Social Education Research
http://ojs.wiserpub.com/index.php/SER/

Research Article

Explaining Gold-Mining and Non-Gold Mining Areas' Inequalities in Learning Achievements in Burkina Faso's Primary Education: A Decomposition Analysis

Jean-Baptiste M. B. Sanfo[1*], Keiichi Ogawa[2]

[1]Institute for Promotion of General Education, the University of Shiga Prefecture, 2500 Hassaka, Hikone, 5228533, Japan
[2]Graduate School of International, Cooperation Studies, Kobe University, Kobe, Japan
Email: sanfo.j@office.usp.ac.jp

Received: 6 August 2021; **Revised:** 25 October 2021; **Accepted:** 28 October 2021

Abstract: Research shows that learning achievements inequalities exist between students from gold mining areas and those from non-gold mining ones. However, there is no evidence on factors that explain this "new" geographic educational inequality. Exploiting the gold mining boom in Burkina Faso, this study employed re-centered influence function decomposition to explore students' background and school factors which explain these learning achievements inequalities and also estimate the proportion of inequalities explained by unmeasured factors. Findings suggest that, relative to student background factors, most of the learning achievements inequalities between the two types of areas are explained by school factors. Moreover, unmeasured educational factors explain a non-negligible proportion of the inequalities, higher for students on the lower and upper tails of the learning achievements distribution. Suggestions for policymakers are discussed based on the findings of the present study.

***Keywords*:** learning achievements, resource curse, RIF decomposition, Oaxaca decomposition, learning inequalities

1. Introduction

Education benefits the individual, the community, and the country (Dziechciarz-Duda & Król, 2013; Pelinescu, 2015). Research shows that learning outcomes are related to earnings, income distribution, and economic development (Hanushek & Woessmann, 2008). Policies, schools, and families contribute to creating educational opportunities (Tawil et al., 2012), and researchers have also been interested in location-based educational evidence. As such, many studies investigated educational outcomes in rural and urban locations (Myhr et al., 2017; Nieuwenhuis & Hooimeijer, 2016). Findings on rural-urban research show that location also affects educational outcomes, resulting in its inclusion in the debate on the determinants of educational outcomes (Hanushek, 2014; Strand, 2014). In addition to rural-urban research, inspired by findings suggesting a natural resource curse on human capital, scholars are increasingly interested in educational outcomes of students from schools located in mining areas relative to non-mining ones (Adu-Gyamfi, 2014; Rau et al., 2015; Ouma et al., 2017). From this perspective, Botchwey and Crawford (2016) found that students in gold mining areas (GMA) have lower learning achievements than their peers in non-gold mining areas (NGMA) in Ghana. In general, contexts of gold mining in developing countries in Africa or other regions tend to be reasonably similar. In such

DOI: https://doi.org/10.37256/ser.3120221087

countries, populations are directly involved in or exposed to artisanal/small-scale mining, and this has strong negative educational and social implications.

Previous studies, though scarce, suggest inequalities in learning achievements between students in GMA and those in NGMA, providing evidence of "new" geographic learning inequalities. This study is based on a previous work by Sanfo (2021) which demonstrated that GMA is negatively related to student learning outcomes in Burkina Faso, similar to its neighboring country, Ghana. The context of gold mining in Burkina Faso is similar to that of other developing countries. Specifically, adults and children of school age are directly and sometimes illegally involved in mining activities dominated by artisanal/small-scale mining (Soré & Maiga, 2015; Sanfo, 2019). Consequently, implications of artisanal/small-scale mining observed in similar countries seem to be consistent with what is observed in Burkina Faso. However, the context of mining in developing countries is often different from the one in developed countries. In the latter, gold mines are usually industrially extracted, companies or individuals in the sector abide by extraction regulations, and the sector seldom involves children. Therefore, the negative social and economic implications of gold mining in developed countries are minimized.

Research shows that GMA is negatively related to students' educational outcomes (Santos, 2018). In Sanfo (2021), we demonstrated a negative association between GMA and students' learning outcomes in the context of Burkina Faso, even after controlling for relevant student and school factors. However, despite providing evidence of lower learning achievements in GMA relative to NGMA, no previous studies investigated factors that account for the learning gap between the two types of areas. This seems to be a result of the widespread use of methodological approaches which do not allow to capture factors accounting for inequalities between groups (e.g., hierarchical linear modeling). Consequently, existing literature does not provide enough knowledge to promote learning without inequalities. Yet, in the context of the current learning crisis (World Bank, 2018), Burkina Faso and most similar countries are engaged in reaching the Sustainable Development Goal (SDG) 4, which obliges them to address any types of learning inequalities. Therefore, it is critical to provide evidence that will give the necessary direction to address learning achievements inequalities between GMA and NGMA. There is a discourse on factors explaining rural-urban learning achievements inequalities framed around rural areas being worse-off in educational inputs relative to urban ones, and differences in inputs account for the inequalities (Azano & Biddle, 2019). Since GMA in developing countries is dominantly in rural areas, it might be tempting to adopt the same discourse, but there is little evidence to support this.

The aim of the study is to investigate factors that account for learning achievements inequalities between GMA and NGMA. Its specific objectives are 1) to investigate student's background (individual and family) and school factors that account for learning achievements inequalities between GMA and NGMA; 2) to estimate the proportion of GMA-NGMA learning inequalities explained by unmeasured educational factors. Results from the analysis reveal that school factors account for most of the GMA-NGMA learning inequalities and student background factors explain little of these inequalities. Also, unmeasurable education factors account for a non-negligible proportion of the inequalities between the two types of areas, and this "unexplained portion" is higher for students on the lower and upper tails of the learning achievements distributions.

This study is significant because it contributes to creating "new" knowledge that has educational policy and theoretical implications. It empirically highlights factors that account for the learning achievements inequalities revealed by previous studies. Moreover, the method employed helps unpack student heterogeneity and reveals hidden mechanisms or circumstances in the effect of factors explaining learning inequalities (Konstantopoulos et al., 2019). From a policy perspective, factors and mechanisms highlighted by this study will likely contribute to reducing geographical learning achievements inequalities. From a theoretical perspective, the study is significant because it provides an explanation of the natural resource curse on education found by previous studies.

The remaining sections of the article are structured as follows. Section 2 presents an overview of education and gold mining in Burkina Faso. Potential factors that may explain learning achievements inequalities between GMA and NGMA are reviewed in Section 3. The data and method used are explained in Section 4. In Section 5, the results of the analysis are presented and discussed. The study is concluded in Section 6 with suggestions for future studies.

2. Primary education and gold mining in Burkina Faso

Burkina Faso is pursuing the educational objectives of the SDGs. Its education system went under a major reform, producing the "Education Orientation Law" (EOL) that adopted its legal framework. Education has since then become a national priority and a right for all individuals in the country. Currently, primary education in Burkina Faso is mandatory and free (only school fees) for all children. Basic education covers primary and junior high school (the latter called post-primary education) and is compulsory for children from 6 to 16 years old. The education system of the country is composed of preschool, basic education, secondary education (high school), and tertiary education.

Burkina Faso implemented the Plan Décennal de Développement de l'Enseignement de Base (PDDEB) [Basic Education 10-year plan] from 2000 to 2009, a plan which helped increase enrolment rates, build and equip schools, and recruit more teachers. However, the plan did not address the education quality issues the country has been facing (Sore, 2015). After PDDEB, Burkina Faso has been implementing the Programme de Développement Stratégique de l'Education de Base (PDSEB) 2012-2021 [Program for the Strategic Development of Basic Education]. Previous and ongoing efforts increased net enrolment rates from 47.7% in 2005/2006 to 65.7% in 2015/2016. Additionally, within the same period, primary education completion rates increased from 34.1% to 58.4%, and education access gender gaps were reduced. There have been many achievements from a perspective of education access, but many children in Burkina Faso still do not complete even basic education or acquire the basic learning skills that children from their levels are supposed to have acquired (Kouraogo & Dianda, 2008; Ouattara & Sore, 2016; Paré-Kaboré, 2012). In such a context and regarding the limited resources, the central government is seeking to identify what and where to give priority to in the allocation of intervention resources.

