0.73.

Total observations: 3951, uncensored observations: 728

References

Alderman H., H. Hoogeveen and M. Rossi, 2006. Reducing Child Malnutrition in Tanzania Combined Effects of Income Growth and Program Interventions Economics and Human Biology 4(1):1-23.

Alderman, H. Hoddinott, J. and B. Kinsey. 2006. Long Term Consequences of Early Childhood Malnutrition, Oxford Economic Papers 58(3):450-474.

Alderman H., Behrman L., V. Lavy and R. Menon, 2001a. Child Health and School Enrollment. A Longitudinal Analysis, Journal of Human Resources, vol 36 (1)

Alderman, Harold, Jere R. Behrman, Hans-Peter Kohler, John Maluccio, and Susan Cotts Watkins. 2001b. "Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing-Country Samples." Demographic Research 5(4).

Beegle, Kathleen, Rajeev Dehejia, Roberta Gatti. 2006. Child Labor and Agricultural Shocks. Jounral of Development Economics. 81: 80-96.

Behrman J.R., Harold Alderman and John Hoddinott, 2004. Hunger and Malnutrition, Copenhagen consensus challenges and opportunities, In Bjorn Lomborg (ed.) Global Crises, Global Solutions. Cambridge (UK): Cambridge University Press.

Behrman J.R., 1996. Impact of Health and Nutrition on Education, World Bank Research Observer, 11(1) pp 23-27.

Deaton, Angus. 1997. The Analysis of Household Surveys. Baltimore: Johns Hopkins University Press.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. 1998. An analysis of sample attrition in panel data. Journal of Human Resources, 33, 251-299.

Glewwe P. and Jacoby H, 1994. Student Achievement And Schooling Choice In Low-Income Countries. Evidence From Ghana, Journal of Human Resources, pp 844-864.

Glewwe, P., Jacoby, H., 1995. An Economic Analysis Of Delayed Primary School Enrollment In A Low Income Country: The Role Of Childhood Nutrition. Review of Economics and Statistics, 77 (1), 156–169.

Glewwe P. Jacoby and King, 2001. Early Childhood Nutrition And Academic Achievement: A Longitudinal Analysis. Journal of Public Economics, n.81 pp. 345-368.

Glick P. and D. Sahn, 2000, Schooling Of Girls And Boys In A West African Country: The Effects Of Parental Education, Income, And Household Structure, Economics and Education Review, (19), pp 63-87.

Hoddinott, John and Bill Kinsey, 2001. Child Growth In The Time Of Drought. Oxford Bulletin of Economics and Statistics, 63, 409-436.

27

Long J.S. 1997. Regression Models for Categorical and Limited Dependent Variables, Sage Publications, Thousand Oaks.

Maddala G. 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press, Cambridge.

Maluccio, John A., John Hoddinott, Jere R. Behrman, A. Quisumbing, R. Martorell and A. D. Stein, 2006, The Impact of an Experimental Nutritional Intervention on Education into Adulthood in Rural Guatemala, Philadelphia-Washington-Atlanta: University of Pennsylvania, IFPRI, Emory, processed.

Shrimpton R., Victora C., de Onis M., Costa Lima R., Blössner M., Clugston, G., 2001, Worldwide timing of growth faltering: implications for nutritional interventions. Pediatrics, 107, 75-81.

Smith R. and R. Blundell. 1986. An Exogeneity Test for a Simultaneous Equation Tobit Model, Econometrica, 54(2), pp. 679-685.

Stock J.H. and M. Yogo, 2004. Testing for Weak Instruments in Linear IV Regression, Harvard University, mimeo.

Strauss, J., and D. Thomas. 1998: Health, Nutrition, and Economic Development, Journal of Economic Literature, 36(2): 766-817.

Timpone R.J. 2006. Concerns with Endogeneity in Statistical Analysis: Modeling the Interdependence *Between* Economic Ties and Conflict, Political Science, The Ohio State University, mimeo.

Yamauchi F. 2006. Early Childhood Nutrition, Schooling and Sibling Inequality in a Dynamic Context: Evidence from South Africa, International Food Policy Research Discussion Paper # 203.

