Benefits of Enhanced Access to Education in Tanzania

by

Livini Donath, Oliver Morrissey and Trudy Owens

JEL Classification: I21, O55

Keywords: Universal Primary Education, youth, household welfare, Tanzania.

Abstract

Motivation: Expanding access to education is a desirable policy. Tanzania implemented Universal Primary Education (UPE) in 2001 and expanded secondary in 2006. However, evidence that UPE delivers benefits in educational attainment or earnings is scarce. We investigate if the enhanced access to education (EAE) in the 2000s benefitted the youth able to complete more schooling, those aged 15-25 in 2018.

Purpose: There is evidence that expansion of primary education in Tanzania in the 1970s increased the incomes of individuals who benefitted. We compare households headed by youth aged 15-25 and by adults (aged over 35) in 2001 and 2018 to identify benefits of EAE.

Approach and methods: Household welfare is measured as consumption per adult equivalent relative to the national poverty line using Household Budget Survey data for 2001 and 2018. We examine whether welfare differences comparing youth headed and adult headed households in 2001 and 2018 are attributable to differences in educational attainment of youth due to EAE using regression analysis and decomposition methods.

Findings: The increase in youth educational attainment by 2018 is a significant factor explaining the increase in welfare of youth headed households between 2001 and 2018. If the youth in 2001 had the same education endowment as their 2018 counterparts, their relative welfare would have been about a quarter higher. If adults had the same level of educational attainment as the youth, their welfare would have been about a third higher in 2018.

Policy implications: Expanding access to education via EAE had a positive effect on the welfare of youth, represented by those who were heads of households, in Tanzania due to increased years of schooling, providing support for implementing policies that improve access to schooling. Benefits are greatest for younger youth, aged 19-24 in 2018. Returns to schooling declined, suggesting that growth in demand for skilled labour did not match the supply of educated entrants to the labour force.

1. Introduction

The nationwide free and Universal Primary Education (UPE) of 2001 in Tanzania and the secondary school expansion program (known as 'ward secondary schools') that began in 2006 are the two most significant reforms to educational access in Tanzania in over 30 years; we refer to both as enhanced access to education (EAE) for convenience. Primary school gross enrolment increased from 84% in 2001 to 109.9% (implying ten per cent are above primary school age) in 2005 as the number of primary schools increased from 11,873 to 14,257 (United Republic of Tanzania [URT], 2005). Secondary school gross enrolment was much lower but increased from 20% in 2006 to 37% in 2012, with an almost doubling in the number of secondary schools (URT, 2008, 2013, 2016). The merit of expanding and easing access to education is an established policy aim (Colclough et al., 2003), and UPE has been championed by UNESCO in the *Education for All Global Monitoring Report* since 2002 with Chris Colclough as founding Director. Delesalle (2021) found that the earlier Tanzanian UPE reforms in the 1970s increased earnings in 2002 of heads of households likely to have benefited. This paper investigates if the more extensive enhanced access to education (EAE) in 2001 and 2006, UPE plus the expansion of secondary schools, increased incomes of the individuals likely to have benefited.

This period of expanded access to education was also a period of improvements in household welfare. Poverty incidence according to national poverty lines declined from 36% in 2001 to 26% in 2018 (URT, 2002, 2019). Did the expansion of access to education in the early 2000s contribute to the improvement in welfare by 2018? To address this important question on the impact of EAE involves comparing a set of individuals who benefitted from the reforms to a set who did not, a difficult challenge given the limited data available for earnings of individuals. As a proxy to compare youth income before the reforms with youth income after, the approach adopted here compares the welfare of households according to the age cohort of the head using the 2001 and 2018 Tanzania Household Budget Surveys (HBS). The HBS are chosen because they have relatively good data for nationally representative samples of households that allow us to compare welfare of youth and adults both before and after EAE (22,022 in 2001 and 9,552 in 2018 for the analysis).

The primary focus is to examine the extent to which differences in educational attainment between two cohorts explain the welfare difference between the two periods: the youth cohort aged between 15 and 25 in 2001, who would have completed primary school before the UPE reforms, and the corresponding cohort in 2018, who should have benefited from the reforms.[1] This is a narrower definition of youth than the official Tanzanian definition of those aged 15 to 35 but is chosen to restrict analysis to those who could have derived full benefit from the 2001 reforms. In addition to increased educational attainment for the youth, the effect on returns to education (at a household level) is considered. These two effects are not the same: beneficiaries of EAE should be more educated and therefore have higher earnings, but as the number of educated workers increases the returns to a given level of education may decline. Using data from the HBS, the analysis compares youth headed households in 2001 and 2018 and compares youth to adult headed households in both years. As consumption in HBS is measured at the household level, we assign it to the head of the household and thus comparison is between

1 The classification of youth begins at 15 (those under 15 are children and omitted from the analysis). In Tanzania, students typically have seven years of primary school beginning at age 7 so the age group 15-25 in 2018 are considered as the 'treated' group. As some youth aged over 25 in 2018 may have benefitted, by repeating a year or two of primary school or from secondary school expansion, Donath et al. (2022, Appendix B) present analysis for broad youth (15-35) and the older youth (26-35), capturing those who could have only benefited from secondary expansion, to compare with the analysis for younger youth (15-25) here. The analysis shows that the welfare gains from education are greatest for younger youth.

groups of households distinguished by the age of the head and education is the household head's level of education measured in years. While the HBS are the only currently available data providing sufficient samples to compare youth and adults in 2018 and 2001, restricting the focus to welfare (in terms of consumption expenditure) of households imposes limitations: only education of the head is considered; it is not possible to incorporate other household members or estimate individual returns to education (available data on individual earnings are inadequate for 2001); and it is not feasible to estimate separately for more disaggregated categories than youth versus adults (such as gender of the head, sector of employment or narrow age ranges), as the sub-samples are too small. Furthermore, we can only partly capture the effects of secondary expansion as many beneficiaries may still be in education in 2018.

Delesalle (2021) also classifies for household heads in a comprehensive study using national household surveys to address the effect of education expansion on household welfare in Tanzania. She considers the UPE programme of 1974-78 aimed at levelling access to education across districts so expansion was concentrated in districts that were initially disadvantaged. As the timing and scale of that expansion varied across districts depending on initial status, with little or no expansion in districts initially well provided with schools, the potential benefit varied according to distrct. Identification is thus based on age and location at the time of reforms as determining the potential treatment effect, exploiting variation in the scale of expansion over time and residence district. She estimates that predicted consumption in 2002 of household heads exposed to 1974-78 programmes was increased by six per cent. Although we also consider effects for household heads, that identification strategy is not feasible as UPE 2001 did not vary intentionally in timing or intensity across districts. However, rather than using methods to predict consumption our approach avails of household consumption reported in the HBS.

