// NO.22-022 | 06/2022

DISCUSSION PAPER

// GUIDO NEIDHÖFER, NORA LUSTIG, AND PATRICIO LARROULET

Nowcasting the Impact of COVID-19 on Education, Intergenerational Mobility and Earnings Inequality in Sub-Saharan Africa

Leibniz Association

ZEW

Nowcasting the Impact of Covid-19 on Education, Intergenerational Mobility and Earnings Inequality in Sub-Saharan Africa

Guido Neidhöfer **Nora Lustig** **Patricio Larroulet**

This version: June 2022

Abstract

Using microsimulations, we nowcast the impact of learning losses caused by COVID-19 on secondary school completion rates, intergenerational mobility of education, and long-run earnings inequality in eight countries Sub-Saharan Africa. On average, secondary school completion rates decrease by 12 percentage points overall and by 16 points for children with low-educated parents. Interestingly, in most countries the gender gap diminishes because for men the projected decrease in secondary school completion is higher. However, a small additional impact on girls' education due to the Covid-19 induced rise in teenage pregnancy is observed in some countries. Intergenerational mobility of education decreases from 1 to close to 50 percent, depending on the country. As a result of the heterogeneous reduction in average years of schooling for advantaged vs. disadvantaged children, earnings inequality could increase between one and four Gini points, depending on the assumptions.

Keywords: COVID-19, lockdowns, human capital, school closures, intergenerational persistence, education, inequality, Africa. *JEL Codes:* I24, I38, J62.

Contacts and affiliations: Guido Neidhöfer is Senior Researcher at ZEW – Leibniz Centre for European Economic Research, Mannheim (guido.neidhoefer@zew.de). Nora Lustig is Samuel Z. Stone Professor of Economics and Director of the Commitment to Equity Institute at Tulane University (nlustig@tulane.edu). Patricio Larroulet is Junior Researcher at the Commitment to Equity Institute, Tulane University.

Acknowledgements: Paper prepared for the project "Measuring Fiscal Equity in the post-COVID-19 World." The authors are grateful to the Bill & Melinda Gates Foundation for its generous support. Guido Neidhöfer acknowledges funding by the Volkswagen Foundation within the grant call "Corona Crisis and Beyond – Perspectives for Science, Scholarship and Society".

1 Introduction

In Sub-Saharan Africa (SSA) the Covid-19 pandemic resulted in a significant contraction of economic activity. After an unprecedented decline of GDP per capita of, on average, 4.5% in 2020, the region slightly recovered in 2021 and is expected to expand again in 2022 and 2023 (IMF, 2021; World Bank, 2022). However, these figures do not consider the persistence of the impact of the pandemic on some dimensions of human development and their potential long-run consequences. One of these dimensions is the effect of the pandemic on education.

As almost everywhere in the world, in SSA schools closed their doors to mitigate the spread of the virus. In most countries, they remained closed over long periods in both 2020 and 2021, adding up to a loss of more than 50% of instruction time in a number of cases (UNESCO, 2022). Available evidence for other parts of the world, suggests large negative effects of school closures on schooling achievements, particularly among children from disadvantaged background (see e.g. the reviews by Hammerstein et al., 2021; Werner and Woessmann, 2021; Zierer, 2021). The scant existing evidence for SSA countries confirms this negative effect on education, as revealed by the simulation of associated learning losses (Angrist et al., 2021; Azevedo et al., 2021), comparing data on reading pre- and post-school closures (Ardington et al., 2021), or reports on learning activity (or lack thereof) of children during school closures collected through household-level data from telephone-surveys (Dang et al, 2021).

In this paper, we use the framework proposed by Neidhöfer et al. (2021) to project the impact of Covid-19 related education disruptions in eight Sub-Saharan African countries: Ethiopia, Ghana, Kenya, Liberia, Malawi, Nigeria, South Africa, and Tanzania. We use pre-pandemic nationally representative household surveys that include information on educational attainment of individuals and their parents' educational background. Furthermore, we collect data on the duration of school closures and mitigation policies implemented by governments to support learning from home during periods of school closures. In addition, we estimate the distribution of access to the internet in each country. We begin by projecting the impact of instructional losses due to school closures on education, intergenerational mobility and secondary school completion rates for the country as a whole, by parental educational background and by gender. We then simulate the capacity of policies to mitigate the impact of school closures on education through the provision of remote learning. When estimating the differences of the impact by gender we also consider the potential increase in teenage pregnancy rates caused by the pandemic. Finally, we estimate the potential effect of the predicted changes in education on future earnings inequality.

Our main results highlight two important aspects: First, that the pandemic is likely to have a significant and persistent negative effect on educational inequality and its intergenerational persistence in Sub-Saharan Africa. Second, that the education channel will also be a significant determinant of future changes in economic inequality. These findings stress the importance of remedial actions that should be taken to prevent the prospected unprecedented decrease in human capital that could potentially offset years of favorable development in creating opportunities.

The remainder of the paper is structured as follows: Section 2 presents the methodology adopted to estimate the impact of the pandemic on education. Section 3 briefly describes the data. Section 4 reports our results on the impact of Covid-19 on education in Sub-Saharan Africa, while highlighting heterogeneities by parental background, gender and presenting estimates on the additional effect of increasing teenage pregnancy due to the pandemic on education. Section 5 evaluates the mitigation effect of policies. Section 6 reports our projections for long-run inequality. Section 7 concludes.

2 Simulation of the impact of Covid-19 on educational achievements and intergenerational persistence

We follow the simulation methodology proposed by Neidhöfer et al. (2021) to nowcast the impact of instructional losses on educational achievements and the intergenerational mobility of education through a counterfactual exercise.[1] The approach builds on measuring the amount of instructional time lost due to the closure of educational facilities (see e.g. Abadzi, 2009; Adda, 2016) and considers asymmetries in the response of countries and in the capabilities of families to cushion instructional disruptions. As done elsewhere in the literature, we assume that there is a human capital production function where the production factors are schooling, through in-person classes and remotely, and the family. The main component driving the shock to human capital is the loss in instructional time suffered by children due to school lockdowns during the pandemic. Families may hereby partly or perfectly substitute schooling depending on the highest educational degree attained by parents. An additional channel causing instructional losses is the likelihood of infection and death within the family, approximated by the relative number of Covid-19 cases and deaths per inhabitant.

The core of the simulation exercise is the construction of a counterfactual human capital equivalent for each individual.[2] This counterfactual quantifies, for people belonging to cohorts that went to

[1] Strictly speaking, we are estimating the intergenerational persistence of human capital.

[2] Monroy-Gomez-Franco et al. (2021) extend the methodology proposed by Neidhöfer et al. (2021) to, first, include parental income as further component of the model and, second, consider the additional impact of instructional losses

school before the pandemic, the individual's level of education if he or she would have experienced the pandemic while still in school in his or her country of residence. The counterfactual education $\hat{e}$ is defined as:

$$\widehat{e_{ijc}} = e_{ijc} - \kappa_{ijc}. \tag{1}$$

Here, e are the actual reported years of education of individual i, with parental educational background j, living in country c.[3] κ is the human capital loss, measured as the share of instructional time lost during the pandemic:

$$\kappa_{ijc} = \frac{\left(t_c - M_{ijc} + \tau_{ic}\right) \cdot (1 - p_j)}{T_c}. \tag{2}$$

This human capital loss is defined by several components and is divided by the total days of school in a regular school year of schooling in the country (T_c).[4]

The components of equation (2) are defined in the following way.

t_c is the instructional loss due solely to school closures; namely the number of days that schools were closed, or partly closed, due to the pandemic. Hereby, days in which schools were partly closed count as half of a day of instructional loss and days of full closures as entire days of instructional loss. We obtain the data on the total duration of school closures from UNESCO (see Section 3 and Appendix A in the Supplemental Material).

M_{ijc} are the days of instructional loss compensated by government policies geared to support not-in-person schooling:

$$M_{ijc} = \delta_1 t_c f_c + \delta_2 t_c n_c \cdot P(A_{ijc} = 1). \tag{3}$$

This term specifies that the days in which schools were closed (t_c) can be compensated either by remote learning through offline resources (such as TV, radio, cellphone or printed copies) and

in the short-run on cumulative learning losses. For the first extension, information on parental income is required, while for the second extension data on test scores (or a comparable measure of learning) is needed. Since we do not have these information available, we consider the original version of the methodology. Due to the usually high correlation between parental education and earnings, we do not expect that applying the first extension would change our results significantly. Applying the second extension could indeed lead to estimates of even higher learning losses. Hence, our results should be considered as lower bound estimates of the persistent learning losses caused by the Covid-19 pandemic in Sub-Saharan Africa. In Section 6, we provide simple estimates of cumulative educational losses deriving from our estimates.