Educational challenges are general in Burkina Faso but sometimes they can be specific to given locations, creating inequalities that need to be addressed (MoE, 2015). This seems to be the case in small-scale gold mining areas. The country has a long history of gold mining, but the mining sector was not significant until the recent change. Its number of gold mines increased from 103 to 416 (362 small-scale mines and 54 industrial ones) between 2006 and 2016. However, the sector is dominated by small-scale gold mines which attract people around them. Many of the communities around these mines rely mainly on mining activities for their livelihood. These activities have implications for the livelihoods of local populations, as they can improve socioeconomic standards due to the economic returns they have. From this perspective, in a country said to be one of the poorest in the world, communities and households exposed to mining activities might afford more educational inputs for their children, which will result in better educational outcomes for the latter. However, the relatively high revenue from the economic boom can also negatively change the aspirations for education of communities and households, which will result in lower educational outcomes. Therefore, both measured educational inputs (tangible) and unmeasured (intangible) ones are potential factors that can explain learning inequalities between GMA and NGMA areas.

3. Potential factors explaining learning inequalities between GMA and NGMA

Studies show that students' learning achievements are associated with family, school, and community factors (Wokadala, 2016; Coulidiati-Kielem, 2016). Family factors often shown as related to students' learning achievements are socioeconomic status (SES), parental education, and having textbooks at home (Cain & Hendryx, 2010; Sikora et al., 2019; Thomson, 2018; Dudaitė, 2016). Literature usually conceptualizes them as family status factors (e.g., income, parental education, social class) and family process factors (e.g., parental involvement, disciplinary factors, parenting) (Bowen et al., 2012; Gubbins & Otero, 2016). Some family status factors are not easily changed (e.g., ethnicity, race), while family process factors are alterable with relevant interventions. Family process factors are usually unmeasured and are likely to have high importance in explaining GMA-NGMA learning achievements differences because evidence suggests that they determine much of the educational decisions of individuals living in GMA (Sanfo, 2019).

Schools are the primary place where learning takes place and thereby school factors influence learning achievements (Ayeni & Adelabu, 2012; Krishnaratne et al., 2013). School resources are related to students' learning achievements (Wokadala, 2016; Ishiguro, 2018). In addition, research has shown that teacher qualifications (e.g., pre-service training, in-service training, and experience) are related to the learning achievements of students (Munawaroh, 2017; Bartilol & Ng"eno, 2016). Better qualified teachers have effective and efficient teaching processes that will make a difference in students' educational achievements. On school management factors, school principals can influence students' learning achievements through the allocation of human and material resources and the creation of a school learning culture (Lortie, 2009; Sanfo, 2020a). Studies on the effect of these factors on students' learning achievements suggest that actions of teachers and principals in GMA are often disrupted (Owusu & Dwomoh, 2012; Adu-Gyamfi, 2014), which implies that they might explain learning achievements differences between GMA and NGMA.

Community factors can influence students' learning achievements through the inputs they provide to schools. Specifically, they can provide schools with the necessary financial and non-financial resources needed to improve teaching programs, and this will likely reflect on students' outcomes (Coly, 2014). Communities can also support school management, known to be related to student performance. A study by Sanfo (2020b) on small-scale gold mining areas showed for example that communities can be a source of school management support and provide financial educational inputs arranged from mining exploitation management, which will transform into better student learning performance. As such, because NGMA does not relatively easily have access to such financial inputs, community factors might explain higher learning achievements for students in GMA.

Many studies have identified characteristics that influence student learning achievements, but being advantaged or disadvantaged in a given educational input does not necessarily explain learning inequalities among groups. There are no typical decomposition studies on learning differentials between GMA and NGMA (to the best of our knowledge), but this study is very close to the traditional geographical (rural-urban) decomposition studies on learning achievements and insights can be borrowed from them. Such studies have demonstrated that students from rural areas perform lower than their counterparts in urban ones, but findings on factors accounting for the learning gap between the two areas are mixed. For example, while some studies support that family factors explain the learning differential between the two types of locations (Ramos et al., 2016), some others argue that school factors explain this learning gap (Cartwright & Allen, 2002; Sullivan et al., 2013). Furthermore, there is a debate on which measured characteristics and unmeasured ones explain the rural-urban learning gap more (Lounkaew, 2013). These mixed findings and debates on the factors explaining geographical learning differentials are particularly interesting with regard to the types of findings that will be revealed for GMA and NGMA.

4. Data and method

4.1 *Data*

This study employs data from the 2014 Program d'Analyse des Systemes Educatifs de la Confemen (PASEC) [Program for the Analysis of CONFEMEN Education Systems]. PASEC measures elementary school students' learning achievements across countries that are dominantly French-speaking ones, and results from the assessment are used to evaluate education effectiveness in participating countries. The data provides information on grades 2 and 6 students and their respective schools and communities. This analysis uses only data on grade-6 students, motivated by some context-specific reasons. Grade-6 is elementary school's final year in the context of Burkina Faso. In that grade, students have to take the national elementary leaving certificate exam which, to some extent, is meant to assess students' overall elementary education learning achievements. This makes it a high-stake grade in the educational system, and education stakeholders pay much attention to it. In addition, this study is interested in child labor as a potential explanatory factor of learning achievements inequalities between the two types of areas studied. It is known that grade-6 students (around 11 years old) are more likely to be involved in labor activities compared to those in grade 2 (around seven years old).

PASEC data is nationally representative of the population of grade-6 students in Burkina Faso. The program used a geographical cluster sampling technique (regions, provinces, communes), followed by probability proportional to size (PPS) sampling which allowed to provide a sample of 200 schools. Within each school, one grade-6 classroom was randomly selected and 20 students were randomly selected within each classroom. However, when a school had only one grade-6 classroom, that classroom was automatically selected and the same process was applied when a classroom had 20 or fewer students. PASEC set a validity benchmark in terms of minimum participation rates for schools and students sampled for the data to be acceptable. This benchmark was satisfied, and as a whole, the data provides information on 3298 students.

4.2 *Variables*

This study uses grade-6 students' reading and mathematics achievements as dependent variables. They are provided as continuous variables by PASEC. The reading achievements variable assesses students' text and document comprehension skills and also their abilities in extracting information from a literary text. However, the variable does not measure students' oral speaking skills. The mathematics achievements variable assesses students' performance in arithmetic, geometry, and measurement. GMA relative to NGMA distributions in reading and mathematics scores is presented in Figure 1. The two panels indicate that the learning achievements differences between the two areas exist and are constant across the whole distribution.

Predictors were selected following previous research showing factors which are likely to be relevant to the context of this study (e.g., Ikeda & Emma, 2014; Scheerens, 2016; Zabaleta, 2011). We used variables such as age, gender, and family socioeconomic factors, as they influence students' learning achievements in the context of Burkina Faso (Coulidiati-Kielem, 2016). However, for originality, this study also uses some "new" variables like child health. This factor is not typically used in studies on factors that affect learning achievements. However, it might be relevant to learning achievements due to children from GMA being exposed to chemicals that might affect their health and make them underperform. As such, this factor might explain geographical differences in learning achievements. Health is multidimensional and difficult to capture by a single variable, so we constructed a health problem synthetic index variable by applying principal component analysis to health-related items. The items used to construct the health indicator variable are presented in Table A-1 in the appendix.