The paper makes two contributions to the literature on the link between education and household welfare. First, studies of trends in welfare in Tanzania have focussed on differences according to gender, employment, and rural/urban categories (Belghith et al, 2020; Khan & Morrissey, 2020), and less attention has been paid to age groups (specifically, youth versus adults). This paper examines how the expansion of education under EAE between 2001 and 2018 is associated with household welfare changes for youth (aged 15-25) compared to adult headed households over the period. Secondly, this paper examines at the household head level the attribution of welfare differences between 2001 and 2018 to changes in years of schooling completed (endowments) and to changes in returns to education. As EAE increases the average education level of entrants to the labour market there may be effects on returns to education.

To separate the effect of endowments and returns, the analysis employs decomposition analysis using the recentered influence function (RIF) to show if differences in educational attainment between youth and adults significantly explain the difference in welfare between the two groups in both years. Agyire-Tettey et al (2018), using decomposition analysis for Ghana, found that differences in educational attainment significantly explained the welfare gaps between rural and urban households. We find it is relevant for gaps between youth and adult headed households. If adults had the same level of educational attainment as the youth, their welfare would have been about a third higher in 2001 and 2018. The findings also suggest that if the youth in 2001 had the same education endowment as their 2018 counterparts, their welfare would have been about a quarter higher. Although there appears to have been a decline in returns to education for the youth, consistent with entrants to the labour force being more educated, this adverse effect was more than compensated by increased endowments.

Section 2 reviews selected related literature, with a focus on Tanzania, followed by an outline of the empirical method in Section 3. Section 4 describes the data used in our analysis and provides descriptive statistics for the main variables. Section 5 presents the results and discussion, and Section 6 concludes.

2. Education, Earnings and Household Welfare

The literature on the effects of UPE tends to focus on educational outcomes and finds that even if enrolment increases, UPE has little if any effect on educational attainment, and that public education expenditure is not clearly correlated with educational outcomes (Al-Samarrai, 2006). For example, Kan & Klasen (2021), using data from household surveys over 2005 to 2014, find no evidence that Uganda's free primary education (the main component of UPE) increased the number of years in, or likelihood of completing, primary school, although there was some improvement in progression to secondary school. These studies do not address effects on earnings or households.

Delesalle (2021) used the first three waves (2008/09, 2010/11 and 2012/13) of the Tanzania National Panel Survey (TNPS) in combination with the 2002 Tanzania Population and Housing Census to estimate the effect of the UPE programme of 1974-78 on consumption in 2002. The estimated effect of covariates of consumption in TNPS are matched with the census to generate predicted consumption in 2002. The UPE increased years of and returns from education, with the largest earnings benefits for wage labour and self-employed, so 'that the returns to [an additional year of] education for the entire Tanzanian population are about 6 per cent' (Delesalle, 2021, p2). Khan & Morrissey (2020) also use data from the first three waves of the TNPS and, although the focus is on income diversification (the number and different types of work household members engage in), find that education does facilitate access to higher earning work: an extra year of education for the household head is associated with about 1.2% higher level of consumption. These are much lower estimates than Delesalle (2021) but this is unsurprising given different time periods, estimation strategies, samples, and dependent variables.

A larger literature investigates the effect of education on employment and earnings. Most evidence of positive returns to education is based on data on earnings of individuals, and estimates vary considerably across studies (Psacharopoulos & Patrinos, 2018). Typically, wage employees have higher returns than the self-employed and agricultural workers, as observed in Tanzania (Al-Samarrai & Reilly, 2008). Serneels et al (2017) found that estimated returns in Tanzania for a year of post-primary school ranged between zero and 20% for men and 30-50% for women. Donath et al (2021, pp 18-23) note that returns per year of schooling vary according to the pay period: workers paid monthly, likely to be in wage employment, have positive returns after two years of education that then increase steadily, whereas workers paid daily or weekly require up to six years of school before returns are positive and although returns then increase steadily they remain at least ten percentage points lower than for workers paid monthly. However, available surveys on earnings are not suitable for our purpose of assessing the impact of EAE due to absence of data for the early 2000s to compare with recent estimates of returns (see Donath et al, 2022, Appendix A).

Some studies consider the household level and find that a measure of household education (often for the head or a selected member) is positively associated with household welfare although the effects are heterogeneous. For example, Arsalan et al (2021) found that an increase in the number of working-age household members with secondary schooling was associated with higher household expenditure and lower poverty. Recent studies on Tanzania have addressed the association between education and household welfare. Belghith et al (2020) note the increase in primary and secondary enrolment between 2012 and 2018, and employed Oaxaca-Blinder decomposition to examine how much poverty reduction can be attributed to changes in the endowments of household characteristics and the amount due to changes in the returns to these characteristics using data from HBS 2012 and 2018. The findings suggest that gains in education have benefited the better-off more than the poor and that the returns to education, while increased for the better-off, significantly declined for the poor. They observe

that while education is positively associated with earnings and household welfare the expansion in access and attainment means 'that the rewards for years of schooling below a certain level have declined [so that] the gains in income, and consumption, associated with primary education have become minimal' (Belghith et al, 2020, p8).

3. Empirical Strategy

To estimate the contribution of education to welfare over time and between groups the decomposition methodology based on the seminal work of Oaxaca (1973) and Blinder (1973) as extended by Firpo et al (2009, 2018) and Fortin et al (2011) is employed, specifically the recentered influence function (RIF). As noted above, available data do not permit estimating returns to education for individuals or applying the identification strategy of Delesalle (2021). As the structure of the data permit a comparison of youth and adult headed households in two periods, a decomposition method based on the characteristics of each household type is appropriate as it allows comparison of types of households 'as if' they had features (characteristics and returns on those characteristics) of another type. The empirical methodology follows Firpo et al (2009, 2018) RIF based decomposition for the mean difference between two groups, such as youth and adults. For a given dependent variable y and independent variables x, RIF decomposition uses RIF regression in combination with reweighting to decompose the overall effect (the difference between y for group 1 and y for group 2) into the difference due to endowments (characteristics x or composition effect) and the difference attributed to the relationship between y and x (coefficient effect or return effect). The method also estimates the contribution of each explanatory variable on the two components of the overall effect.