[3] To avoid measurement error and harmonize the measure of education across countries, following most of the literature we do not use the actually reported years of schooling, but impute the regular years of education necessary to obtain the attained educational level. Hence, incomplete primary education is equivalent to three years of schooling, complete primary to five, incomplete secondary to eight, complete secondary to 12, and more than secondary education to 15. We assign years of parental education following the same procedure.

[4] Since we consider the years 2020 and 2021, T is the sum of regular days of schooling in two years.

online resources. f and n are indices that range from zero to one, constructed to measure the alternative supply of education with offline and online resources in each country, respectively. We derive these two measures from the UNESCO-Survey of National Education Responses to Covid-19 School Closures (see Section 3 and Appendix A in the Supplemental Material). δ_1 and δ_2 are weights that define the relative efficiency of these tools, which in our baseline specification have both the value of 0.5: i.e., the combination of offline and online resources is potentially able to fully substitute a day of in-person classes. Online learning is, however, only available if the household is connected to the internet. Hence, we interact the term capturing the mitigation through online learning with the likelihood that the individual lived in a household with access to internet in his or her childhood. We approximate this likelihood by the probability that a household with education of the household head j in country c has access to internet $(A_{ijc} = 1)$, which we estimate using household survey data for each country in the sample (see Section 3 and Appendix A in the Supplemental Material).

τ measures the additional instructional loss associated with health shocks suffered by households:

$$\tau_{ic} = \tau^q \cdot P(q_{ic} = 1) + \tau^d \cdot P(d_{ic} = 1). \tag{4}$$

τ considers the human capital loss due to infection (q) of one of the household members with Covid-19 and death (d) due to the latter. To estimate the likelihood of this to happen, the number of infections and deaths per inhabitant in the country is multiplied by the average country-level household size. τ^q and τ^d are the numbers of days of schooling lost due to the time the child cannot dedicate to learning in case he/she or someone in the household is infected, sick, or in case of death of a family member.[5] Following Neidhöfer et al. (2021), we set τ^q to the average days of symptom duration, to around one week (i.e. five days of school), and τ^d to a three week loss of instructional time (i.e. 15 days).

This three components that contribute to the loss and recovery of instructional time during the pandemic can be compensated to a certain degree by parental inputs, captured by the term $1 - p_j$ in equation (2). Here, p_j is the parental factor of substitution. It is defined as:

$$p_j = \frac{e_j^p}{\max\left(e_j^p\right)}, \tag{5}$$

[5] This assumption follows the literature on the impact of household health shocks on education (e.g. Gertler et al., 2004).

where e_j^p are the completed years of schooling of parents with education level j and $\max(e_j^p)$ the maximum level that can be attained (in this analysis, more than a completed secondary degree, equivalent to 15 years of schooling). Consequently, $1 - p_j$ is zero for children with high-educated parents that may fully substitute the instructional losses (i.e. $\kappa = 0 \ \forall t, M, \tau$), and one for children of illiterate parents who completely depend on the provision of remote learning by the education system. For other levels of parental education, the value of $1 - p_j$ lies within this interval. Generally, $1 - p_j$ is aimed to capture the higher capabilities of parents with higher levels of education to support their children's education when schools are closed, as evidenced in several empirical studies (e.g. Engzell et al., 2021; Jaume and Willen, 2019; Maldonando and De Witte, 2020).

For the simulation of changes in intergenerational mobility, the instructional losses resulting from the equations above are attributed randomly to the share of the sample with parental education j that mirrors the likelihood that the household in which the individual grew up has the characteristics displayed above (likelihood to have access to the internet, likelihood of death and infection). The parameter $1 - p_j$ instead can be understood as a proportional instructional loss experienced by all children of parents with a j level of education, or as a certain share of children of parents with a j level of education who suffer the entire instructional loss, while the rest are unaffected. In the first scenario $1 - p_j$ is the degree in which parents with educational level j are able to substitute schooling; in this case, for the simulation the shock is distributed to the degree $(t_c - M_{ijc} + \tau_{ic}) \cdot (1 - p_j)$ evenly to all individuals with the respective parental educational background. In the second scenario $1 - p_j$ is the likelihood of parents with educational level j to perfectly substitute schooling; in this scenario, a shock of the amount of $(t_c - M_{ijc} + \tau_{ic})$ is attributed to a randomly selected share $1 - p_j$ of the population within the group of individuals with parental education j. We report the results of the simulations for the second scenario in the main body of the text and for the first scenario in Appendix B of the Supplemental Material.

Once obtained the counterfactual post-pandemic education of individuals following the procedure explained above, we estimate two standard measures for the intergenerational mobility of education: the intergenerational slope coefficient (a measure of relative intergenerational persistence) and the probability of (absolute) upward mobility (see e.g. Neidhöfer et al., 2018). The first indicator is obtained by regressing the education of children on the education of parents. The second indicator is obtained by estimating the likelihood of individuals with low-educated parents to complete secondary schooling. In the post-pandemic counterfactual, secondary school completion changes for individuals with a completed secondary degree, and not more, whose counterfactual post-pandemic education lies under the threshold of 12 years of education. We also estimate the same likelihood for children of high-educated parents to provide a benchmark for comparison. Following the literature on educational mobility in African countries, we define low-educated parents as those who did not complete primary schooling and those that completed at least primary schooling as "high-educated" parents (Alesina et al., 2021).[6] In order to simulate and quantify the potential impact of the pandemic on intergenerational mobility, we estimate the two indicators using the actual years of education of individuals. Then, we re-estimate them with the counterfactual education. Finally, we measure the difference between the two resulting measures for each indicator.

3 Data

3.1 Household surveys

Our primary source of individual data for each country are nationally representative household surveys. To avoid co-habitation bias in our intergenerational mobility estimates, we only use surveys that include retrospective questions on parental education (see Emran et al., 2015), as well as information on the level of education of respondents. To ensure that individuals completed at least secondary education, we restrict the sample to respondent aged 19 or older. Furthermore, in order to analyze a cohort of individuals whose average level of education and educational mobility is as close as possible to those who are in the education system in 2020 and 2021, we restrict the sample to people born between 1987 and 1994. Our final sample comprises 26,884 individuals, ranging from 1,530 individual observations for Kenya to 5,209 for Ghana. We harmonize education levels across countries using the ISCED scale conversions provided by UNESCO. Appendix A in the Supplemental Material describes the household survey data and the exact classification of education used in each country.

3.2 Country-level data on school closures and other characteristics

To compute the single measures for t, M and τ described in Section 2, we complement the household survey data with country-level data on school closures in 2020 and 2021, information on educational mitigation strategies, and epidemiological parameters on the number of Covid-19 cases and deaths per inhabitant. Data on school closures is retrieved from UNESCO. Information

[6] Results using a different threshold to define low-educated and high-educated parents (completed secondary schooling) are included in Appendix B of the Supplemental Material.

on educational mitigation strategies are retrieved from the UNESCO-Survey of National Education Responses to Covid-19 School Closures. Data on Covid-19 cases and deaths from the website *Our World in Data*. As mentioned, we multiply this last two indicators by the average number of people living in the same household in each country, retrieved from data by the United Nations, Department of Economic and Social Affairs, Population Division. Furthermore, we estimate the distribution of internet access by level of education of the household head – our measure for A in equation (3) – using household survey data for each country in the sample. Appendix A in the Supplemental Material describes the country-level data and its sources more in detail.

4 The impact of COVID-19 on education

4.1 Instructional losses, educational inequality and intergenerational persistence

Applying the methodology explained in Section 2, we obtain estimates for the impact of the pandemic on educational achievements and intergenerational mobility. Table 1 provides a first estimate of the inequality in instructional time losses during the pandemic. Considering the days of school closures in 2020 and 2021 and the educational mitigation measures, the application of our methodology shows that disadvantaged children (children whose parents did not complete primary education) lost between 31 and 118 days of instruction, while their peers with a more favorable parental educational background (children whose parents completed at least primary education) only between 10 and 33. The country with the strongest inequality is South Africa, where we observe a difference of 85 days of instruction between the two groups.

These instructional time losses have also detrimental repercussions on the predicted changes in the likelihood of children to complete secondary schooling. Figure 1 shows the predicted change in secondary completion due to the instructional loss. As is evident, the share of individuals with complete secondary education decreases substantially due to the pandemic's impact on school closures. On average across the eight Sub-Saharan countries in our sample, secondary school completion rates decrease by 12 percentage points.