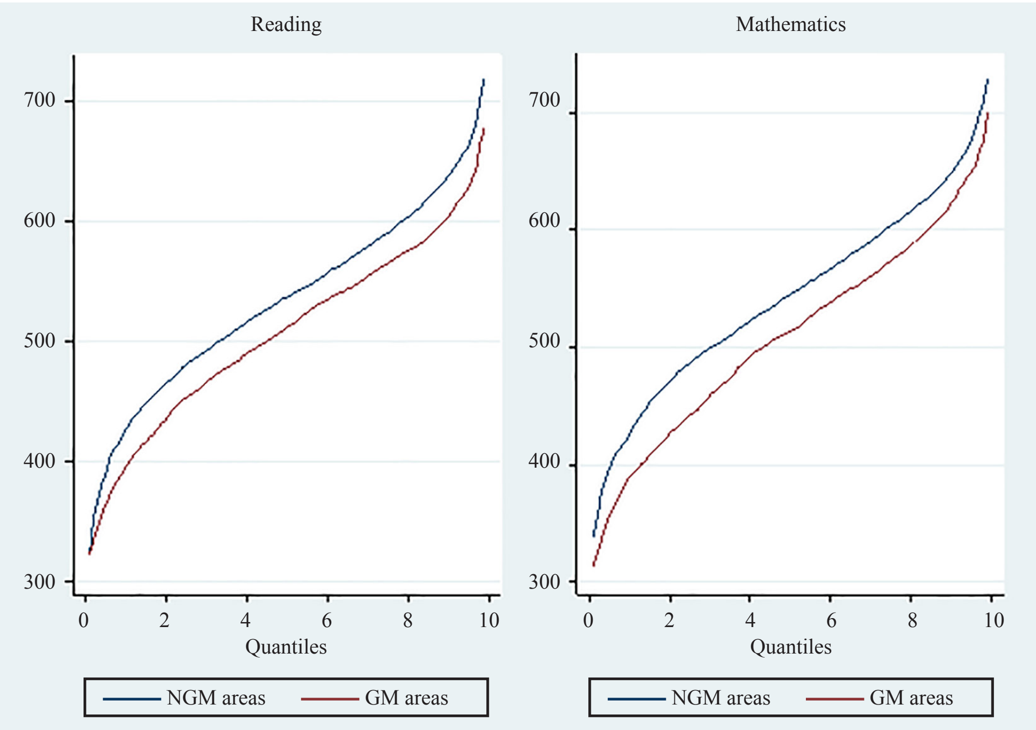

Figure 1. Reading and mathematics distribution by area across quantiles

Some of the variables PASEC provides in its dataset capture latent concepts. The program used item response theory to create these latent variables. For example, it used Rasch one-parameter model with originally dichotomous variables to compute some index variables. The latent variables provided in the dataset and used in this study are family SES (constructed using student household characteristics), community involvement (created based on variables related to community participation in school life), classroom pedagogical resources (created based on the availability of a certain number of pedagogical resources that the teacher can use in the classroom), and teacher perceived social advantages index (created based on the availability of a number of socioeconomic advantages for the teacher). Table 1 presents the summary statistics of the variables used in the study.

Table 1. Summary statistics of variables used in the analysis

	Definition	GMA		NGMA		NGMA-GMA
Number		1273		2025		
Variables		Mean	SD	Mean	SD	
Dependent variables						
Reading score	Student reading score	507.253	80.899	534.662	83.185	27.409***
Math score	Student math score	509.162	89.808	542.533	85.713	33.371***
Student background variables						
Gender	Male = 1	0.496	0.500	0.481	0.500	-0.014
Age	Student age in years	13.361	1.298	13.241	1.306	-0.119***
Grade repetition status	Repeated a grade = 1	0.515	0.500	0.557	0.497	0.0412**
Labor status	Student works = 1	0.267	0.442	0.245	0.430	-0.0219
Health index	Health problems index	-0.003	0.633	0.002	0.608	0.0029
Family SES	Family SES index	48.931	6.613	50.756	7.960	1.824***
School variables						
Classroom resources	Classroom res. index	51.515	11.192	53.713	8.751	2.198***
Teacher pre-serv. training	Has training = 1	0.430	0.495	0.363	0.481	-0.067***
Teacher experience	Years of experience	10.465	4.735	12.258	6.737	1.792***
Teacher absence	Days of absence/week	1.624	2.003	1.141	1.482	-0.483***
Teacher social advantages	Social advantage index	45.402	7.568	47.110	7.609	1.707***
Principal has initial training	Has training = 1	0.177	0.382	0.191	0.393	0.014
Principal experience	Years of experience	8.309	5.010	8.999	6.165	24.850***
School feeding program	Has feeding program = 1	0.918	0.274	0.885	0.318	-0.032***
Community involvement	Community involv. index	51.183	8.181	49.266	10.224	-1.917***
Rural or urban school	Rural = 1	0.304	0.460	0.554	0.497	0.249***
GMA or NGMA (all)	Gold mining = 1	0.386		0.614		

Note: *** p < 0.01, ** p < 0.05, * p < 0.1
GMA = gold mining area NGMA = non-gold mining area.
Source: PASEC 2014 data (own computation). T-test for differences between GMA and non-GMA areas are provided in the column 7 (extreme right side) and for all variables.

4.3 *Gold mining area identification strategy*

The main objective of this paper is to explore what factors account for learning achievements differences between GMA and NGMA, but a readily available administrative division between the two types of areas does not exist in Burkina Faso. Consequently, to identify GMA, we use a common approach from literature (Ahleru et al., 2020; Bazillier & Girard, 2020). It consists of using gold mines Global Positioning System (GPS) coordinates to determine GMA location, to which households or communities (based on their geographical location information) within a certain radius of the location are matched and considered as from GMA. The dataset used does not provide information on the location of students' households, and consequently, communes (the lowest administrative tier) where schools are located were considered as where students live. In the context of Burkina Faso, students usually go to schools located within their respective commune of residence, so this choice does not create issues for the analysis. PASEC 2014 data obtained contains such geographical information of schools, and information containing GPS data of gold mines were obtained for the purpose of the research [GPS coordinates of gold mines were provided by MinEx Consulting, a private gold mining consulting company. The company estimates that the data cover 99% of giant-sized gold deposits, 95% of major deposits, 70% of moderate deposits and 50% of minor deposits, and were up-to-date as of 2015].

Information on gold mines indicates whether a given gold mine extraction method is industrial or artisanal. The study considered students in schools within the radius of 20 km of artisanal gold mines as those from GMA, but it also included students in schools within the radius of 40 km of an industrial gold mine as from GMA. The inclusion of the latter was motivated by previous studies which suggest that industrial mines also attract communities around them; however, when settling, such communities have to respect a given distance set by industrial gold mines (Sanfo, 2020b; Zabsonré et al., 2018). Being around an industrial gold mine makes it easier for communities to engage in activities to look for residual gold left by industrial gold mining companies. The choice of the radiuses is somewhat arbitrary, but this is due to inconsistency in the literature on this choice. For example, some studies use a radius of 100 kms while others use a radius of 80 km or even less (Aragon & Rud, 2013; Loayza & Rigolini, 2016). However, our choice does not cause serious issues because the farther we go away from extraction mines, the weaker their effect on communities (Aragon & Rud, 2013). Moreover, there may be other types of mining activities in NGMA (e.g., zinc), but such mining activities are industrial and do not attract local communities, so they do not have strong implications for education.

4.4 *Analysis*

The strategy of the empirical analysis of this study starts from a traditional Oaxaca-Blinder decomposition (OBD), a technique commonly used in studies on wage inequalities based on gender or race (Blinder, 1973; Oaxaca, 1973). In OBD, a linear education production function of GMA and NGMA is estimated at the mean, and the difference in the outcomes between the two areas is decomposed. Our education production function can be summarized as:

$$A_i = \beta' X_i + \varepsilon_i \tag{1}$$

Where A_i represents the PASEC test achievement of student i, X_i represents a vector that contains student individual, family, and school characteristics. The term ε_i represents the error term and β represents a vector of all associated coefficients and the intercept.

The OBD of the achievement difference between students from *GMA* and *NGMA* can be written as:

$$A^{NGMA} - A^{GMA} = \beta^{NGMA\prime} X_{NGMA} - \beta^{GMA\prime} X_{GMA} \tag{2}$$

Where A^{NGMA} denotes the mean of the PASEC test achievement for children from schools in *NGMA* and A^{GMA} denotes the mean of the same outcome but for children from schools in *GMA*.

Following Jann (2008), the OBD of equation (2) can be re-written as follows:

$$\begin{aligned} A^{NGMA} - A^{GMA} &= \beta^{NGMA\prime} X^{NGMA} - \beta^{GMA\prime} X^{GMA} \\ &= \{(X^{NGMA} - X^{GMA})\}' \beta^* + \{(X^{NGMA})(\beta^{NGMA} - \beta^*)' + (X^{GMA})'(\beta^* - \beta^{GMA})\} \\ &= E + U \end{aligned} \tag{3}$$

Where the first component $E = \{(X^{NGMA} - X^{GMA})\}'\beta^*$ denotes the explained portion of learning achievements difference (explained by differences in observed characteristics). The second component $U = \{(X^{NGMA})(\beta^{NGMA} - \beta^*)' + (X^{GMA})'(\beta^* - \beta^{GMA})\}$ denotes the unexplained part and captures a portion of the difference in the outcome explained by unmeasured educational factors. The term β^* is a reference coefficient vector and is given by the linear combination $\beta^* = \Omega\beta^{NGMA} + (I - \Omega)\beta^{GMA}$.