The baseline regression for the decomposition is a standard household consumption model:

$$\ln C_{it} = \boldsymbol{\alpha} \boldsymbol{s}_{it} + \boldsymbol{\beta} \boldsymbol{x}_{it} + \varepsilon_{it} \qquad (1)$$

Where C is the household consumption to poverty line ratio (CPL), our preferred welfare measure to account for the price differences (inflation) between surveys given the absence of good price deflators; $\boldsymbol{s}$ a vector of schooling of the household head and its square (in years); $\boldsymbol{x}$ is a vector (including a constant) of individual/household characteristics; $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are regression parameters; ε is a standard error term; and i and t index individual and time, respectively. As the decomposition method is not suited to address endogeneity, (1) is estimated using OLS and the RIF decomposition reweights the coefficient estimates to separate the two components; our estimation follows the procedure of Rios-Avila (2020) and a more detailed exposition can be found in Donath et al (2022).

Our interest is in estimating how the vectors of covariates ($\boldsymbol{s}$ and $\boldsymbol{x}$), specifically years of education, contribute to difference in welfare between the two groups (youth and adult headed households) and the two years (2001 and 2018). This involves establishing the counterfactual; for example, what would the welfare of youth headed households in 2001 (or adult headed households in 2018) have been if they had the characteristics of youth headed households in 2018? Specifically, how important is education in explaining the difference? As the counterfactual cannot be observed, it is approximated using the conditional probabilities derived from estimating (1) and reweighting these by the sample shares in different groups. For example, knowing the endowments (values of variables) and contribution of factors to youth welfare in 2001 and 2018, one can estimate the probability that a household is in the 2001 group given the observed household characteristics and similarly for 2018 (or for youth versus adult headed households). The effect of each variable can then be reweighted by the group's share of the sample to obtain the outcome for that group if it had the characteristics of another group.

The overall effects can then be decomposed into two main components, each with two sub-components. The first component is called the (aggregate) composition effect and it consists of

the pure composition (endowment) effect and the specification error. The second is called the coefficient effect and consists of the pure coefficient effect and the reweighting error. The pure effects are our main focus. The error components are used to assess the quality of the reweighting and specification of the regression function: smaller and insignificant errors indicate more robust results (Firpo et al, 2018). The RIF decomposition for the mean of log consumption relative to the poverty line (CPL) is estimated in Stata using the command *Oaxaca_rif* (Rios-Avila, 2020).

4. Data and Descriptive Statistics

The Tanzanian Household Budget Surveys (HBS) for 2001 and 2018, conducted by the National Bureau of Statistics, are among the most extensive household surveys for Tanzania Mainland (omitting Zanzibar). Data collection for HBS 2001 took place from May 2000 to June 2001 and for HBS 2018 from December 2017 to November 2018. Both surveys employed a multi-stage cluster sampling to obtain representative samples of 22,176 and 9,552 households in 2001 and 2018, respectively. Despite the sample for 2018 being significantly smaller than its 2001 counterpart, the sampling mechanism still ensured representativeness at the national (Mainland) level (URT, 2019). A total of 154 households in 2001 had missing information on assets ownership and were excluded from the analysis, leaving us with a sample of 22,022 households. All households in 2018 had complete information.

Households are categorised according to the head's age group: youth and adults. Youth households include all households headed by a youth aged between 15 and 25, who may have benefitted from EAE by 2018, and adult households include all households headed by an adult aged over 35 (who won't have benefitted from EAE). The welfare indicator is measured at the household level as the ratio of the per adult equivalent household consumption to the national poverty line (CPL), both as reported in the respective HBS (URT, 2002; URT, 2019), and is assigned to the head of household.[2] Given the absence of good price deflators covering 2001 to 2018, especially to capture spatial price variation, the ratio is an appropriate if simple way to represent the relative welfare of the household at the time of the survey and helps to account for inflation and trends in earnings between the surveys for comparing welfare in 2001 to 2018. Overall, it is consistent with the trends in poverty: the ratio of per adult household consumption to the national poverty line improved on average from 1.79 in 2001 to 2.28 in 2018, equivalent to a 27% increase similar to the percentage decline in poverty (URT, 2002, 2019). An implication is that our measure is relative – the comparisons are for changes in relative welfare of youth headed and adult headed households.

Education is the household head's level of education measured in years (the variable in the HBS). This is appropriate as our comparison is between groups of households distinguished by the age of the head – youth who benefited from EAE by 2018 and adults who didn't. However, endogeneity of education is a concern to the extent that factors associated with higher welfare,

2 We use the official per adult equivalence scale (AES) reported in URT (2009, p82), which attaches lower weights to infants and children to adjust for household composition and is used to construct the national poverty line. Consumption includes expenditure on food and regular non-food items but excludes durables and infrequent purchases. A specific concern is that the official AES may understate the welfare of female-headed households and households with relatively many adult females because females aged 15 and over are given a lower weight than males of the same age. Khan & Morrissey (2020) show for Tanzania over 2008-13 that measuring household welfare using equivalence scales with higher weights on adult females gives almost identical results to using the official AES for determinants of household consumption, so using AES should give us reliable results.

such as education, may also be associated with a higher likelihood of youth to form a household. As methods to incorporate such endogeneity in a RIF decomposition have not been established (Firpo et al, 2018), Donath et al (2022, Section 5.2) estimate determinants of a youth aged 15-25 being a head of household and find that more educated youth were more likely to be a household head.[3] Nevertheless, the size effect is negligible as an additional year of education increases the probability of becoming head of a household by no more than 0.5%. As mean years of education for youth heads are relatively low (six years in 2001 and seven in 2018), such a small effect is unlikely to alter our main results.

Section 4.1 provides definitions of the main variables used in the analysis, followed by Section 4.2 with descriptive statistics comparing youth and adult headed households in 2001 and 2018. In addition to the analysis here for younger youth aged 15-25, Donath et al (2022, Appendix B) provide estimates for the broad 15-35 and older 26-35 youth age groups. The findings are similar but the benefits for broad and older youth from EAE are somewhat smaller than for younger youth, consistent with those aged over 26 including many who could not have benefited from EAE. The relevant result is that gains in schooling are greatest for younger youth able to continue to secondary school (even if not completed): compared to the 26-35 group, the 15-25 group has higher returns after eight years (completed primary) and the gap increases after 11 years (completed secondary).

4.1. Definition of the Main Variables

All variables are taken from the HBS (URT, 2002, 2019). Basic definitions for the variables are listed below.