Figure 2 shows the heterogeneous impact of the pandemic on the likelihood to complete secondary education of children of low and high-educated parents. Children with high-educated parents are affected, but the shock hits strongly the already rather low likelihood to complete secondary education of children with low-educated parents in Sub-Saharan countries. In all countries this likelihood decreases substantially. On average across all countries, the decline equals 16 percentage points. In most countries the chances of secondary school completion of disadvantaged children reach a level even lower than ten percent; i.e. less than one of ten children with low-educated parents leaves education with at least a secondary school degree. In Malawi, the likelihood even decreases to less than one percent. This effect is driven by children at the margin, namely those that in the pre-Covid scenario completed secondary education but did not continue with tertiary education afterwards. Since this condition mostly applies for children with low-educated parents, the impact of the pandemic is particularly strong at the lower bottom of the distribution and contributes to higher inequality in educational achievements.

Based on our projections, the intergenerational mobility of education is expected to fall in all countries under analysis. Conventionally, intergenerational mobility is measured by the "persistence" coefficient, which measures the partial correlation between the years of schooling of parents and children. The larger (smaller) the coefficient, the lower (higher) is mobility. Figure 3 shows the slope coefficient of intergenerational persistence measured pre-COVID and the post-pandemic counterfactual. Since this measure considers the entire distribution of years of education, and not just a marginal threshold as captured by the likelihood of secondary school completion discussed above, the size of the effects is generally smaller. However, relative to the size of the slope coefficient, the effect ranges between a one percent increase in persistence, and an increase by almost 50 percent.

4.2 Impact by gender

We estimate the effect separately for men and women and evaluate the impact of the pandemic on gender inequality in educational attainments. We here mainly focus on the overall results. The results for children of low and high-educated parents separately are available in Appendix B in the Supplemental Material.

Figure 4 shows that in all countries but the cases of Kenya and South Africa, the likelihood of male children belonging to the 1987-1994 cohort to complete secondary education is substantially higher than the likelihood of their female peers. In Kenya completion rates are similar among males and females, while in South Africa females have a slightly higher completion rate than males. Our estimates show that secondary completion rates of male children may decrease by more in absolute terms than completion rates of female children.

Figure 5 shows the resulting impact on the gender gap in the likelihood to complete secondary schooling. While in most countries the gender gap diminishes because of the stronger decrease in secondary school completion rates among men, in Kenya the gender gap stays at a constant level, but changes the pattern: before the pandemic, completion rates in Kenya were higher for male than for female, while after the pandemic this trend is reversed. In Ghana the gender gap stays virtually the same as before.

The explanation for these differential effects by gender depends on the different quantity of boys and girls that in the pre-COVID scenario completed secondary education but did not continue with tertiary education afterwards. In most countries, although boys are more likely to complete secondary education, among those that complete secondary education, the likelihood to continue with tertiary education after completing secondary education is higher for female. This explains why the effect of the pandemic on the fundamental threshold of secondary school completion is less strong for female than male students.

4.3 Additional effect of increases in teenage pregnancy on girls' education

In this part of the analysis, we estimate the additional impact on secondary school completion of a further potential consequence of the pandemic, which affects the well-being of young girls and their educational achievements: the rise in teenage pregnancy. Teenage pregnancy reduces the probability of receiving a high school diploma and enroll in tertiary education, while increasing the likelihood of leaving school without a qualification (e.g. Fergusson and Woodward, 2000; Fletcher and Wolfe, 2009). Confinement and deprivation during lockdowns is believed to dramatically worsen the situation regarding child abuse and teenage pregnancy, especially among vulnerable families. Indeed, media reports, statements by local NGOs and reports of international organizations in several African countries provide anecdotal evidence that anticipate an increase in adolescent birth rates during the Covid-19 pandemic (UNICEF, 2020). Recent reports confirm this picture. In Malawi, for instance, an 11% increase in teenage pregnancies was recorded from January to August 2020 compared to the same period in 2019 (UNFPA, 2021). Scientific studies document a dramatic rise in the risk of young girls becoming pregnant, for instance, in Kenya (Zulaika et al., 2022), Nigeria (Musa et al., 2021) and other African countries (see the review by Willie, 2021).

To take into account the effect of teenage pregnancy on education, we extend the exercise described in Section 2 to account for this additional shock affecting the human capital formation of girls during the pandemic. We do so by including a further component into the model that allows to simulate the impact of an increase in the likelihood of young girls to become pregnant on secondary school completion rates in each country.

Formally, the counterfactual post-pandemic education becomes:

$$\widetilde{e_{ijc}} = \widehat{e_{ijc}} - P(\zeta_{ijc} = 1) \cdot Z. \tag{6}$$

$\widehat{e_{iJc}}$ are the counterfactual years of schooling defined in equation (1). $P(\zeta_{ijc} = 1)$ is the likelihood of a girl with parental background *j* in country *c* to drop out from school due to pregnancy during the Covid-19 pandemic. To account for socio-economic differences in the probability of this event to occur and its consequences for the educational career of girls, $P(\zeta_{ijc} = 1)$ is obtained by multiplying the percentage point increase in teenage pregnancy in the country due to the pandemic, with one minus the parental factor of substitution, defined in equation (5).[7] Z quantifies the consequences of pregnancy on education, which we set to a loss equal to the entire amount of two years of education.

To estimate the parameters of the model, we collect data on the increase in adolescent birth rates for each country. However, due to underreporting during the pandemic, birth registry data from low-income countries remains incomplete and no clear conclusions can be drawn of how COVID-19 affected births, in general, and teenage pregnancy rates (UNFPA, 2021). Hence, we rely on estimates on the increase in teenage pregnancy from different reports (see a more detailed description in Appendix A in the Supplemental Material). Among these, we choose the worst case scenarios, which indicates an increase in teenage pregnancy due to the pandemic of about 75%.[8] Based on this figure for the potential increase, and on data on adolescent birth rates before the pandemic, for each country we project the percentage point increase in adolescent birth rates. Then, we simulate the impact of this additional factor, besides of school closures, on the education of girls. Finally, we estimate, likewise to the estimations performed before, the resulting impact on secondary school completion rates.

Figure 6 shows the results. The first bar shows, for each country, the likelihood of secondary school completion of girls born between 1987 and 1994, while the second bar shows the post-pandemic counterfactual; both already shown in Figure 4. The third bar, shows the post-pandemic counterfactual also considering the estimated increase in teenage pregnancy rates. It turns out that the potential increase in teenage pregnancy contributes marginally to further increase educational drop-out. In three countries, the impact on the decline in secondary school completion rates is around one percentage point, while in the other countries it is lower than one percentage point.

[7] Hence, in this simulation the increase in the risk of getting pregnant is set to be equal for girls of all parental backgrounds, namely equal to the average increase in the adolescent birth rate in the country of residence. What varies by parental background is the risk that this pregnancy leads to drop out from education, which is obtained by multiplying with one minus the parental factor of substitution. This means that the likelihood to drop out from school in case of pregnancy of girls whose parents are highly educated is lower than the likelihood of girls whose parents have low education.

[8] Estimates based on this worst case scenario can be considered an upper bound. Indeed, in this part of the analysis we are interested in understanding the maximum contribution of teenage pregnancy on the top of the effect of school closures. Estimates based on the other scenarios show an even lower impact of teenage pregnancy and are available upon request.

The main effect of the pandemic on education is, hence, confirmed to be driven by school closures. Although this may come as a surprise to some, our result is in line with studies that found that for the majority of young women, pregnancy occurred after dropping out from school, rather than the opposite (Fergusson and Woodward, 2000). In conclusion, concentrating resources on keeping girls in school could also help to counteract the problem of teenage pregnancy as part of the post-COVID remedial actions. A caveat is in order, however. The pandemic created unique circumstances. First, school closures could be seen as equivalent to "forced dropouts" especially for girls of low socioeconomic background. In that case, the findings by Ferguson and Woodward (2000) would apply. Furthermore, lockdowns exacerbated the circumstances within the household that can lead to teenage pregnancy given that members were forced to stay at home for lengthy periods.

5 Decomposition of effects and simulation of policy alternatives

The impacts shown in the previous section refer to the bundled effect of school closures (adding the potential impact of Covid-related health issues) and the mitigating impact of remote learning policies. We now evaluate the potential of educational mitigation policies to reduce instructional losses and temper the decrease in intergenerational mobility. First, we decompose the impact of the enacted policies vis-a-vis the effect that school closures would have had without any mitigation measure. Then, we simulate different scenarios either improving the policies or the infrastructures that interact with the effectiveness of these policies.