While OBD can provide important insights into what characteristics determine learning achievements between GMA and NGMA and subsequently the portion of the difference explained by unmeasured educational factors, its estimations are only at the mean value of the learning achievements of the two groups. In other words, OBD does not inform us whether factors explaining learning achievements between the two groups differ across learning achievements distributions. In order to go beyond the limitations of the mean analysis, the study uses re-centered influence function decomposition (RIFD). Similar to OBD, RIFD estimates the education production function of the two groups and then decomposes the difference into the components already presented. However, RIFD estimates across learning achievements distributions. The education production estimation in RIFD uses unconditional quantile regression (UQR) based on the re-centered influence function (RIF) developed by Firpo et al. (2007a; 2007b). RFI is defined as:

$$RIF(Y;\ q_\tau;\ F_Y) = q_\tau + \frac{\tau - 1(Y \le q_\tau)}{f_Y(q_\tau)} \tag{4}$$

Where q_τ is the value of the dependent variable y at quantile τ. $f_Y(q_\tau)$ is the density function of Y at q_τ. $1(Y \le q_\tau)$ is the indicator function and identifies whether the value of the dependent variable Y for an individual is below q_τ.

RIFD mirrors OBD but decomposes the group differences along with distributions of the learning achievements. As such, the initial OBD in equation 3 can be specified to show learning achievements differences between students from schools in GMA and those from schools in NGMA at quantile τ:

$$A_\tau{}^{NGMA} - A_\tau{}^{GMA} = \beta^{NGMA\prime} X_\tau{}^{NGMA} - \beta^{GMA\prime} X_\tau{}^{GMA}$$

$$= \{(X^{NGMA} - X^{GMA})\}'\beta^*{}_\tau + \{(X^{NGMA})(\beta_\tau{}^{NGMA} - \beta^*{}_\tau)' + (X^{GMA})'(\beta^*{}_\tau - \beta^{GMA}{}_\tau)\} \tag{5}$$

Where $\beta^*{}_\tau = \Omega_\tau \beta_{NGMA\tau} + (I - \Omega_\tau)\beta_{GMA\tau}$ still similar to standard OBD, the first component of equation (5) $\{(X^{NGMA} - X^{GMA})\}'\beta^*{}_\tau$ denotes the portion of learning achievements difference attributable to differences in observed characteristics between the two groups (e.g., family and school factors), but at quantile τ. The second component $\{(X^{NGMA})(\beta_\tau{}^{NGMA} - \beta^*{}_\tau)' + (X^{GMA})'(\beta^*{}_\tau - \beta^{GMA}{}_\tau)\}$ denotes the portion of the difference explained by unmeasured factors and can be interpreted as the potential effect of unmeasured educational factors.

A concern that needs to be mentioned is that studies on learning achievements point out the clustering of education systems (e.g., students nested in schools) which, in addition to observed characteristics, have a non-negligible effect on learning achievements (Sanfo, 2020a). When the clustering effect is not accounted for, estimations of standard errors are too small, leading to inconsistency in the estimated results (Hox, 2010). This study accounts for the clustering effect (students within schools) by using school-cluster-adjusted standard errors when fitting the model (Woessmann, 2003). Furthermore, teacher and school principal variables fitted in quadratic terms in order to capture their non-linear characteristics.

5. Results and discussion

5.1 *Student background and school factors that account for learning achievements inequalities between GMA and NGMA*

As already discussed in the analysis section, RIFD decomposes the GMA-NGMA learning achievements differential into explained and unexplained components across the learning achievements distributions. The former is of main interest to the first research objective of this study. A detailed decomposition is used to have a detailed contribution of individual factors which account for the learning achievements differential. Tables 2 and 3 present the results of the detailed decomposition across the learning distributions for reading and mathematics achievements, respectively.

For reading achievements, results of the detailed decomposition reveal that for student background factors, student age accounts for a disadvantage in learning achievements for students in GMA, but only for those from the 50th percentile and above. However, grade repetition explains reading achievements inequalities between students in this area and their peers in NGMA for those at the 70th percentile and above. Moreover, family socioeconomic status (SES) significantly accounts for lower reading achievements differences for students from GMA relative to those from NGMA and across all percentiles.

Table 2. Detailed decomposition of factors accounting for reading achievements inequalities between GMA and NGMA

VARIABLES	Mean	Quantile 10	Quantile 20	Quantile 30	Quantile 40	Quantile 50	Quantile 60	Quantile 70	Quantile 80	Quantile 90
Age	-0.402*	0.695	0.414	-0.185	-0.454	-0.665**	-0.873**	-1.101**	-1.282**	-1.025**
	(0.238)	(0.468)	(0.331)	(0.259)	(0.290)	(0.339)	(0.403)	(0.485)	(0.554)	(0.492)
Grade repetition	0.308	-0.113	0.284	0.241	0.062	0.059	0.277	0.577*	0.686*	1.032*
	(0.204)	(0.310)	(0.269)	(0.237)	(0.200)	(0.189)	(0.226)	(0.328)	(0.377)	(0.535)
Family SES	-2.068***	-1.955**	-1.306*	-1.066*	-1.266**	-2.489***	-2.268***	-2.044***	-1.692**	-2.014***
	(0.525)	(0.883)	(0.702)	(0.619)	(0.593)	(0.656)	(0.644)	(0.648)	(0.658)	(0.777)
Teacher advantages	-1.408***	-2.156***	-1.139*	-1.445**	-0.908*	-0.725	-0.182	-0.276	-0.567	-0.967
	(0.457)	(0.822)	(0.607)	(0.577)	(0.501)	(0.481)	(0.467)	(0.483)	(0.532)	(0.687)
Classroom resources	-4.311***	-6.955***	-4.308***	-3.959***	-3.384***	-3.062***	-3.072***	-2.757***	-3.183***	-3.287***
	(0.923)	(1.577)	(1.039)	(0.926)	(0.839)	(0.793)	(0.798)	(0.778)	(0.819)	(0.926)
Teacher has training	-0.212	1.700**	0.394	0.135	-0.338	-0.762*	-1.150**	-1.112**	-1.003**	-1.063**
	(0.333)	(0.839)	(0.573)	(0.481)	(0.447)	(0.460)	(0.504)	(0.499)	(0.497)	(0.514)
Teacher experience	-3.697***	-10.212***	-5.572***	-4.392***	-2.767**	-2.038**	-2.336**	-2.224**	-1.724*	-3.042***
	(0.907)	(2.155)	(1.504)	(1.263)	(1.119)	(1.020)	(0.977)	(0.972)	(0.945)	(1.036)
Teacher absence	-0.359	3.032***	2.867***	0.927	-0.548	-0.419	-1.141	-1.323*	-2.391***	-2.500***
	(0.539)	(1.162)	(0.996)	(0.866)	(0.823)	(0.765)	(0.751)	(0.751)	(0.802)	(0.881)
Principal experience	1.181**	3.952***	2.667***	2.166***	1.788**	2.010***	1.492**	0.849	-0.836	-1.355
	(0.520)	(1.220)	(0.925)	(0.819)	(0.771)	(0.761)	(0.716)	(0.704)	(0.749)	(0.916)
Feeding program	-0.780**	-0.363	-0.722*	-0.716*	-0.750*	-0.583*	-0.614*	-0.603*	-0.939*	-1.843**
	(0.393)	(0.302)	(0.380)	(0.373)	(0.383)	(0.316)	(0.334)	(0.340)	(0.488)	(0.916)
Comm. involvement	0.981**	1.338	1.560**	0.232	0.476	0.363	0.587	0.347	0.954*	1.799***
	(0.412)	(0.843)	(0.654)	(0.508)	(0.467)	(0.427)	(0.434)	(0.442)	(0.494)	(0.674)

Notes: comm. = community. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Source: PASEC 2014 data (own computation). The coefficients are from RIF decomposition regressions across quantiles and standard errors are in parentheses. The dependent variable is students' reading scores. Each regression for each quantile includes all variables presented in the descriptive statistics, but only statistically significant ones presented in this table.