Household characteristics

- **CPL:** the ratio of household consumption expenditure per adult equivalent to the poverty line, both based on official AES, in logarithm form (to adjust for skewness of income)
- **poor:** = 1 for households below the basic needs poverty line (0 otherwise)
- **rural:** = 1 for households in rural area (0 otherwise)
- **hhsize:** total number of usual members in the household

Household head characteristics

- **education:** years of schooling of the household head
- **noeducation:** = 1 if household head completed less than three years of primary education (0 otherwise)
- **someprimary:** = 1 if head completed 4-6 years primary (0 otherwise)
- **primary:** = 1 if head completed the seven years of primary (0 otherwise)
- **somesecondary:** = 1 if head completed 2-3 years of secondary education (0 otherwise)
- **secondary:** = 1 if head completed the four years of lower secondary education (0 otherwise)
- **postsecondary:** = 1 if head has more than lower secondary education (0 otherwise)
- **age:** age of the household head in years
- **female:** = 1 if the head is female (0 otherwise)
- **married:** = 1 if the head is married (0 otherwise)

4.2. Descriptive Statistics

Table 1 shows the characteristics of youth headed households (Youth, aged 15-25) compared to adult headed households (Adult) for each year and compares the characteristics of youth

3 Income appears more relevant as wage earners are more likely to be a household head than those in agriculture. This is not independent of schooling as educated youth are more likely to be in non-agricultural self or wage employment, which are associated with higher household welfare (Khan & Morrissey, 2020).

household heads (Head) to all other youth in the sample (Other). The share of households headed by a youth in the total sample decreased by a third from 7.5% in 2001 to 5% in 2018 (there were 473 youth headed households in 2018). This could be due to multiple factors (including changes in regional sampling as a larger share of the smaller sample is urban in 2018). In both years youth headed households had significantly higher welfare (CPL), were less likely to be poor, and were much smaller than adult headed households. One would expect adult-headed households to be larger as they are older and in general larger households are poorer, so household size is included as a control to account for additional size effects in the analysis. The CPL increased and the proportion poor and household size decreased for both groups in 2018, reflecting the general reduction in poverty compared to 2001. The differences declined: youth headed households had relative welfare (CPL) about 30% higher than adult headed in 2001 but this fell to about 20% in 2018. This is consistent with youths who are not heads being more educated on average than heads in 2018 (less educated youth appear more likely to be heads even if more schooling is associated with being a youth head), and potentially having higher earnings (reflected in the welfare of their household).

Table 1: about here

Compared to youths that were not heads, a much higher proportion of youth heads were married, although more than a quarter were not, and they were almost four years older on average than other youth. A quarter of youth heads were female in 2001, declining to about a fifth in 2018, and a significant share of other youth were married (many would be the spouses of youth heads). Although youth heads had slightly more years of education in 2001, this was reversed by 2018: education of youth heads increased by about one year on average compared to about two years on average for other youth. In 2001 youth heads were more likely to have completed primary but also more likely to have no education; by 2018 the proportion of youth heads with post primary education was lower than for other youth. This is consistent with a tendency for more educated individuals to delay forming households. However, many factors interact to motivate household formation and education is only one of these (and is not very significant so unlikely to confound our results, as shown in Donath et al, 2022).

Figures 1 and 2: about here

Donath et al (2022, Table 2) provide comparable data for adult headed households, showing that educational attainment and household welfare were higher 2018, whereas household size, the proportion rural and in poverty all fell. Youth had more years of education than adults in both years – 6 compared to 4 in 2001 and 7 compared to less than 6 in 2018. Analysis of years of education by age in 2018 in Figure 1 confirms that (i) benefits of EAE are greatest (over 7 years of school) for the 19-24 age group (and up to age 29); and (ii) the 15-18 age group have less than 7 years education on average although many may still be in education and very few are heads (overall less than 1% of heads are still in education). The density plots for household CPL in Figure 2 show that youth headed households have higher CPL and lower poverty rates than adult headed households in both years. The distribution for the youth (15-25) is to the right of that for adults and the proportion of youth households above the poverty line (logCPL>0) is greater, especially in 2018.

5. Education and Youth Welfare: Results and Discussion

The core question is whether returns to education for youth (15-25), in terms of household welfare, are different from those of adults; and whether they have changed between 2001 and 2018. Table 2 presents the results for CPL for each year based on OLS regression estimates of equation (1) – returns to education are therefore in terms of relative household welfare, not individual earnings. Years of schooling (*Sch*) of the household head is positively and significantly correlated with welfare for adult headed households only, but Sch^2 is positive and significant for youth and adults, in both years. The significance of schooling squared implies a strong convex relationship between education and welfare – each extra year of schooling is associated with higher welfare than the previous year – and this is stronger for adults but slightly weaker for youth in 2018.

Table 2 and Figure 3: about here

The presence of Sch^2 complicates the interpretation of the coefficients of schooling variables so we plot the implied welfare returns to education from Table 2 in Figure 3.[4] In 2001, youth headed households had lower returns up to 12 years education (completing lower secondary) and beyond this adult headed households had lower returns. By 2018 returns become positive for youth (adults) after about four (three) years of school; returns are slightly higher for adults up to 10 years of education, after which returns are higher with an increasing gap for youth.[5] Most other included regressors have the expected sign; coefficients on *Age* and Age^2 are only clearly significant for adults in 2001(implying a threshold age before earnings rise) and *Married* is insignificant for youth in 2018. In contrast, coefficients are positive and significant for *Female* (except adults in 2001), but negative and significant for rural and household size (both associated with lower welfare, more pronounced in 2018).[6]

Education and Welfare Changes

The factors contributing to differences in mean household welfare between youth and adult headed households for each year are investigated using RIF regressions to decompose into two parts as explained in Section 3: the part due to differences in characteristics or endowment (the explained part) and that due to differences in returns to these characteristics (the unexplained part). Each of the two parts are broken down: the explained part into pure explained and the specification error; and the unexplained part into pure unexplained and reweighting error.

Table 3 presents the decomposition results by year. To simplify interpretation, the coefficients

4 Donath et al (2021, Tables 11 & 14), using data on some 8000 individuals over 2008-2016, report results for *Sch* and Sch^2 consistent with those in Table 2 and show that completed level and years of education give qualitatively consistent results and returns for secondary are considerably higher than for primary, as illustrated in Figure 3. AME (*Sch*) in Table 2 shows a decline in (household) returns to education in 2018 compared to 2001(consistent with entrants to the labour force being more educated) – returns to primary education almost halved whereas returns to completing secondary fell by about a quarter.

5 The pattern is different for broader youth (15-35): in 2001 returns to adults fell below youth after 8 years of education, whereas in 2018 returns fell for both groups and were slightly higher for adults (Donath et al, 2022, Appendix Figure B1). This supports the view that younger youth (15-25) benefited most from EAE.