Figure 7 ranks the countries in our sample by the estimated decrease in intergenerational mobility, measured by the slope coefficient, taking into account the closure of schools and the enacted offline and online educational mitigation strategies. The figure also shows the projected decrease in mobility due to school closures that would have taken place without mitigation measures. As can be seen clearly in the graph, although the effect of mitigation measures is sizable in most countries, it is not sufficient to close the gap caused by the closure of schools.

As a next step, we simulate which combination of measures and infrastructural improvements would allow to cushion the negative effects of the pandemic on intergenerational mobility. We successively change the parameters for online learning and internet coverage, keeping offline level at the current level, and measure the impact it would have on the difference between the pre-pandemic level of intergenerational mobility and the post-pandemic counterfactual. Figure 8 shows the results of this policy exercise. The first bar in the graph shows the baseline situation, namely the decrease in intergenerational mobility given the current distribution of internet coverage and online learning tools provided by the education systems. The change is always displayed in points of the partial correlation between parents' and children's education (i.e. the slope coefficient). The second bar shows the estimates obtained with improved online learning – i.e. setting the index for online learning tools, i.e. *n* in equation (3), to one – while keeping constant the distribution of internet access in the population. In the third and fourth bar, instead, results are obtained by improving internet coverage while keeping the current value of online learning constant: first, internet coverage is improved in such a way that each individual in the population has twice the likelihood to have access to the internet; then, universal internet coverage is granted, i.e. the likelihood of internet access is set to 100% for all individuals in the sample. The last bar shows the results of granting both, full internet access and improved online learning.

The analysis shows that in all countries, given the current distribution of internet access, an improvement of the provided online learning resources would have no sizeable effect on reducing the negative effect of the pandemic on intergenerational mobility. At the same time, improving internet access alone, even granting universal internet access, would not be enough to close the gap. An unrealistically strong and costly effort by states would have been necessary, both improving online learning tools and the current infrastructure, to fully mitigate the impact of the closure of schools on educational disruptions by offering online remote learning.[9]

6 Long-run inequality

Here we estimate the potential long-run repercussions of the effect of the pandemic on education on long-run earnings inequality. This exercise follows three steps. First, we simulate the counterfactual post-pandemic years of education for each individual in the affected group following the methodology explained in Section 2. Second, we estimate the earnings of each individual applying the returns to education to the actual years of education and the counterfactual ones. Thus, we obtain two earnings distributions: one in the absence of COVID-19 and one with our estimated post-pandemic effect. Third, we estimate the degree of inequality in these two distributions and compare them. The impact of the pandemic on long-run inequality is then identified as the difference between the inequality in the actual earnings distribution and its post-pandemic counterfactual.

Importantly, for this exercise we include a further identifying restriction to the learning losses. While in the main analysis, years of education are understood as a human capital equivalent measure

[9] The negative sign of some coefficients, i.e. a higher level of intergenerational mobility in the counterfactual with respect to the pre-pandemic scenario, derives from the fact that with universal internet coverage children from low-educated families benefit most from the mitigation measures, while the likelihood of infection and death in the family may affect individuals over the entire distribution. However, the size of the difference shows that this equalizing effect would be negligible.

attached to the instructional time, which is shocked by the pandemic, for the analysis of earnings estimates it is more meaningful to focus on completed full years of education. Hence, we translate our measure of instructional losses into full years of education. Hereby, we assume that an individual that suffers a moderate instructional loss, which following Kubitschek et al. (2005) we define to be less than 25% of a school year, may complete the school year. Instead, individuals who lost between 25% and 50% of instructional time in the two school years are assumed to lose an entire year of schooling. Individuals who lost more than 50% of instructional time lose the complete two years of irregular schooling time.[10]

Figure 9 shows the average years of education of the 1987-1994 cohort in each country and the post-COVID counterfactual scenario estimated with our methodology and the further restrictions explained above. We call this a scenario of "cumulative instructional losses", which should display the upper bound of the effect, in contrast to the "baseline scenario", which displays the lower-bound. In all countries the decrease in years of schooling is substantial, both in absolute and in relative terms. Furthermore, the average decrease is particularly high among the children of low educated parents (graphs showing the heterogeneities by parental background for each country are displayed in Appendix B of the Supplemental Material).

To measure how these changes in average years of schooling translate into changes in earnings inequality, we estimate the permanent earnings of each individual.[11] Earnings y are defined as:

$$\hat{y}_{isc} = \overline{w}_{sc} \cdot r_{sc} \cdot e_{isc} \tag{7}$$

Where e are the years of education of individual i. $\overline{w}_{sc}$ are the average monthly earnings of male workers $(s = m)$ or female workers $(s = f)$ in country c, which we retrieve from ILO data on labor statistics for each country. r are the wage returns to a year of full education for men or women in country c, which we retrieve from the literature review by Psacharopoulos and Patrinos (2018).[12] We estimate y using both, the actual years of education and the post-pandemic counterfactual

[10] Estimates of changes in intergenerational mobility due to the pandemic using this alternative specifications are included in Appendix B of the Supplemental Material.

[11] An alternative to obtain predicted earnings would be to run a Mincer regression including years of education on labor earnings and then predict earnings using the counterfactual education of each individual. However, since most surveys either do not report labor earnings or have many missing values, we opt for this alternative estimation procedure to obtain a predicted measure for permanent earnings which varies by gender and education.

[12] For this exercise, we have the necessary information for all countries but Liberia, which hence is excluded from this part of the analysis. More information on the data that we use for these two measures is included in Appendix A of the Supplemental Material.

obtained applying the method explained in Section 2. Finally, we estimate the Gini coefficient of both earnings distributions.[13]

Table 2 shows the results of the exercise. The values in the first column show the Gini coefficients derived from the earnings distribution estimated using the actual years of education of each individual. In the second column, the Gini coefficient is estimated on the counterfactual distribution of earnings, i.e. those estimated using the counterfactual post-pandemic years of education. The third column shows the difference between the two Gini coefficients and the fourth column the standard error of this difference. The upper panel of the Table shows the scenario of cumulative instructional losses explained above, the lower panel shows the baseline scenario. We observe that in the former, which can be considered an upper bound, the impact on earnings inequality is between one and four points of the Gini. In the latter, which can be considered a lower bound, the impact is on average halved, but still around one Gini point in most countries.

Of course, the presented analysis has several caveats. It is based on several simplifying assumptions, some of them based on parameters that are also expected to change as a consequence of the pandemic. In addition, the estimate for earnings inequality that our analysis provides, measures the intra-cohort level of earnings inequality, which is not comparable with usual measures of wage inequality based on a wider population. Furthermore, since the statistics on monthly wages and returns to education belong to different years, the estimates are also not comparable across countries, but only within countries. Hence, the findings should be understand as merely indicative for the potential economic significance and persistence of the effect of the education channel in the long-run. Nevertheless, the results of this simple exercise are conspicuous. They suggestively show that the education channel is a factor that will play an important and determinant role for the impact of the pandemic on long-run inequality.

7 Conclusions

Using microsimulation and the framework proposed by Neidhöfer et al. (2021), we estimated the impact of the pandemic-related instructional losses on educational achievements and intergenerational mobility in eight Sub-Saharan African countries. We focused on the asymmetric effects of school closures on the education of children with different socio-economic backgrounds. We analyzed the potential mitigating impact of policies and proposed a straightforward first-order

[13] Of course, this analysis gives only a static representation of the potential long-run consequences, abstracting from dynamic and endogenous effects on (i) earnings and (ii) returns to education, which are both likely to change as well as a consequence of the pandemic. However, estimating the effect of the pandemic on these two parameters would go beyond the scope of this work.

estimate of the potential effect on future earnings inequality caused by the pandemic's predicted impact on education.

Our findings show that the pandemic is likely to have a significant negative effect on average years of schooling and secondary school completion rates. Intergenerational mobility of education is bound to be lower as well. Educational inequality is expected to rise because the effect of school closures is stronger for disadvantaged children, who are at a higher risk to drop out from the education system without completing a secondary school degree.

On average, intergenerational mobility of education in the eight Sub-Saharan countries in our sample is expected to decrease by 10%, while the likelihood of children from low-educated families to complete secondary education could decrease, on average, by 16 percentage points. Our simulation shows that, for several SSA countries, the likelihood of children from disadvantaged families to complete secondary education may even become lower than ten percent. This means that in these countries less than one of every ten children with low-educated parents affected by the COVID-19 crisis may leave education with a secondary schooling degree. This alarming picture mirrors the projections of Neidhöfer et al. (2021) for Latin America, which found some confirmation in current analyses of education drop-outs using administrative data for Brazil (Lichand et al., 2021).