Table 3. Detailed decomposition of factors accounting for mathematics achievements inequalities between GMA and NGMA

VARIABLES	Mean	Quantile 10	Quantile 20	Quantile 30	Quantile 40	Quantile 50	Quantile 60	Quantile 70	Quantile 80	Quantile 90
Age	0.025	1.177*	0.497	0.190	0.039	-0.229	-0.475	-0.821**	-0.508	-0.450
	(0.192)	(0.610)	(0.366)	(0.272)	(0.250)	(0.266)	(0.314)	(0.409)	(0.326)	(0.328)
Grade repetition	0.325	0.248	0.034	-0.002	0.081	0.171	0.268	0.391	0.630*	0.684*
	(0.217)	(0.344)	(0.258)	(0.220)	(0.215)	(0.226)	(0.245)	(0.277)	(0.360)	(0.396)
Family SES	-1.311***	-1.279	-1.063	-1.408**	-1.624**	-1.110*	-1.225*	-0.760	-0.622	-1.454**
	(0.491)	(0.881)	(0.764)	(0.685)	(0.665)	(0.644)	(0.653)	(0.631)	(0.646)	(0.742)
Teacher advantages	-0.715*	-1.726**	-0.863	-0.549	-0.149	-0.242	0.013	0.321	0.032	-0.316
	(0.426)	(0.842)	(0.627)	(0.546)	(0.522)	(0.523)	(0.543)	(0.565)	(0.574)	(0.655)
Classroom resources	-4.578***	-6.971***	-5.461***	-4.658***	-4.006***	-4.114***	-3.708***	-3.393***	-3.254***	-3.218***
	(0.958)	(1.619)	(1.238)	(1.056)	(0.938)	(0.943)	(0.894)	(0.858)	(0.848)	(0.910)
Teacher has training	-0.169	1.702**	0.658	-0.175	-0.339	-0.847	-0.420	-0.610	-1.171**	-0.728
	(0.348)	(0.868)	(0.600)	(0.499)	(0.482)	(0.517)	(0.476)	(0.482)	(0.540)	(0.463)
Teacher experience	-6.269***	-12.825***	-8.775***	-6.643***	-6.376***	-5.901***	-5.902***	-5.217***	-3.809***	-3.040***
	(1.092)	(2.408)	(1.725)	(1.414)	(1.332)	(1.280)	(1.264)	(1.194)	(1.108)	(1.022)
Teacher absence	0.359	3.675***	3.336***	1.272	0.358	-0.353	-0.709	-1.164	-1.318	-1.582**
	(0.575)	(1.231)	(1.069)	(0.913)	(0.854)	(0.834)	(0.838)	(0.839)	(0.838)	(0.804)
Principal experience	2.450***	3.611***	3.949***	3.490***	3.141***	2.558***	2.613***	2.120**	1.191	0.451
	(0.680)	(1.246)	(1.130)	(0.993)	(0.936)	(0.870)	(0.880)	(0.831)	(0.776)	(0.820)
Feeding program	-0.794**	-0.719*	-0.873*	-0.760*	-0.563*	-0.569*	-0.651*	-0.652*	-0.849*	-1.394*
	(0.401)	(0.423)	(0.455)	(0.400)	(0.320)	(0.326)	(0.360)	(0.366)	(0.455)	(0.715)
Comm. involvement	0.566	0.655	0.827	0.131	0.003	-0.090	-0.103	0.090	-0.013	1.107*
	(0.401)	(0.832)	(0.659)	(0.532)	(0.495)	(0.482)	(0.482)	(0.484)	(0.499)	(0.605)

Note: comm. = community. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
Source: PASEC 2014 data (own computation). The coefficients are from RIF decomposition regressions across quantiles and standard errors are in parentheses. The dependent variable is students'mathematics scores. Each regression for each quantile includes all variables presented in the descriptive statistics, but only statistically significant ones presented in this table.

On school factors, results show that GMA-NGMA learning achievements inequalities can be explained by teacher perceived social advantages but only for students from the 10th to the 40th percentile, with a disadvantage for students from GMA. A disadvantage in learning achievements for students in GMA is also explained by classroom pedagogical resources and teacher experiences factors but across all learning achievements distribution. Similar results are also observed on the teacher training status, but for students from the 50th percentile and above. Teacher absence shows mixed results for the lower and upper tails of the reading achievements distribution. On the lower tail (10th and 20th percentiles), this factor accounts for an advantage for students in GMA, but a disadvantage for them on the upper tail (70th to 90th percentiles).

Furthermore, school principal experience significantly accounts for learning achievements inequalities between GMA and NGMA. This factor gives an advantage for students in the former for those from the 10th to the 60th percentiles. Oppositely, the factor school feeding program accounts for learning achievements disadvantages for students from GMA, across almost the whole learning achievements distribution. Community involvement in school life explains learning achievements differences between GMA and NGMA, but results are mixed, as they are significant only at percentiles 20, 80, and 90. However, students from GMA are advantaged from the perspective of this factor.

Results on the student background and school factors that account for mathematics achievements inequalities between GMA and NGMA are fairly similar to those on reading achievements in terms of the factors statistically explaining these inequalities (Table 3). Grade repetition explains the learning differential between the two areas for students on the upper tail, similar to age, but the latter also seems to explain differences for students at the lower end of the tail (10th percentile). The family SES factor also accounts for differences for students from the 30th to 90th percentiles (except percentiles 70 and 80).

On school factors, teacher advantages explain GMA-NGMA learning achievements differences only for students at percentile 10, significant at 5% level. Similarly, teacher training is significant at 5% level but for students at percentiles 10 and 80. Teacher absence explains mathematics achievements inequalities between the two areas for the students at the lower tail (10th and 20th) and the upper tail (90th) of the distribution. Classroom resources and teacher experience factors explain a learning disadvantage for students in GMA across all the distribution, and results are statistically significant at 1% level for both factors. Likewise, having a school feeding program explains mathematics achievements inequalities for all students across the learning achievements distribution and is statistically significant at 10% level. On school principals, their experience explains an advantage for students from GMA and from the 10th to the 60th percentiles, statistically significant at 1% level. Community factors only explain GMA-NGMA mathematics achievements for students at the 90th percentile, statistically significant at 10% level.

5.2 *The proportion of GMA-NGMA learning achievements inequalities explained by unmeasured educational factors*

The unexplained component of the RIFD captures the portion of the learning achievements differential which is not explained by differences in measured characteristics. The unexplained portion was used to compute the proportion of learning achievements inequalities explained by unmeasured factors at the mean and across learning achievements distributions. Results for both the mean and percentiles are presented in Tables 4 and 5 for reading and mathematics achievements, respectively.

From a mean analysis perspective of reading achievements, results indicate that unmeasured factors account for 27.68% of the reading achievements differential between GMA and NGMA, significant at 5% level. Disaggregating the decomposition into learning achievements percentiles reveals an increase of the proportion of the GMA-NGMA reading achievements difference to 34.50% at the 10th percentile, significant at 10% level. Furthermore, the increase in this proportion is consistent as we move to the 20th percentile (41.16%) but significant at a 5% level for this percentile. However, from the 30th to the 70th percentiles, reading achievements inequalities explained by unmeasured educational factors are not statistically significant. This might explain why their respective proportions decrease to be between 14.58% and 21%. Consequently, across these levels of the distribution, it can be said that unmeasured educational factors do not account for differences in students' reading achievements, and it implies that all differences are explained by measured characteristics. The upper tail of the learning distributions shows a similar pattern to the lower tail in the proportion of GMA-NGMA reading achievements differential explained by unmeasured educational factors. Specifically, unmeasured educational factors account for 34.58% of the differential for students at the 80th percentile, statistically significant at 5% level. This proportion increases to 57.33% for students at the 90th percentile, statistically significant at 1% level.