6 Using HBS, we can only identify the activity sector of the household head at a general level (agriculture, self-employed, wage) and separate estimates by sector of activity, gender of head or narrower age ranges are not feasible owing to the small sample sizes (especially in 2018). We do control for rural (versus urban) location to partly captures effects insofar as agriculture is primarily rural and wage employment primarily urban.

of the education variables (*sch* and sch^2) are aggregated into one variable 'education'; the coefficient of *age* and age^2 into 'headage'; and ownership of assets, livestock per capita, and dummies for regions of residence into 'other controls'.[7] The top panel of Table 3 shows the contribution of the explained and unexplained parts to the total difference in log welfare. Only the explained component is significant in both years (for the reweighted model) implying that it is only the difference in endowment that explains differences in welfare between the two age groups. Both the specification and reweighting errors are insignificant, suggesting that the model is correctly specified and reweighed. Although the reweighting error is insignificant, it is large relative to the raw difference and the 2001 counterfactual is not significantly different from zero. To assess if the reweighting has an undue effect on the estimates, unweighted results are also reported (in this case there is no counterfactual and the explained part is insignificant). The coefficients for the explained and unexplained parts are much larger, consistent with the unweighted model not accounting for the effects of specification and reweight errors, and although the size and in some cases significance of variables alters, the education variable remains significant.

The breakdown of the 'pure explained' component reveals that the coefficient on education is positive and significant (although smaller and weakly significant in the unweighted model), confirming that the youths heading households have significantly better education attainment than adults heading households. About a third of the pure explained welfare difference is attributable to differences in educational attainment between youths and adults (the remainder is largely attributable to differences in household size). In other words, if an adult had the same level of educational attainment as a youth heading a household, their welfare would have been about 35% (0.101/0.292) higher in 2001 and 34% higher in 2018. Household size is the only other characteristic that has a consistently significant contribution to the explained difference. These results are consistent with Table 2. The difference in returns to education (unexplained), however, does not have a significant effect on welfare (also consistent with Table 2 and Figure 2). No variables are significant for the unexplained difference.

Table 3: about here

Table 4 provides the RIF decomposition for the youth heading households between 2001 and 2018, reported by gender and rural/urban residence. Again, both the specification and reweighting errors are insignificant, and in this case the reweighting error is not so large relative to the raw difference while the counterfactuals are significantly different from zero so unweighted results are not reported (but yield similar coefficients for the explained and unexplained parts). We can have more confidence in the reweighted model for decomposing effects over time for youth headed households than for the comparison with adult headed households.

The results for the pooled model in Table 4 suggest that the difference in welfare between the two periods is attributed to differences in characteristics (the explained part) – other controls (such as assets) are the main factor but education accounts for about 25% of the explained difference while household size is slightly more important for male heads – and coefficients are

7 The Stata command *oaxaca_rif* is calibrated for this common approach in decomposition (and the aggregations fit with the specification).

insignificant.[8] Although the raw difference is at best weakly significant, the counterfactual is relatively large. The share of education in the explained part implies that if the youth in 2001 had the same education endowment as their 2018 counterparts, their welfare would have been about 25% higher (slightly lower for males).

Table 4: about here

There is evidence of heterogeneity by place of residence: education (and other controls to a lesser extent) is less important for rural than urban households; whereas education accounts for about 37% of the explained difference in urban areas it only accounts for 22% of the difference in rural areas. This is consistent with education endowments being higher in urban areas and the greater association between education and welfare for urban households is consistent with the greater availability of wage employment. In rural areas, in contrast, education endowments are lower and there is less wage employment – a greater share of employment is in lower skilled agriculture. No variables are significant for the unexplained part (in the reweighted model), implying that the decline in returns to education had no detrimental effect on welfare.

6 Conclusions

This paper focuses on investigating how much of the welfare difference between households headed by youth (aged 15-25) in 2001 and 2018 and the difference between youth and adult headed households in each of the years can be attributed to changes in educational attainment and returns to education. The aim is to assess the impact of increased participation in education following EAE, Universal Primary Education introduced in 2001 and extended secondary education in 2006, which mainly benefited the youth aged 15–25 years in 2018. The study examines both how much of the welfare differences between 2001 and 2018 can be attributed to changes in the association between education and welfare over the period, and how much can be attributed to changes in educational attainment. Samples of household heads from the 2001 and 2018 HBS are investigated using RIF decomposition of the mean.

The analysis decomposes the welfare differences between cohorts into the part attributable to differences in characteristics (education is the focus) and the part attributable to returns to these characteristics (again the focus is on returns to education). The decomposition of welfare between youth in 2001 (pre reform) and in 2018 (post reform) sheds light on the effect of the large expansion of education. If a significant part of the differences in welfare between the two youth cohorts can be attributed to the differences in educational attainment, there is evidence that EAE improved welfare through increasing educational attainment.

Proxying welfare by household (per adult equivalent) consumption expenditure relative to the national poverty line, the analysis shows that youths, having more education, enjoy higher welfare levels (relative to the poverty line) than adults in both years; if adults had the same education as youth their welfare would be about a third higher. Comparing youth cohorts across years, the youth in 2018 have higher education and welfare levels than their 2001 counterparts; increased educational attainment between 2001 and 2018 can explain 24% of the increase in welfare of youth headed households. Despite the decline in returns to education, we find no evidence that this reduced welfare – the adverse effect was more than offset by increased years

8 Data are too limited to capture gender differences as there are too few female-headed youth households, especially in 2018. Nevertheless, gender is mostly insignificant in Table 3 and estimates for males only in Table 4 are not qualitatively different from the pooled sample.

of schooling. The benefits of EAE appear greatest for youth with some or completed secondary education. Although endogeneity of education or endogenous selection of youth to be heads of households are concerns, education is a minor factor determining whether a young person is a head of household and therefore this is unlikely to qualitatively alter the main finding.

The analysis is based on comparisons of youth and adult headed households, given the difficulty in measuring education and consumption of individuals at the household level to compare 2018 with 2001. As a result, we could not analyse the welfare effects for the youth who live in households headed by an adult. Small samples mean we couldn't explore effects for households at a more disaggregated level, such as by gender of the head (although it was included as a characteristic). Nevertheless, we add to the evidence in Delesalle (2021) that the early UPE in the 1970s increased incomes in 2002 by showing the the more extensive EAE in the 2000s contributed to increased income (measured by consumption spending at the household level) in 2018. The benefit of increased levels of education (endowments) offset the decline in returns to schooling. Addressing the low returns requires increasing the quality of education and training and expanding labour market opportunities.

Data Statement

The cleaned and constructed data used in the analysis, and the do files, are available on request.