Furthermore, our projections of the impact on long-run inequality suggest that the education channel could become an important determinant of earnings inequality in the future. Although these findings have surely their limitations, they highlight the importance of remedial actions that should be taken as quickly as possible to temper the impact on learning and educational losses especially for the more disadvantaged children.

References

Abadzi, H. (2009). Instructional time loss in developing countries: Concepts, measurement, and implications. The World Bank Research Observer, 24(2), 267-290.

Adda, J. (2016). Economic activity and the spread of viral diseases: Evidence from high frequency data. The Quarterly Journal of Economics, 131(2), 891-941.

Alesina, A., Hohmann, S., Michalopoulos, S., & Papaioannou, E. (2021). Intergenerational mobility in Africa. Econometrica, 89(1), 1-35.

Angrist, N., de Barros, A., Bhula, R., Chakera, S., Cummiskey, C., DeStefano, J., ... & Stern, J. (2021). Building back better to avert a learning catastrophe: Estimating learning loss from COVID-19 school shutdowns in Africa and facilitating short-term and long-term learning recovery. International Journal of Educational Development, 84, 102397.

Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. International Journal of Educational Development, 86, 102480.

Azevedo, J. P., Hasan, A., Goldemberg, D., Geven, K., & Iqbal, S. A. (2021). Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. The World Bank Research Observer, 36(1), 1-40.

Dang, Hai-Anh; Oseni, Gbemisola; Zezza, Alberto; Abanokova, Kseniya. (2021). Impact of COVID-19 on Learning: Evidence from Six Sub-Saharan African Countries. World Bank, Washington, DC.

Emran, M. S., Greene, W., & Shilpi, F. (2018). When measure matters coresidency, truncation bias, and intergenerational mobility in developing countries. Journal of Human Resources, 53(3), 589-607.

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. Proceedings of the National Academy of Sciences, 118(17).

Fergusson, D. M., & Woodward, L. J. (2000). Teenage pregnancy and female educational underachievement: A prospective study of a New Zealand birth cohort. Journal of Marriage and Family, 62(1), 147-161.

Fletcher, J. M., & Wolfe, B. L. (2009). Education and labor market consequences of teenage childbearing evidence using the timing of pregnancy outcomes and community fixed effects. Journal of Human Resources, 44(2), 303-325.

Gertler, P., Levine, D. I., & Ames, M. (2004). Schooling and parental death. Review of Economics and Statistics, 86(1), 211-225.

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of COVID-19-Related School Closures on Student Achievement-A Systematic Review. Frontiers in Psychology, 4020.

IMF (2021). Regional economic outlook. Sub-Saharan Africa : one planet, two worlds, three stories. Washington DC: International Monetary Fund (IMF).

Jaume, D., & Willén, A. (2019). The long-run effects of teacher strikes: evidence from Argentina. Journal of Labor Economics, 37(4), 1097-1139.

Kubitschek, W. N., Hallinan, M. T., Arnett, S. M., & Galipeau, K. S. (2005). High school schedule changes and the effect of lost instructional time on achievement. The High School Journal, 89(1), 63-71.

Lichand, G., Dória, C. A., Neto, O. L., & Cossi, J. (2021). The impacts of remote learning in secondary education: Evidence from brazil during the pandemic. Inter-American Development Bank Technical Note N. IDB-TN-02214.

Maldonado, J. E., & De Witte, K. (2022). The effect of school closures on standardised student test outcomes. British Educational Research Journal, 48(1), 49-94.

Michael Mncedisi Willie. (2021). Teenage Pregnancy During a Pandemic. International Journal of Women's Health Care, 6(3), 218-219.

Monroy-Gómez-Franco, L., Vélez-Grajales, R., & López-Calva, L. F. (2022). The potential effects of the COVID-19 pandemic on learnings. International journal of educational development, 102581.

Musa, S. S., Odey, G. O., Musa, M. K., Alhaj, S. M., Sunday, B. A., Muhammad, S. M., & Lucero-Prisno, D. E. (2021). Early marriage and teenage pregnancy: The unspoken consequences of COVID-19 pandemic in Nigeria. Public Health in Practice, 2, 100152.

Neidhöfer, G., Lustig, N., & Tommasi, M. (2021). Intergenerational transmission of lockdown consequences: prognosis of the longer-run persistence of COVID-19 in Latin America. The Journal of Economic Inequality, 19(3), 571-598.

Neidhöfer, G., Serrano, J., & Gasparini, L. (2018). Educational inequality and intergenerational mobility in Latin America: A new database. Journal of Development Economics, 134, 329-349.

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. Education Economics, 26(5), 445-458.

The Economist (2022). Tracking covid-19 excess deaths across countries. https://www.economist.com/graphic-detail/coronavirus-excess-deaths-tracker

UNESCO (2022). Global monitoring of school closures.

UNFPA (2021). How will the COVID-19 pandemic affect births? Technical Brief.

UNICEF (2020). COVID-19: A Catastrophe for Children in Sub-Saharan Africa. Technical Report.

Werner, K., & Woessmann, L. (2021). The legacy of covid-19 in education. CESifo Working Paper No. 9358.

World Bank (2022). Global Economic Prospects: Sub-Saharan Africa.

Zierer, K. (2021). Effects of pandemic-related school closures on pupils' performance and learning in selected countries: A rapid review. Education Sciences, 11(6), 252.

Zulaika, G., Bulbarelli, M., Nyothach, E., van Eijk, A., Mason, L., Fwaya, E., ... & Phillips-Howard, P. A. (2022). Impact of COVID-19 lockdowns on adolescent pregnancy and school dropout among secondary schoolgirls in Kenya. BMJ global health, 7(1), e007666.

Tables and Figures

Table 1 – Average days of instructional time lost considering mitigation strategies

	Days of instructional time lost in 2020 and 2021		… in percentage of the regular school years	
	disadvantaged children	advantaged children	disadvantaged children	advantaged children
Ethiopia	101	33	25%	8%
Ghana	99	30	26%	8%
Kenya	76	23	17%	5%
Liberia	59	17	13%	4%
Malawi	53	16	13%	4%
Nigeria	86	28	22%	7%
South Africa	118	33	27%	8%
Tanzania	31	10	8%	2%

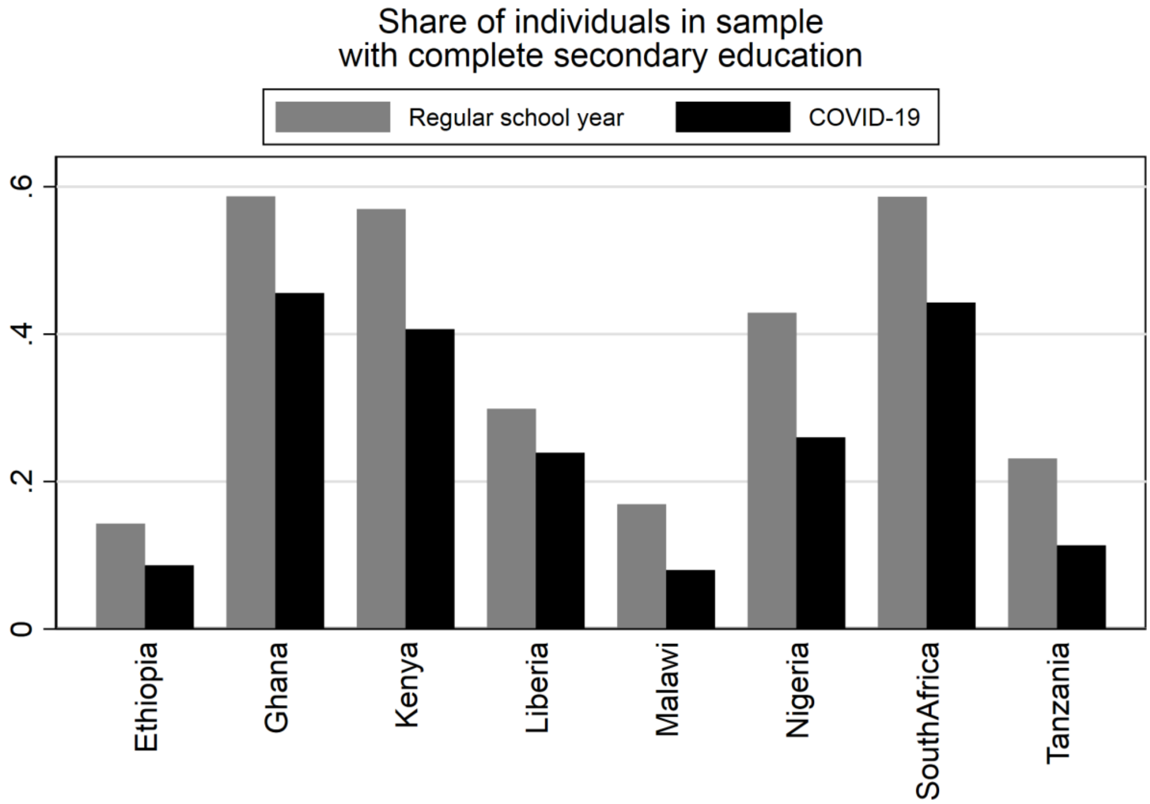
Notes: Disadvantaged children are children whose parents did not complete primary education. Advantaged children are those with at least one parent who completed primary education.