Table 4. Reading achievements inequalities explained by unmeasured education factors

Estimate	GMA	NGMA	Difference	Explained	Unexplained	Unexplained (%)
Mean	507.715***	534.622***	-26.907***	-19.458***	-7.449**	27.684
	(2.295)	(1.887)	(2.971)	(1.884)	(2.914)	
Perc. 10	397.505***	428.885***	-31.380***	-20.554***	-10.826*	34.500
	(4.181)	(3.441)	(5.415)	(3.133)	(5.974)	
Perc. 20	436.162***	467.125***	-30.963***	-18.217***	-12.746**	41.165
	(3.904)	(2.800)	(4.804)	(2.431)	(5.086)	
Perc. 30	467.758***	493.444***	-25.686***	-21.940***	-3.746	14.584
	(3.358)	(2.569)	(4.228)	(2.244)	(4.491)	
Perc. 40	490.577***	516.820***	-26.243***	-21.930***	-4.313	16.435
	(3.248)	(2.451)	(4.069)	(2.184)	(4.233)	
Perc. 50	512.596***	537.549***	-24.952***	-21.157***	-3.795	15.209
	(3.206)	(2.368)	(3.986)	(2.147)	(4.042)	
Perc. 60	535.468***	558.030***	-22.562***	-19.606***	-2.956	13.102
	(2.983)	(2.357)	(3.802)	(2.174)	(3.825)	
Perc. 70	555.230***	579.150***	-23.920***	-18.898***	-5.022	20.995
	(2.802)	(2.416)	(3.700)	(2.241)	(3.718)	
Perc. 80	575.846***	603.035***	-27.189***	-17.787***	-9.402**	34.580
	(2.779)	(2.548)	(3.770)	(2.371)	(3.720)	
Perc. 90	604.801***	638.518***	-33.718***	-14.385***	-19.333***	57.337
	(3.229)	(2.998)	(4.406)	(2.798)	(4.320)	

Notes: Perc. = percentile *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ standard errors in parentheses
Source: PASEC 2014 data (own computation).
Numbers in columns GMA, NGMA, Explained, Unexplained are obtained from RIF regressions at the mean or the respective quantiles used. Difference = mean GMA – mean NGMA; Percentage unexplained = (unexplained/difference) × 100.

Figure 2 is a visual representation of the proportion of GMA-NGMA reading achievements differential explained by unmeasured educational factors as already presented. The figure shows a low and inconsistent proportion of differential accounted for by unmeasured factors for the percentiles that are not statistically significant (30^{th} to the 70^{th}). However, on both tails of the reading achievements distributions, i.e., the statistically significant portion, the proportion of reading achievements differential explained by unmeasured educational factors is relatively higher. On the upper tail, these factors explain more than half of the reading achievements difference between GMA and NGMA.

From a mean analysis perspective of mathematics achievements, results indicate that about half (50.03%) of achievements differential between GMA and NGMA is accounted for by unmeasured educational factors, statistically significant at 1% level. Similarly, the disaggregation into percentiles indicates that about half of the differential is explained by unmeasured educational factors, from percentiles 10 to 30, statistically significant at 1% level. However, the proportion explained by unmeasured educational factors is between 35.70% and 48.50% from percentiles 40 to 70.

This proportion increases to 56.46% and 46.11% for percentiles 80 and 90, respectively.

Table 5. Mathematics achievements inequalities explained by unmeasured education factors

Estimate	GMA	NGMA	Difference	Explained	Unexplained	Unexplained (%)
OLS (Mean)	509.679***	542.840***	-33.161***	-16.569***	-16.592***	50.035
	(2.551)	(1.948)	(3.209)	(1.847)	(3.242)	
Perc. 10	390.876***	428.384***	-37.507***	-18.466***	-19.042***	50.769
	(3.946)	(3.629)	(5.362)	(3.170)	(6.079)	
Perc. 20	427.699***	471.890***	-44.190***	-18.811***	-25.380***	57.434
	(3.918)	(3.015)	(4.944)	(2.630)	(5.334)	
Perc. 30	460.294***	500.132***	-39.838***	-18.024***	-21.814***	54.757
	(4.222)	(2.649)	(4.984)	(2.279)	(5.189)	
Perc. 40	492.622***	523.028***	-30.405***	-18.139***	-12.267**	40.345
	(3.822)	(2.578)	(4.610)	(2.189)	(4.789)	
Perc. 50	514.045***	545.821***	-31.776***	-17.741***	-14.035***	44.169
	(3.485)	(2.582)	(4.338)	(2.240)	(4.510)	
Perc. 60	539.142***	567.443***	-28.301***	-18.196***	-10.104**	35.702
	(3.308)	(2.588)	(4.201)	(2.231)	(4.336)	
Perc. 70	560.710***	590.787***	-30.077***	-15.487***	-14.590***	48.509
	(3.238)	(2.579)	(4.140)	(2.276)	(4.277)	
Perc. 80	586.417***	616.866***	-30.449***	-13.256***	-17.193***	56.465
	(3.372)	(2.593)	(4.253)	(2.278)	(4.383)	
Perc. 90	623.290***	648.146***	-24.856***	-13.393***	-11.463**	46.118
	(3.896)	(2.865)	(4.836)	(2.564)	(4.776)	

Note: Perc. = percentile *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ standard errors in parentheses
Source: PASEC 2014 data (own computation).
Numbers in columns GMA, NGMA, Explained, Unexplained are obtained from RIF regressions at the mean or the respective quantiles used. Difference = GMA – NGMA; Percentage unexplained = (unexplained/difference) × 100.

Figure 3 is a visual representation of the proportion of mathematics achievements differential between GMA and NGMA explained by unmeasured educational factors. The figure shows higher proportions of mathematics achievements differential accounted for by unmeasured factor for students on the lower and the upper tails of the distribution. However, for students who are not on the two tails of the distribution, unmeasured education factors still account for an important proportion of mathematics achievements differences between GMA and NGMA.

We examined the robustness of the above results by using a different radius to define GMA. Specifically, for a new definition of GMA, we used a radius of 10 kilometers for artisanal gold mines and a radius of 20 kilometers for industrial gold mines. After redefining GMA, we fitted a decomposition regression at the mean to estimate the gap between GMA and NGMA and also the factors that explain it. The results of this decomposition regression are presented in Figure 4. This approach of reducing the radiuses we initially used showed that our estimations do not almost change. In other words, the gap between GMA and NGMA is still observed and the factors that explain the gap are almost the same. This robustness is observed both for reading and mathematics achievements.

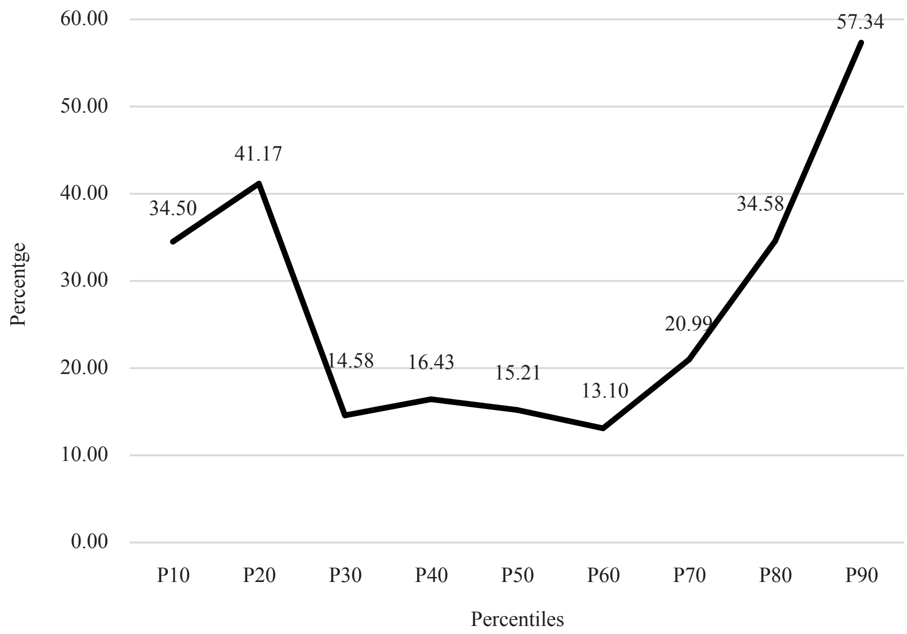

Figure 2. Visual representation of the proportion of reading achievements explained by unmeasured education factors

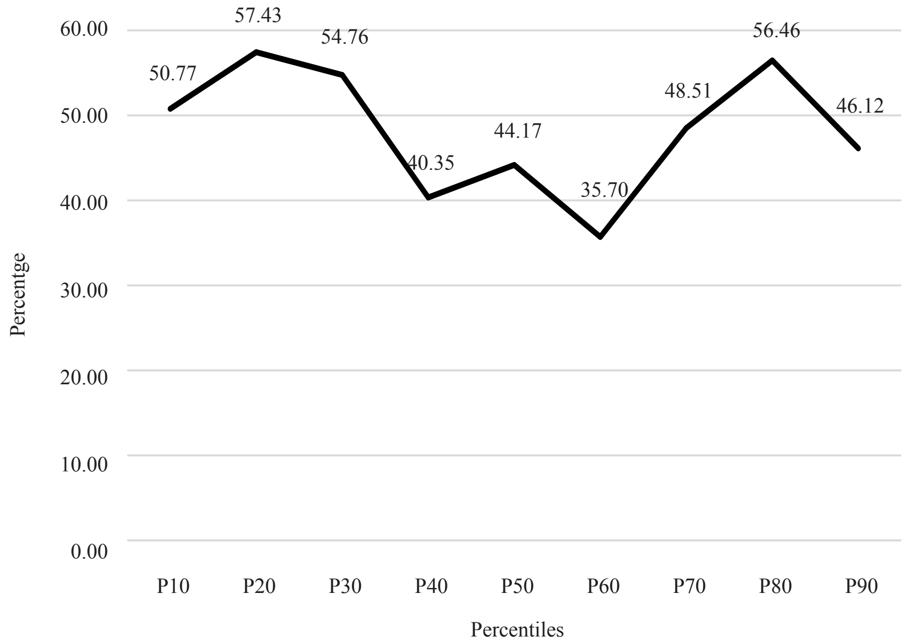

Figure 3. Visual representation of the proportion of mathematics achievements explained by unmeasured education factors