REFERENCES

Agyire-Tettey, F., Ackah, C. G., & Asuman, D. (2018). An Unconditional Quantile Regression Based Decomposition of Spatial Welfare Inequalities in Ghana. *Journal of Development Studies,* **54**(3), 537–556. https://doi.org/10.1080/00220388.2017.1296571

Al-Samarrai, S. (2006), Achieving education for all: How much does money matter? *Journal of International Development*, 18 (2), 179–206. https://doi.org/10.1002/jid.1190

Al-Samarrai, S., & Reilly, B. (2008). Education, employment and earnings of secondary school and university leavers in Tanzania: Evidence from a tracer study. *Journal of Development Studies*, **44**(2), 258–288. https://doi.org/10.1080/00220380701789950

Arsalan, A., Tschirley, D., & Egger, E.-M. (2021). Rural Youth Welfare along the Rural-urban Gradient: An Empirical Analysis across Developing Countries. *Journal of Development Studies*, **57** (4), 544-570. https://doi.org/10.1080/00220388.2020.1808197

Belghith, N. B. H., Karamba, W., Talbert, E., & de Boisseson, P. (2020). *Tanzania Mainland Poverty Assessment: Tanzania's Path to Poverty Reduction and Pro-Poor Growth*. The World Bank. http://documents.worldbank.org/curated/en/254411585030305188/pdf/Part-1-Path-to-Poverty-Reduction-and-Pro-Poor-Growth.pdf

Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources,* **8**(4), 436–455. https://doi.org/10.2307/144855

Colclough, C., Al-Samarrai, S., Rose, P. & Tembon, M. (2003). *Achieving Schooling For All in Africa: Costs, Commitment and Gender*, Aldershot: Ashgate. http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?t2000=020814/(100)

Delesalle, E. (2021). The effect of the Universal Primary Education program on consumption and on the employment sector: Evidence from Tanzania. *World Development,* **142**, 105345. https://doi.org/10.1016/j.worlddev.2020.105345

Donath, L., Morrissey, O., & Owens, T. (2021). Does the Pay Period Matter in Estimating Returns to Schooling? Evidence from East Africa, School of Economics, University of Nottingham: *CREDIT Research Paper 21/01*. https://www.nottingham.ac.uk/credit/documents/papers/2021/21-01.pdf

Donath, L., Morrissey, O., & Owens, T. (2022). Universal Primary Education and Household Welfare in Tanzania, School of Economics, University of Nottingham: *CREDIT Research Paper 22/02*. https://www.nottingham.ac.uk/credit/news/papers/2202.aspx

Firpo, S. P., Fortin, N. M., & Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, **77**(3), 953–973. https://doi.org/10.3982/ecta6822

Firpo, S. P., Fortin, N. M., & Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics,* **6**(2), 1–41. https://doi.org/10.3390/econometrics6020028

Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition Methods in Economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics Volume 4* (pp. 1–102). Amsterdam: Elsevier https://doi.org/10.1016/S0169-7218(11)00407-2

Kan, S., & Klasen , S. (2021), Evaluating universal primary education in Uganda: School fee abolition and educational outcomes, *Review of Development Economics*, **25**, 116–147. https://doi.org/10.1111/rode.12725

Khan, R., & Morrissey, O. (2020). Income diversification and household welfare in Tanzania 2008 – 13, *WIDER Working Papers 2020/110*, Helsinki: UNU-WIDER, https://www.wider.unu.edu/sites/default/files/Publications/Working-paper/PDF/wp2020-110.pdf

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, **26**(5), 445–458. https://doi.org/10.1080/09645292.2018.1484426

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review,* **14**(3), 693–709. http://dx.doi.org/10.2307/2525981

Rios-Avila, F. (2020). Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *Stata Journal,* **20**(1), 51–94. https://doi.org/10.1177/1536867X20909690

Serneels, P., Beegle, K., & Dillon, A. (2017). Do returns to education depend on how and whom you ask? *Economics of Education Review*, **60**, 5–19. https://doi.org/10.1016/j.econedurev.2017.07.010

URT. (2002). *Household Budget Survey 2000/01*. National Bureau of Statistics, Dar es Salaam. https://www.nbs.go.tz/nbs/takwimu/hbs/hbs2000-01/Key_Findings_HBS_2001_Eng.pdf

URT. (2005). *Basic Education Statistics in Tanzania (BEST) 1995 - 2005*. Ministry of Education and Culture, Dar es Salaam. http://tzonline.org/pdf/basiceducationstaticsbest199520051.pdf

URT. (2007). *National Youth Development Policy*. Ministry of Labour, Employment and Youth Development, Dar es Salaam. https://www.youthpolicy.org/national/Tanzania_2007_National_Youth_Policy.pdf

URT. (2008). *The Tanzania Economic Survey 2007*. The Ministry of Finance and Economic Affairs, Dar es Salaam.

URT. (2013). *The Economic Survey 2012*. Ministry of Finance, Dar es Salaam. Accessed 10 September 2011 at http://www.tanzania.go.tz/economicsurvey1/2006/index.html (no longer available).

URT. (2015). *Tanzania Integrated Labour Force Survey 2014*. National Bureau of Statistics, Dar es Salaam. https://www.nbs.go.tz/index.php/en/census-surveys/labour-statistics/141-the-2014-integrated-labour-force-survey-ilfs-analytical-report

URT. (2016). *National Basic Education Statistics in Tanzania (BEST) 2012 - 2016*. Ministry of Education, Science and Technology, Dar es Salaam.

URT. (2018). *National population projections.* National Bureau of Statistics and Office of the Chief Government Statistician, Dar es Salaam and Zanzibar. https://www.nbs.go.tz/nbs/takwimu/census2012/Projection-Report-20132035.pdf

URT. (2019). *Household Budget Survey 2017-18: Key Indicators Report.* National Bureau of Statistics, Dar es Salaam. https://www.nbs.go.tz/index.php/en/census-surveys/poverty-indicators-statistics/household-budget-survey-hbs/413-the-2017-18-household-budget-survey-key-indicators-report

Table 1: Characteristics of Youth (15-25) Headed Households by Survey Year

Characteristics		2001			2018	
% unless stated	Youth	Adult	Difference	Youth	Adult	Difference
Households (by Head)						
CPL (ratio)	1.86	1.44	0.42***	2.27	1.90	0.37***
Poor	0.25	0.38	-0.13***	0.14	0.26	-0.12***
Rural	0.79	0.77	0.02***	0.73	0.64	0.09***
Size (number)	4.26	6.92	-2.66***	3.69	6.64	-2.95***
Youth	Head	Other	Difference	Head	Other	Difference
Education (years)	5.98	5.45	0.53***	6.85	7.15	-0.27***
No education	0.16	0.22	-0.06***	0.15	0.13	0.02***
Some primary	0.05	0.19	-0.14***	0.11	0.10	0.01***
Primary	0.76	0.54	0.22***	0.49	0.46	0.03***
Some secondary	0.01	0.02	-0.01***	0.07	0.12	-0.05***
Secondary	0.02	0.03	-0.01***	0.19	0.17	0.02***
Post-secondary	0.00	0.01	-0.01***	0.00	0.02	-0.02***
Age (years)	23.21	19.46	3.75***	23.03	19.19	3.84***
Female	0.25	0.59	-0.34***	0.19	0.55	-0.36***
Married	0.72	0.30	0.42***	0.75	0.24	0.51***
Observations	1,647	22,481	-	473	8,121	-