Table 2 – Long-run effect of human capital loss on intra-cohort wage inequality

	Intra-cohort wage inequality (Gini coefficient)			
	pre-COVID	post-COVID	Δ	s.e.
	Scenario of cumulative instructional losses			
Ethiopia	0.655	0.675	0.019	0.001
Ghana	0.539	0.552	0.013	0.001
Kenya	0.254	0.293	0.039	0.002
Malawi	0.683	0.708	0.025	0.001
Nigeria	0.466	0.497	0.031	0.002
South Africa	0.161	0.204	0.043	0.002
Tanzania	0.439	0.464	0.025	0.001
	Baseline scenario			
Ethiopia	0.655	0.664	0.008	0.000
Ghana	0.539	0.545	0.006	0.000
Kenya	0.254	0.267	0.013	0.001
Malawi	0.683	0.689	0.006	0.000
Nigeria	0.466	0.478	0.011	0.001
South Africa	0.161	0.184	0.024	0.001
Tanzania	0.439	0.446	0.008	0.000

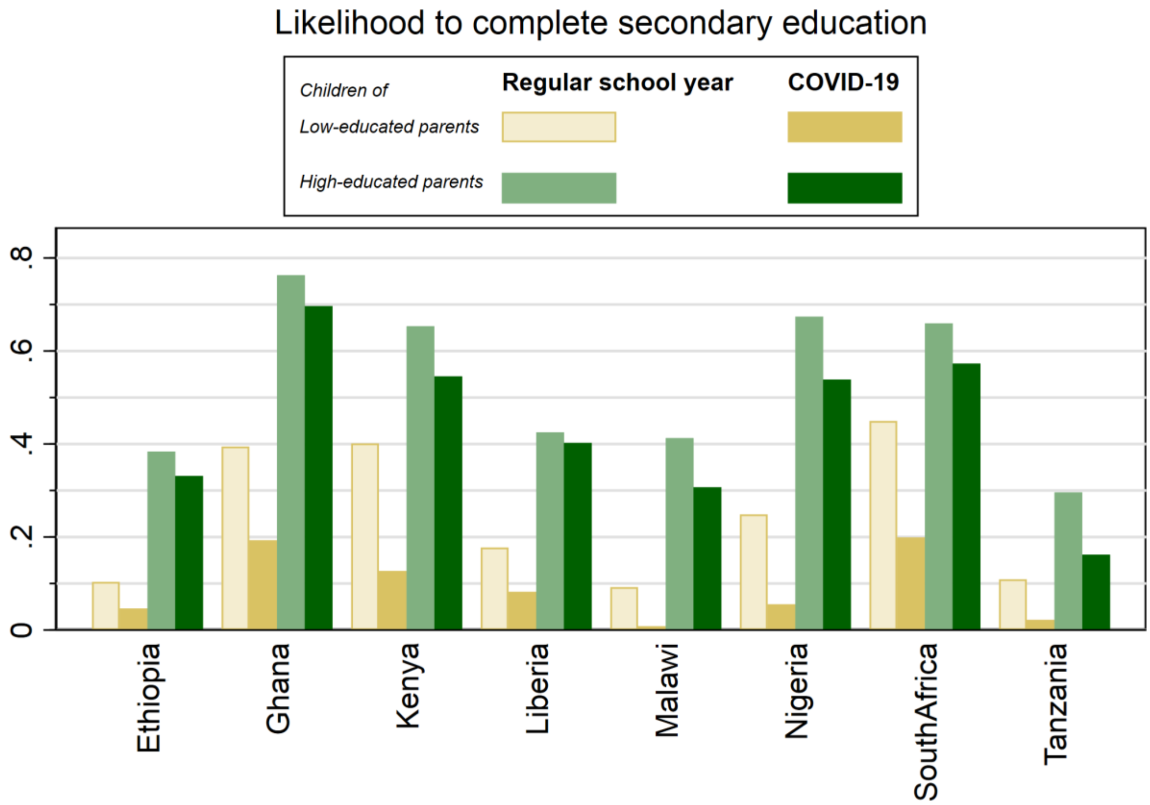
Notes: Columns show predicted intra-cohort earnings inequality, measured by the Gini coefficient, in the pre-pandemic and post-pandemic counterfactual, their difference, and the standard error of the difference. Upper part of the table shows scenario considering cumulative instructional losses, while lower part shows the baseline scenario. Source: National household surveys, own estimates.

Figure 1 – Predicted impact of the pandemic on secondary school completion



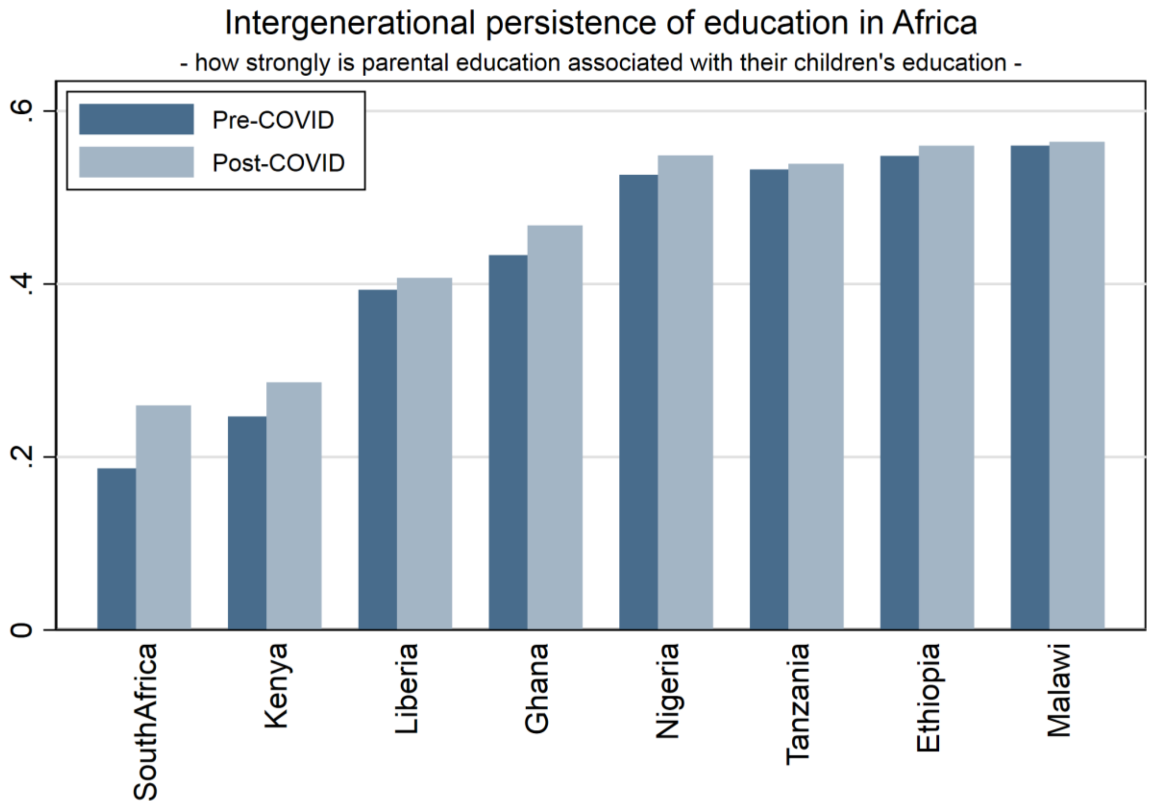
Notes: Completed secondary education is equivalent to 12 full years of schooling. First scenario shows actual share of individuals in sample with completed secondary schooling. Second scenario shows estimates of the same share after simulation of the COVID-19 shock. Source: National household surveys, own estimates.