World Bank. 1993. World Development Report 1993: Investing in Health. Oxford University Press.

Notes to the paper

[1] The literature exploring childhood nutrition on future schooling is inferential, as is this paper. An exception is given by Maluccio et al. (2006), who estimate the impact of a specific nutrition intervention targeted to children before the age of two.

[2] This result is calculated as the effect of postponing of two years the beginning of a stream of equal earnings, and, thus, is not dependent on any particular country or income level.

[3] In our sample the maximum grade obtained is twelve years of school, thus "max" is in our estimates equal to twelve.

[4] See Beegle, Dehejia and Gatti (2006) for another use of this panel data set.

[5] The attrition rate of the panel is much lower than that of some other well-known panel surveys. For example, the attrition rate per year of three surveys summarized in Alderman *et al.* (2001b) ranged from 17.5% to 1.5% per year.

[6] We do not have any information on how the health or the socio-economic conditions of the households and individuals evolved during the time window between the last interview and current one (ten years). However, we observe the initial and final outcomes for the same individuals.

[7] Making comparison among individuals of different age can be potentially misleading. Difference between kids of different age could also reflect cohort effects, instead of being attributed to age passing. In fact, individuals aged *n* in 2004 could differ in their behavior, one year later, from *n+1* years old individuals' behavior, simply due to a cohort effect.

[8] Malnutrition measures are also sensitive to measurement errors. Measurement errors in age would lead to a negative correlation between age and percent of the median nutritional status and hence a spurious correlation between age of school entry and reported nutritional status. There is, however, no systematic bias apparent in reported ages; the age pyramid for the sample is not different from that of Census of Tanzania and, thus, does not point to a particular problem.

[9] In order to control for potential endogeneity of percentage of height and percentage of height squared, we follow the procedure described in Blundell and Smith (1986). We perform the maximum likelihood by including the endogenous variables along with their residuals, which are obtained by regressing the endogenous variables on the instrument set (the first stage regression is illustrated in Table A1). Exogeneity of percentage of median cannot be rejected, given the insignificance of these residuals in the years of school equation, while exogeneity of height is rejected in delay school entry equation.

[10] The shocks used are both at community level and at household level. At community level, we used flood disaster occurred in the community 5-6 years prior to the first round of the KHDS (1991). At household level, we used crop loss happened in the last 5-6 years prior to the interview. Those variables have been interacted with age categories (0-3 and 4-5) and gender, as shown in Table A1.

[11] To the degree that these shocks are long lasting, they would affect income or assets. However, this impact, if any, is packed up in the current expenditure variable included in the regression.

[12] The Hansen J statistics is reported as we run our regression allowing for cluster effects.

[13] This result is in line with that found by Yamauchi (2006) who provides a theoretic justification in that improved health may increase the increase the opportunity cost of schooling as well as increasing the returns to investing in education. These two impacts may be offsetting. In this paper it is shown that increases in nutritional indicator (measure by height z-score) are positively correlated to grade obtained, but the effect becomes negatively correlated for large values of z-scores.

[14] Interactions between gender of the child and nutritional status were used as additional regressors but they were not significant in explaining neither delay in school entry nor years accomplished.

[15] Despite the marked differences in the coefficients in table 3, there is not as pronounced a difference in the simulated values between the instrumented and uninstrumented results. For example, the uninstrumented years of schooling for girls at 85% of median is 6.87 and 7.01 at 90%; The corresponding values for boys are 6.9 and 7.05.

[16] The reason why we take 85% of the median as reference height is because stunted children exhibit, on average, a value of height which is equal to 85% of the reference median.

[17] The first term of the product discounts the stream of benefits to the first year of life of the child, the moment when the program is supposed to happen. The second term is the stream of earnings

discounted at the age of entry into the labor market. Years of delay and completed are taken from estimates in table 4.