Notes: Author's calculations from HBS 2001 and 2018 data weighted using survey weights; mean value for continuous variables and % share of sample for binary indicators. CPL is household consumption relative to the poverty line. 'Difference' for *households* is the value for youth headed households minus the corresponding value for adult headed households; 'Difference' for *youth* characteristics are based on the average for youth heads minus the mean for other youth (with significance * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table 2: Household Returns to Education by Head Age and Survey Year

	2001		2018	
	Youth (15-25)	Adult	Youth (15-25)	Adult
Sch	-0.001	0.034***	-0.020	-0.010***
	(0.013)	(0.003)	(0.021)	(0.004)
Sch²	0.004***	0.001***	0.005***	0.004***
	(0.001)	(0.000)	(0.002)	(0.000)
Age	0.072	-0.011***	0.394*	0.002
	(0.099)	(0.003)	(0.210)	(0.003)
Age²	-0.125	0.008***	-0.870*	-0.002
	(0.227)	(0.002)	(0.475)	(0.003)
Female	0.173***	0.010	0.144**	0.054***
	(0.031)	(0.014)	(0.061)	(0.018)
Rural	-0.145***	-0.147***	-0.215***	-0.183***
	(0.032)	(0.011)	(0.064)	(0.016)
Married	0.076**	0.071***	-0.009	0.052***
	(0.034)	(0.014)	(0.063)	(0.019)
lnSize	-0.550***	-0.451***	-0.493***	-0.473***
	(0.026)	(0.008)	(0.054)	(0.011)
_cons	-0.088	1.153***	-3.577	0.656***
	(1.068)	(0.082)	(2.304)	(0.099)
Controls	Yes	Yes	Yes	Yes
AME(*Sch*)	0.048***	0.043***	0.047***	0.027***
	(0.006)	(0.001)	(0.009)	(0.002)
N	1,647	13,983	473	6,945
R^2	0.40	0.40	0.488	0.44

Notes: Estimated by OLS with household CPL as dependent variable. AME(*Sch*) is the average mean effect of a year of schooling; Household size is in logs (*lnSize*). Other controls included (not reported) are livestock per capita, region of residence and ownership of assets. Standard errors in parentheses (significance indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$).

Table 3: Reweighted and Unweighted RIF Decomposition Within Survey Years

	2001		2018	
	Weighted	Unweighted	Weighted	Unweighted
Overall				
Youth (15-25)	0.420***	0.420***	0.571***	0.571***
Counterfactual	0.067	.	0.411***	.
Adult	0.149***	0.149***	0.374***	0.374***
Difference	0.271***	0.271***	0.197***	0.197***
Explained	0.352***	4.386	0.160***	0.980
Unexplained	-0.082	-4.115	0.037	-0.783
Explained	0.292***	4.386	0.248**	0.980
education	0.101***	0.059**	0.085***	0.032*
headage	0.018**	4.065	0.006	0.837
female	0.000	0.000	-0.006	-0.004
rural	-0.002	0.001	0.007	-0.003
married	-0.000	-0.001	0.001	0.001
lnSize	0.200***	0.281***	0.292***	0.230***
Other controls	-0.025	-0.022	-0.138***	-0.105***
Unexplained	-2.298	-4.115	-8.037	-0.783
education	0.037	-0.007	0.016	0.015
age	-1.681	-1.573	-4.953	-0.865
female	0.008	0.002	0.017	0.017*
rural	-0.013	-0.018	0.014	0.006
married	-0.015	0.016	0.015	-0.021
lnSize	-0.568	-0.247**	0.344	0.063
Other controls	0.347	0.072	0.068	0.056
constant	-0.388	-2.353	-3.499	-0.005
Specification error	0.061		-0.088	
Reweight error	2.216		8.075	
N1 (youth)	1,647	1,647	473	473
N2 (adult)	13,983	13,983	6,945	6,945
N (all observations)	15,630	15,630	7,418	7,418

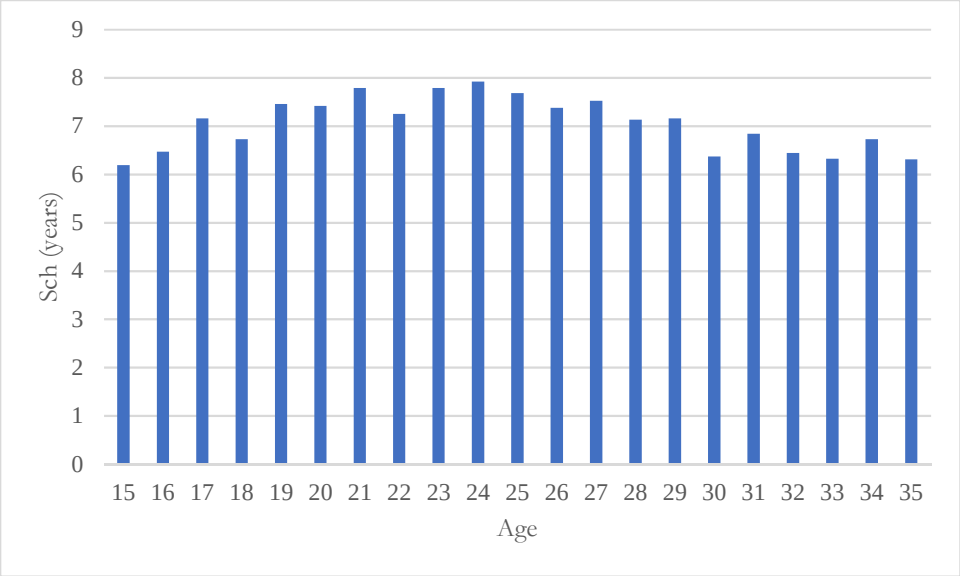
Notes: 'Other controls' is the aggregate effect of livestock per capita, region of residence and ownership of assets. Binary variables are normalised; significance * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: RIF Oaxaca-Blinder Decomposition for Youth (15-25)