Figure 2 – Predicted impact of the pandemic on secondary school completion by parental background



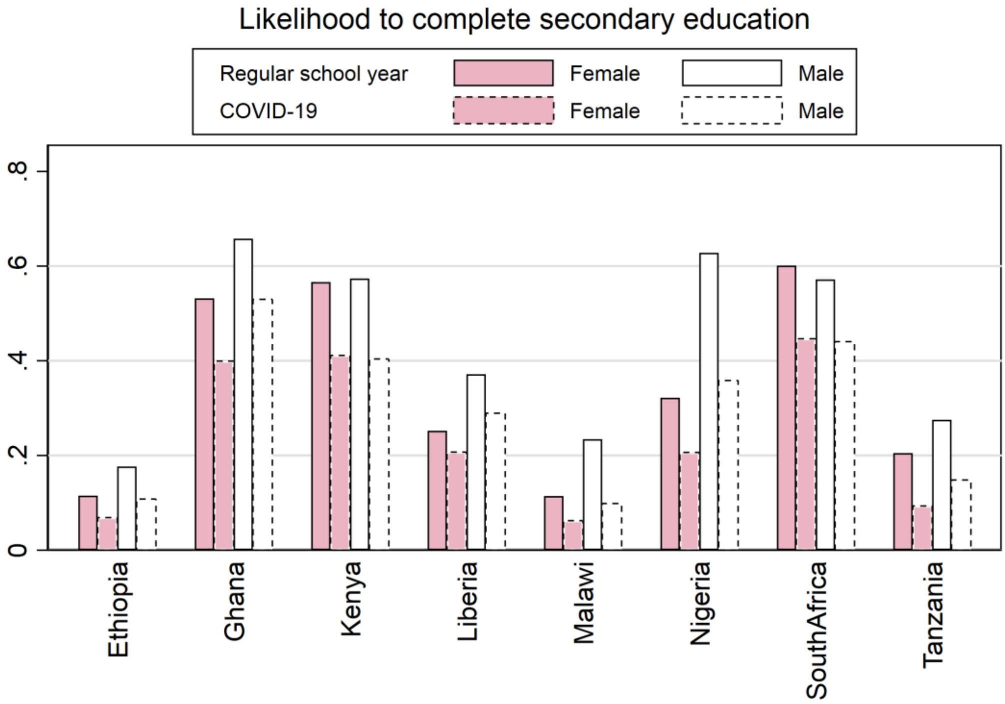
Notes: Bars show the likelihood to complete at least 12 years of schooling before and after simulation of the COVID-19 shock on education. High educated parents have at least completed primary education, low educated parents less than primary secondary education. Source: National household surveys, own estimates.

Figure 3 – Change in the intergenerational persistence of education due to the COVID-19 pandemic



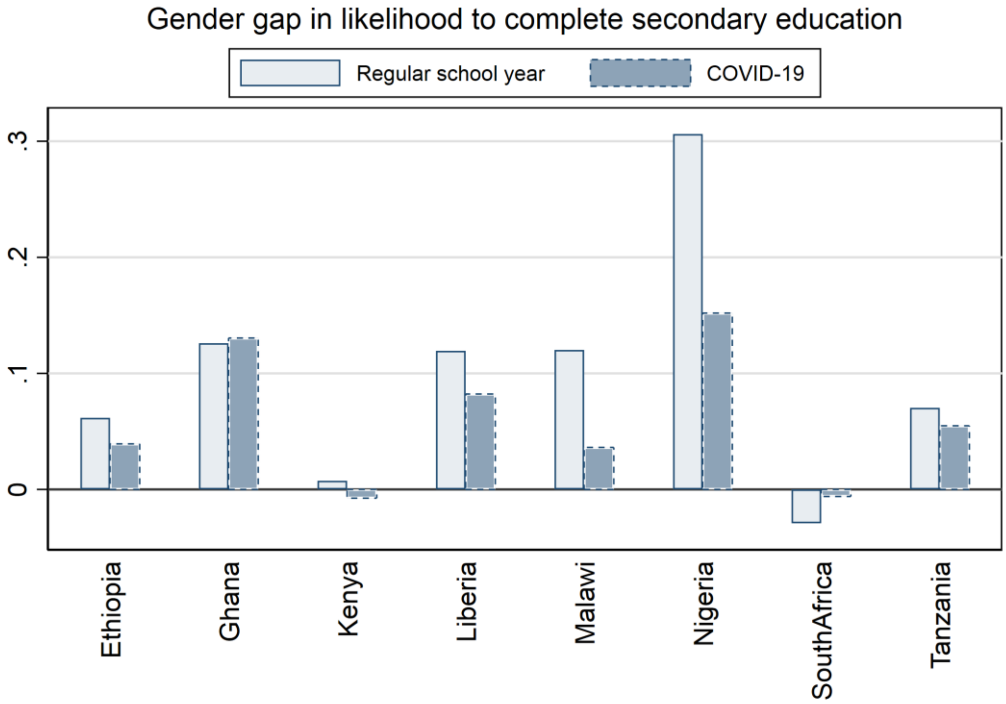
Notes: Bars show the pre-pandemic and counterfactual post-pandemic level of intergenerational persistence of education, measured by the slope coefficient. Source: National household surveys, own estimates.

Figure 4 – Heterogeneous impact of the pandemic on secondary school completion by gender



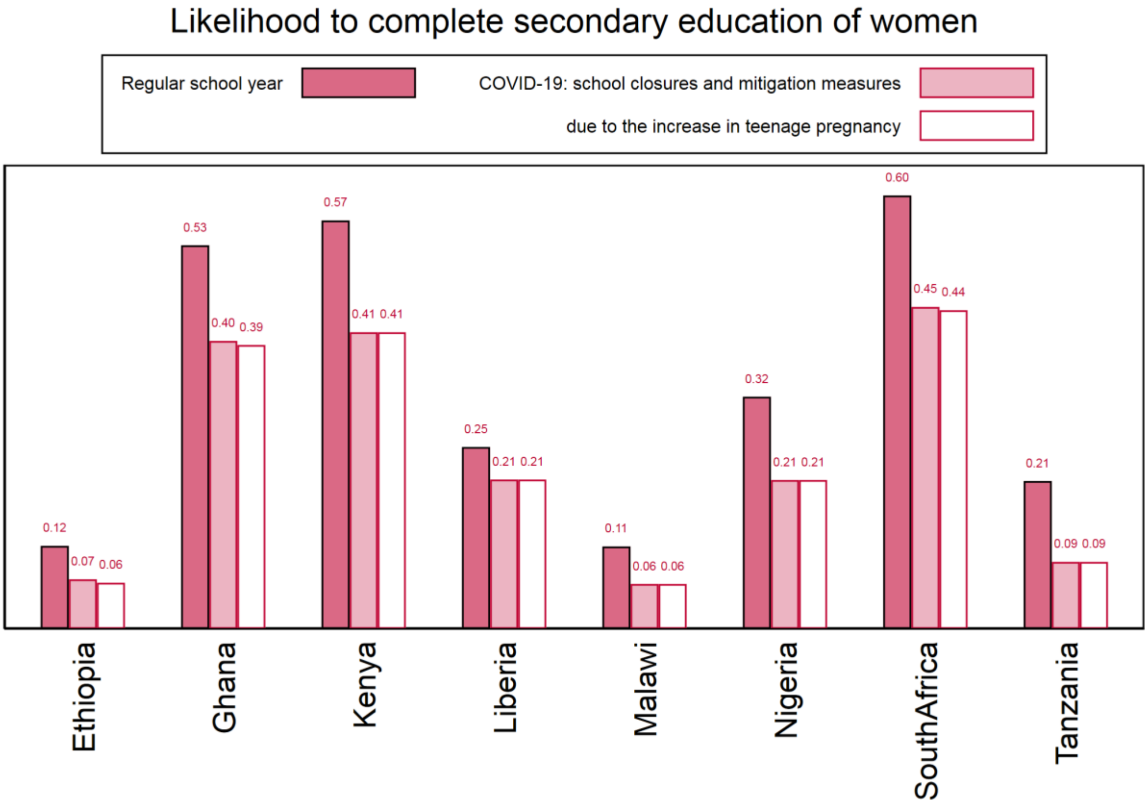
Notes: Bars show the pre-pandemic and counterfactual post-pandemic secondary school completion rate of men and women. Source: National household surveys, own estimates.