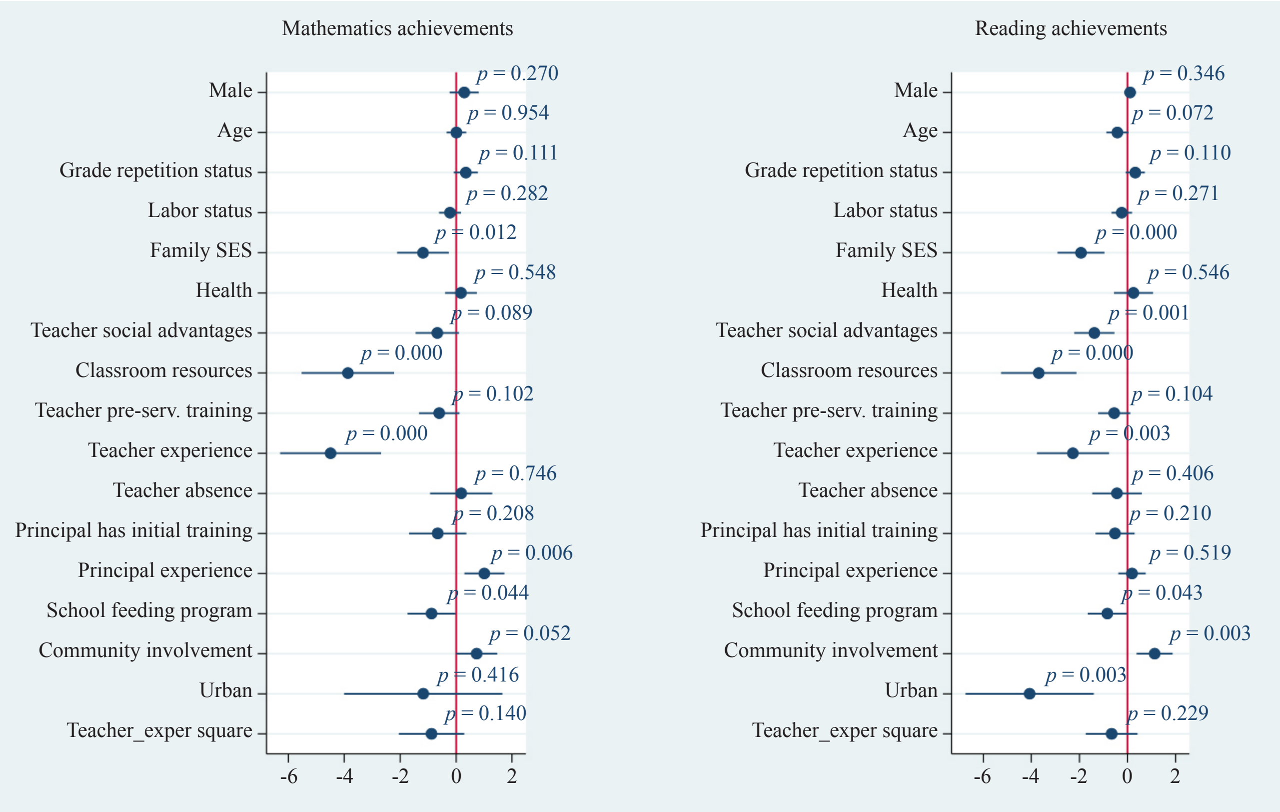

Figure 4. Factors explaining GMA-NGMA learning achievements gaps

Note: Radius of 10 kms of artisanal gold mines and 20 kms of industrial gold mines were used to define GMA

5.3 *Discussion*

From an overall perspective, the findings in this study are surprising because they indicate that most factors accounting for the GMA-NGMA learning achievements inequalities are school-related ones. This suggests that eradicating differences in the significant factors is expected to reduce GMA-NGMA learning achievements equalities. The detailed decomposition revealed that age explains learning achievements differences for main students on the upper tail of the distribution. In GMA, the older the children, the higher the likelihood they get engaged in labor activities and neglect their studies (Botchwey & Crawford, 2016). This is likely to put them at a disadvantage relative to their peers from NGMA. Moreover, repeating a grade explains learning achievements differentials for students on the upper tail of the distribution. Not surprisingly, the study found that SES explains the GMA-NGMA learning achievements difference across all percentiles. This is consistent with some studies on other developing countries suggesting that households in GMA might earn a decent income but still live in poverty and poor conditions (Kim & Lin, 2017; van der Ploeg, 2011). Unfortunately, this also seems to be the case in Burkina Faso, where households in GMA may not be willing to invest in the education of their children because they consider it to be a long and uncertain investment relative to the potential of getting rich overnight (Maiga, 2015; Soré & Maiga, 2015). As such, their children will have a disadvantage in family educational investments, which likely explains related lower learning achievements we found.

Furthermore, the study revealed that school factors also explain GMA-NGAM learning achievements differences. Looking at teacher factors, it found that teacher advantages account for differences in learning achievements among students on the lower tail of the distribution. This finding seems to confirm the importance of teacher advantages for students' learning achievements (Michaelowa, 2002). Students at the lower tail of the learning distribution are usually those that have learning difficulties and need extra teacher support (Aldridge et al., 2012), but disadvantaged teachers will likely not provide this extra support, worsening the learning conditions of these struggling students. Similarly, classrooms resources account for lower learning for all students in GMA. A plausible explanation of this finding is that many GMA in Burkina Faso are rural ones, putting them at a disadvantage in the provision of educational resources. Burkinabe government seems to confirm this as it is recently trying to solve such issues by identifying communes to prioritize in educational resources investments; however, the implementation of the strategy seems to be suffering from many gaps (Sanfo, 2020a). Teacher training accounts for lower learning achievements of students from these areas, mainly for those at the 50th percentile and above. As already discussed, many GMA are rural ones, and it is known that rural areas are less likely to have trained teachers. This finding confirms the importance of qualified teachers and highlights that related differences will likely bring about inequalities. Teacher absence explains GMA-NGMA learning achievements inequalities, but surprisingly students from GMA have an advantage, though it is only for those on the lower tail of the distribution. Teachers from NGMA, dominantly from cities, are more likely than those from GMA to be absent from schools because they may have extra activities to do (e.g., private tutoring). This finding might be explained by teachers in GMA focusing more on low-performers than other students. Consequently, we see that teacher absence explains learning achievements inequalities between GMA and NGMA differently for students on the upper tail of the distribution. This seems to confirm that teachers are more likely not to devote much or quality time to these students on the upper tail of the distribution, resulting in them performing lower than their peers.

School principal experience explains GMA-NGMA students' learning achievements differences, though not for those at the upper tail of the distribution. Evidence has shown that principals' characteristics are linked with students' achievements but their effectiveness also depends on community support (Sanfo, 2020a). Experienced principals are more likely to be efficient in their work, at the same time encouraging communities to support them, and this seems to be the case of Burkinabe communities in GMA. From this perspective, community involvement in school life also favors students in GMA. However, this study found that students in GMA have a disadvantage in relation to the school feeding program. This might be explained by the fact that it is common for schools in GMA and other remote areas to devote school feeding programs to community women who might not be knowledgeable on children's nutrition (Kaboré, 2019). Since what students eat may affect their learning achievements (Neumann et al., 2007), those in GMA will subsequently perform lower than their peers in other areas who are better off, explaining what this study found.