	Pooled	Male	Rural	Urban
Overall				
2018	0.571***	0.535***	0.415***	0.995***
Counterfactual	0.263***	0.228***	0.211**	0.728***
2001	0.420***	0.438***	0.340***	0.722***
Difference	0.152*	0.097	0.075	0.272
Explained	0.309***	0.307***	0.204**	0.267*
Unexplained	-0.157	-0.210*	-0.128	0.005
Explained	0.396***	0.391***	0.273***	0.416**
education	0.097***	0.091**	0.061**	0.155**
age	-0.001	0.001	-0.001	0.008
female	0.004		0.008	-0.013
rural	0.013*	0.018*		
married	-0.003	-0.007	-0.000	0.008
lnhhsize	0.043*	0.050**	0.027	-0.023
Other controls	0.229***	0.227***	0.169***	0.286*
Unexplained	0.073	-0.062	0.095	-0.504
education	0.134	0.123	0.155	-0.162
age	-4.083	-0.353	-5.960	-5.172
female	0.031		0.032	-0.062
rural	-0.132	-0.138		
married	-0.074	-0.122	-0.139	0.092
lnhhsize	0.272	0.268	0.497	-0.831
Other controls	0.033	-0.076	-0.133	0.293
constant	3.924	0.184	5.727	5.336
Specification error	-0.087	-0.084	-0.069	-0.149
Reweight error	-0.230	-0.148	-0.224	0.509
N1 (2018)	473	377	333	140
N2 (2001)	1,647	1,181	608	1,039
N (all observations)	2,120	1,558	941	1,179

Notes: No separate estimates for female heads due to too few observations. 'Other controls' is the aggregate effect of livestock per capita, region of residence and ownership of assets. Binary variables are normalised; significance * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

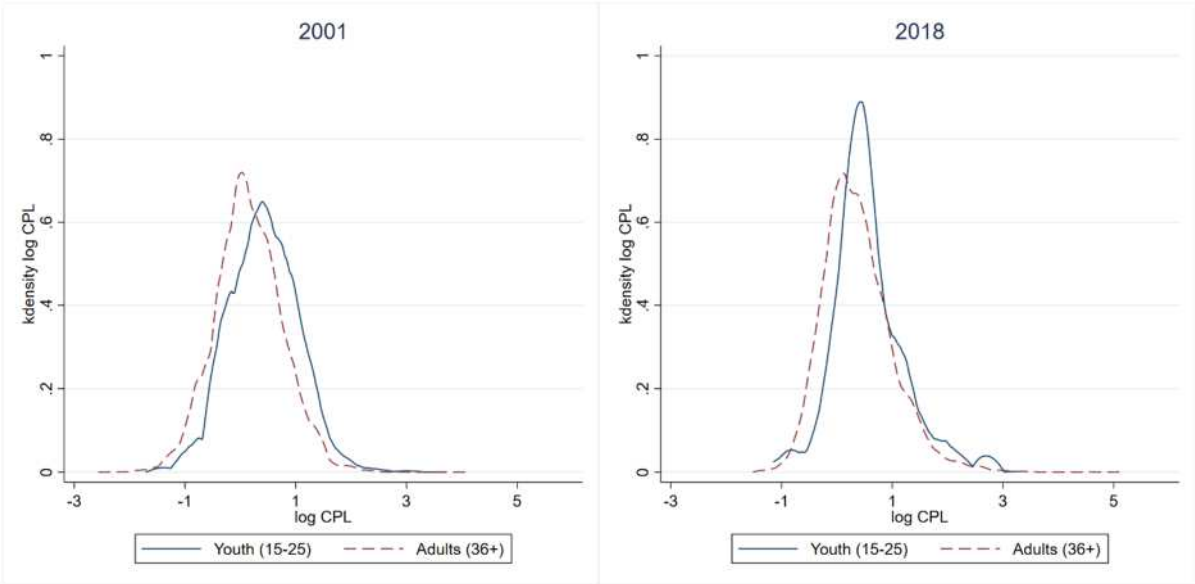
Figure 1: Youth's Years of Education by Age in 2018



Note: Survey weights applied to represent mean years of education for each age in 2018. Applies for all youth in the survey (not only household heads): aged 15-18 benefited from UPE, are still in education or left school without completing secondary; aged 19-24 benefited from UPE, could have completed secondary, may be in further education; aged 25-27 could only have benefited from secondary expansion; and aged 28-35 had no benefit from EAE.

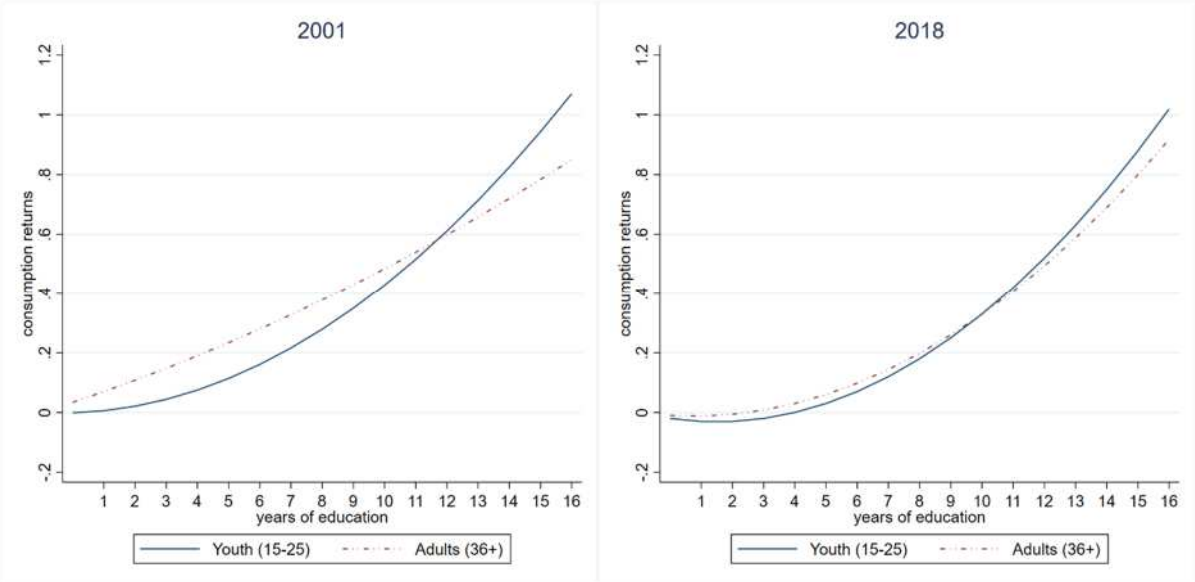
Source: Computed from sample from HBS 2018.

Figure 2: Household Consumption (CPL) by Age Group and Year



Source: Authors' computation from HBS 2001 and 2018 samples.

Figure 3: Implied Returns to Education by Age Group and Year



Source: Authors' estimates from HBS 2001 and 2018 samples.