Figure 5 – Predicted change in the gender gap in secondary school completion due to the pandemic



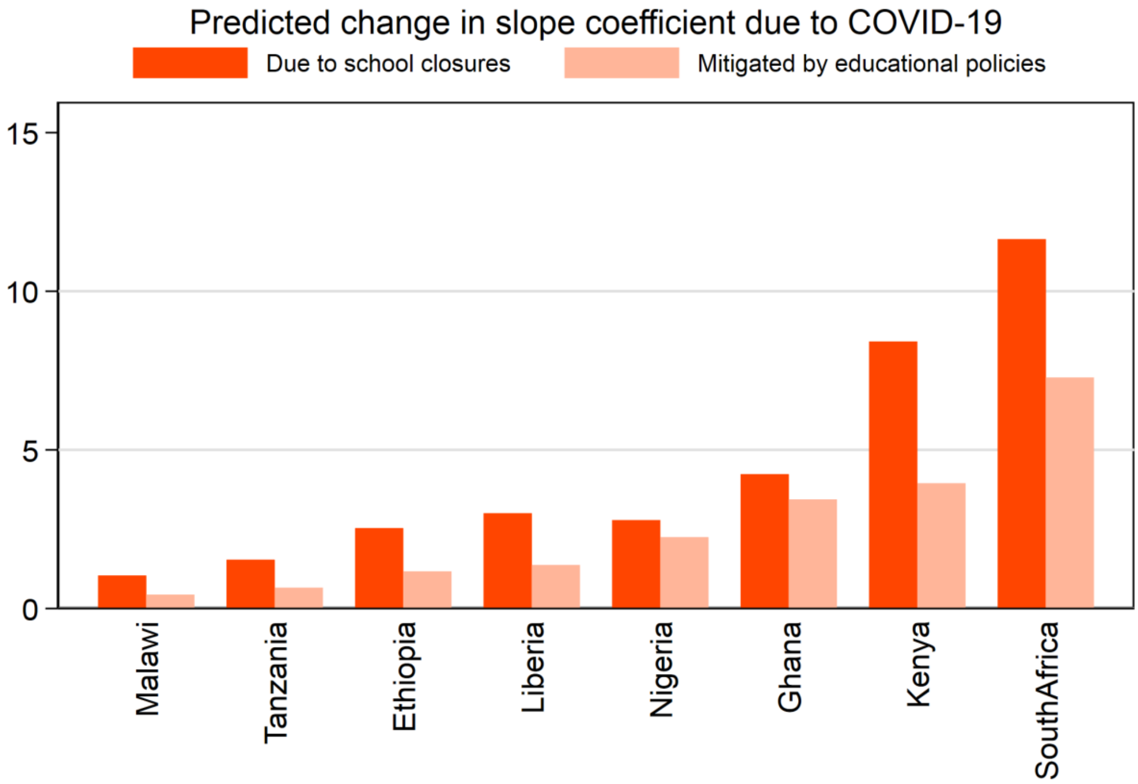
Notes: Bars show the pre-pandemic and counterfactual post-pandemic difference between secondary school completion rates of men and women. Source: National household surveys, own estimates.

Figure 6 – Predicted impact of increasing teenage pregnancy on secondary school completion rates



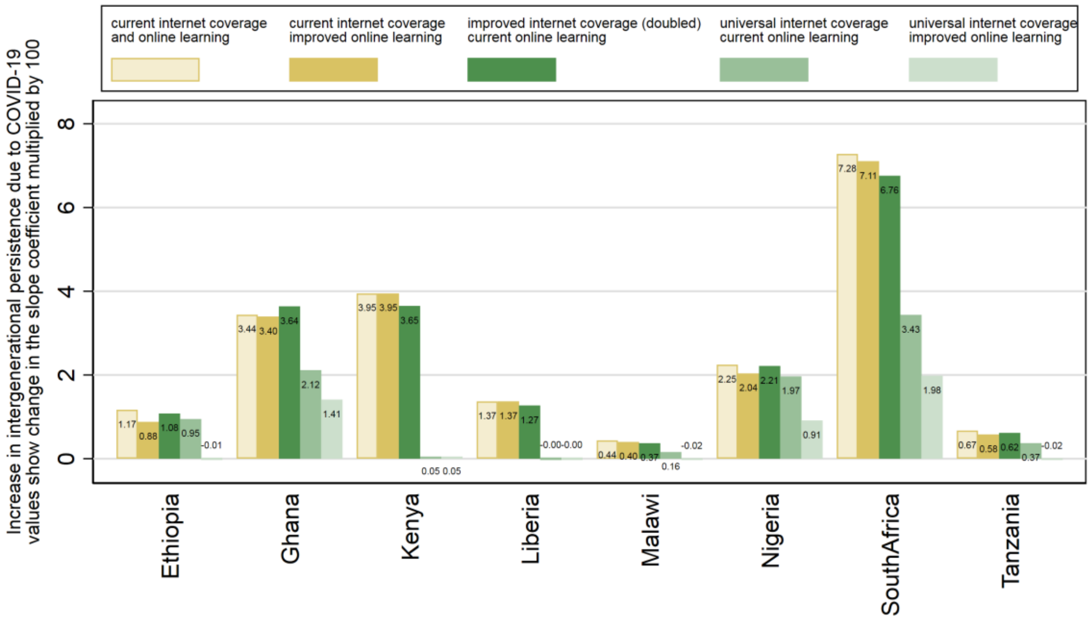
Notes: Bars show the likelihood to complete at least 12 years of schooling before and after simulation of the COVID-19 shock on education. The last bar shows the additional decrease in secondary school completion rates due to the predicted increase in teenage pregnancy. High educated parents have at least completed primary education, low educated parents less than primary secondary education. Source: National household surveys, own estimates.

Figure 7 – Evaluation of the impact of educational policies on reducing the effect of the pandemic on intergenerational mobility



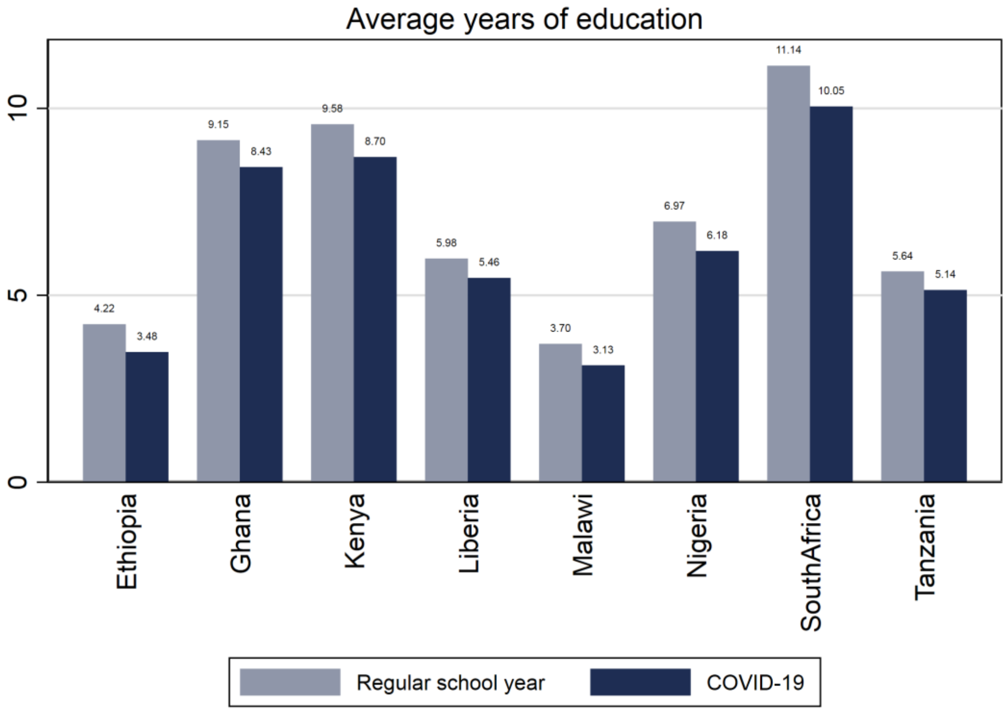
Notes: Bars show the difference (multiplied by 100) between the pre-pandemic and counterfactual post-pandemic level of intergenerational persistence, measured by the slope coefficient, in two scenarios: i) only the effect of school closures on the instructional loss is taken into account; 2) including the mitigating effect of educational policies to provide offline and online remote learning. Source: National household surveys, own estimates.

Figure 8 – Policy simulation exercise to evaluate the combination of measures that would allow to cushion the effects of the pandemic on education



Notes: Bars show the difference (multiplied by 100) between the pre-pandemic and counterfactual post-pandemic level of intergenerational persistence, measured by the slope coefficient, in different scenarios of improved online remote learning and/or internet coverage. The first bar shows the difference between the pre-pandemic level and the baseline estimate for the post-pandemic counterfactual (i.e. given the current online remote learning efforts and distribution of internet access). Source: National household surveys, own estimates.

Figure 9 – Change in average years of education due to the pandemic considering cumulative learning losses



Notes: Bars show the pre-pandemic and counterfactual post-pandemic average years of schooling considering cumulative learning losses. Source: National household surveys, own estimates.

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de