Our papers can be downloaded at:

http://cerp.unito.it/publications

CeRP Working Paper Series

N° 1/00	Guido Menzio	Opting Out of Social Security over the Life Cycle
N° 2/00	Pier Marco Ferraresi Elsa Fornero	Social Security Transition in Italy: Costs, Distorsions and (some) Possible Correction
N° 3/00	Emanuele Baldacci Luca Inglese	Le caratteristiche socio economiche dei pensionati in Italia. Analisi della distribuzione dei redditi da pensione (only available in the Italian version)
N° 4/01	Peter Diamond	Towards an Optimal Social Security Design
N° 5/01	Vincenzo Andrietti	Occupational Pensions and Interfirm Job Mobility in the European Union. Evidence from the ECHP Survey
N° 6/01	Flavia Coda Moscarola	The Effects of Immigration Inflows on the Sustainability of the Italian Welfare State
N° 7/01	Margherita Borella	The Error Structure of Earnings: an Analysis on Italian Longitudinal Data
N° 8/01	Margherita Borella	Social Security Systems and the Distribution of Income: an Application to the Italian Case
N° 9/01	Hans Blommestein	Ageing, Pension Reform, and Financial Market Implications in the OECD Area
N° 10/01	Vincenzo Andrietti and Vincent Hildebrand	Pension Portability and Labour Mobility in the United States. New Evidence from the SIPP Data
N° 11/01	Mara Faccio and Ameziane Lasfer	Institutional Shareholders and Corporate Governance: The Case of UK Pension Funds
N° 12/01	Roberta Romano	Less is More: Making Shareholder Activism a Valuable Mechanism of Corporate Governance
N° 13/01	Michela Scatigna	Institutional Investors, Corporate Governance and Pension Funds
N° 14/01	Thomas H. Noe	Investor Activism and Financial Market Structure
N° 15/01	Estelle James	How Can China Solve ist Old Age Security Problem? The Interaction Between Pension, SOE and Financial Market Reform
N° 16/01	Estelle James and Xue Song	Annuities Markets Around the World: Money's Worth and Risk Intermediation
N° 17/02	Richard Disney and Sarah Smith	The Labour Supply Effect of the Abolition of the Earnings Rule for Older Workers in the United Kingdom
N° 18/02	Francesco Daveri	Labor Taxes and Unemployment: a Survey of the Aggregate Evidence
N° 19/02	Paolo Battocchio Francesco Menoncin	Optimal Portfolio Strategies with Stochastic Wage Income and Inflation: The Case of a Defined Contribution Pension Plan
N° 20/02	Mauro Mastrogiacomo	Dual Retirement in Italy and Expectations
N° 21/02	Olivia S. Mitchell David McCarthy	Annuities for an Ageing World

N° 66/07	Maarten van Rooij Annamaria Lusardi Rob Alessie	Financial Literacy and Stock Market Participation
N° 67/07	Carolina Fugazza Maela Giofré Giovanna Nicodano	International Diversification and Labor Income Risk
N° 68/07	Massimo Guidolin Giovanna Nicodano	Small Caps in International Diversified Portfolios
N° 69/07	Carolina Fugazza Massimo Guidolin Giovanna Nicodano	Investing in Mixed Asset Portfolios: the Ex-Post Performance
N° 70/07	Radha Iyengar Giovanni Mastrobuoni	The Political Economy of the Disability Insurance. Theory and Evidence of Gubernatorial Learning from Social Security Administration Monitoring
N° 71/07	Flavia Coda Moscarola	Women participation and caring decisions: do different institutional frameworks matter? A comparison between Italy and The Netherlands
N° 72/08	Annamaria Lusardi Olivia Mitchell	Planning and Financial Literacy: How Do Women Fare?
N° 73/08	Michele Belloni Rob Alessie	The Importance of Financial Incentives on Retirement Choices: New Evidence for Italy
N° 74/08	Maela Giofré	Information Asymmetries and Foreign Equity Portfolios: Households versus Financial Investors
N° 75/08	Harold Alderman Johannes Hoogeveen Mariacristina Rossi	Preschool Nutrition and Subsequent Schooling Attainment: Longitudinal Evidence from Tanzania