The study found that unmeasured educational factors account for a non-negligible proportion of GMA-NGMA learning achievements differences, and the proportion is higher for students on the lower and upper tails of the distribution. This suggests that only reducing differences in tangible factors is not enough to reduce learning achievements inequalities between the two areas. Also, there are heterogeneous student factors that need to be considered when trying to understand the importance of unmeasured educational factors for learning achievements, mostly for students on the two tails of the distribution. For example, teachers may have higher expectations for high-performing students and will then devote much more efforts to them, "forgetting" low-performing ones. In such a scenario, educational investments need to target the correction of teacher perception and subsequent behavior in relation to students. Moreover, households in GMA might not have high expectations when their children have a low performance. Such households are likely not to provide the investments in educational inputs necessary for their children to succeed in school even if they can afford these inputs. In this case, there is a need for interventions that will change the perception and aspirations that households have for education.

As a whole, the study reveals that measured and unmeasured factors explain learning achievements inequalities. However, the study found that the importance of unmeasured factors is different between reading and mathematics achievements. The two subjects belong to different learning domains, so differences in their relations to learning achievements are often found in macro-level studies. Nevertheless, it might be interesting to investigate, through micro-level studies, why such differences are found. The findings of our study remind us that artisanal gold mining in developing countries does not have a negative impact on only the environment, it also has a negative impact on education and many factors contribute to reducing or worsening this impact. Some of these factors are "unexplained" because they are challenging to capture through a quantitative study like this one. Qualitative studies could be more appropriate to capture the factors which quantitative studies find challenging. Education authorities in Burkina Faso and education practitioners seem to be more and more interested in educational issues in GMA. Let us hope that this interest will contribute to more research to comprehensively unpack the factors that explain these issues.

6. Conclusion

This study examined tangible and intangible factors which account for learning achievements inequalities between GMA and NGMA. To do so, it employed RIFD with PASEC 2014 data for analysis across distributions. The analysis revealed that most of the learning achievements inequalities between the two types of areas are accounted for by school factors and not much student background ones, consistent across learning achievements distributions for mathematics and reading. Moreover, it found that intangible factors account for a non-negligible proportion of the inequalities, and this proportion is higher for students on the lower and upper tails of the learning achievements distribution.

This study reveals that we can borrow from explanations on the factors accounting for rural-urban learning achievements inequalities to explain GMA-NGMA achievements inequalities. Many GMA are in remote rural areas, making them worse-off in the provision of educational inputs. This results in schools in these areas not having the necessary educational resources to function appropriately and will therefore have lower learning achievements for their students. The deficit discourse seems applicable in this case. However, the study reminds us that community involvement and school principals can contribute to giving an advantage in learning achievements to GMA. As such, opposite to what some studies suggest, natural resources are not completely a curse for communities living around GMA, since these activities might provide resources that can give them an advantage over their peers from NGMA.

This study sheds light on factors explaining GMA-NGMA learning achievements inequalities, but its limitations need to be noted. First, it is limited in time and space, as the data used is cross-sectional and covers only one country. Using cross-sectional data and RIFD restricts the study to show correlation and not causality. Therefore, our findings are not necessarily generalizable to other countries or over time and should not be interpreted as causality. Second, the unexplained component in RIFD might be influenced by the choice of the base group when using dummies, but this could not be addressed because there are currently no solutions to this limitation of the RIFD method. Similarly, there may be endogeneity in some of the variables used. RIFD seems not to be able to address this potential endogeneity, so we recommend further studies that can address these issues and investigate the consistency of our findings. Third, student mobility could not be controlled for due to the lack of a variable indicating how many years students have been in their respective schools. In other words, some students considered as from GMA may have just moved to live in these areas (with their families) and will likely differ from the other students who have been there for many years. In such a case, the findings may be affected by endogenous migration. These limitations are potential issues that future studies may need to address.

From an educational policy perspective, to reduce the “new” geographic learning achievements inequalities, evidence from this study suggests that Burkina Faso needs to create relevant reforms and practices by adopting supply-side measures which will improve tangible schools inputs in GMA. However, only improving tangible inputs in these areas is not enough to reduce the inequalities. It is also important to address intangible educational characteristics, mostly in relation to low-performing and high-performing students. At least, potential educational interventions need to find a balance in investment between tangible educational characteristics which improve learning conditions and intangible educational characteristics which improve perceptions and behaviors in favor of learning. This entails a need for studies to investigate the intangible characteristics which explain these geographic learning achievements inequalities. Moreover, this study underlines the need for Burkinabe gold mining measures which will invest more gold revenues in public goods such as education, for example through Corporate Social Responsibility, as compensation for the negative externalities of mining activities.

Conflict of interest

The authors declare no competing financial interest.

References

Adu-Gyamfi, E. (2014). The effect of illegal mining on school attendance and academic performance of junior high school students in Upper Denkyira West district of Ghana. *Journal of Education & Human Development, 3*(1),

523-545. http://jehdnet.com/journals/jehd/Vol_3_No_1_March_2014/28.pdf
Ahleru, P., Baskaran, T., & Bigsten, A. (2020). Gold mining and education: A long-run resource curse in Africa? *The Journal of Development Studies, 59*(9), 1745-1762. https://doi.org/10.1080/00220388.2019.1696959
Aldridge, J. M., Afari, E., & Fraser, B. J. (2012). Influence of teacher support and personal relevance on academic self-efficacy and enjoyment of mathematics lessons: a structural equation modeling approach. *Alberta Journal of Educational Research, 58*(4), 614-633. https://doi.org/10.1037/t38960-000
Aragon, F. M., & Rud, J. P. (2013). Natural resources and local communities: Evidence from a peruvian gold mine. *American Economic Journal: Economic Policy, 5*(2), 1-25. https://doi.org/10.1257/pol.5.2.1
Ayeni, A. J., & Adelabu, M. A. (2012). Improving learning infrastructure and environment for sustainable quality assurance practice in secondary schools in Ondo state, south-west, Nigeria. *International Journal of Research Studies in Education, 1*(1), 61-68. https://doi.org/10.5861/ijrse.2012.v1i1.20
Azano, A. P., & Biddle, C. (2019). Disrupting dichotomous traps and rethinking problem formation for rural education. *The Rural Educator, 40*(2), 4-11. https://doi.org/10.35608/ruraled.v40i2.845
Bartilol, J. S., & Ng"eno, K. J. (2016). Influence of selected teacher factors on pupils' academic performance in public primary schools in Keiyo south sub-county, Kenya. *IOSR Journal of Humanities and Social Science, 21*(6), 35-44.
Bazillier, R., & Girard, V. (2020). The old digger and the machine. Evidence on the distributive effect of the artisanal and industrial gold rushes in Burkina Faso. *Journal of Development Economics, 143*, 102411. https://doi.org/10.1016/j.jdeveco.2019.102411
Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources, 8*(4), 436-455. https://doi.org/10.2307/144855
Botchwey, G., & Crawford, G. (2016). Impact of small-scale gold mining on education and livelihoods in Ghana. *International Journal of Educational Leadership, 7*(1), 1-25. https://www.researchgate.net/publication/305266977_Botchwey_G_and_Crawford_G_2016_Impact_of_Small-scale_Mining_on_Education_and_Livelihoods_in_Ghana
Bowen, G. L., Hopson, L. M., Rose, R. A., & Glennie E. J. (2012). Students' perceived parental school behavior expectations and their academic performance: A longitudinal analysis. *Family Relations, 61*(2), 175-191. https://doi.org/10.1111/j.1741-3729.2011.00695.x
Cain, L., & Hendryx, M. S. (2010). Learning outcomes among students in relation to west virginia coal mining: An environmental riskscape approach. *Environmental Justice, 3*(2), 71-77. https://doi.org/10.1089/env.2010.0001
Cartwright, F., & Allen, M. K. (2002). *Understanding the rural-urban reading gap*. Ottawa, Canada: authority of the Minister responsible for Statistics Canada. https://files.eric.ed.gov/fulltext/ED505294.pdf
Coly, A. (2014). *Les associations de parents d'élèves a travers l'histoire de l'education au sénégal: Pour une nouvelle définition de la politique educative: Le cas de la casamance [Parents' associations throughout the history of education in Senegal: For a new definition of educational policy: The case of casamance]*. Doctoral dissertation, Université de Limoges, Limoges. https://tel.archives-ouvertes.fr/tel-01081292/document
Coulidiati-Kielem, J. (2016). Analyse transversale de l'effet des facteurs familiaux sur les acquis scolaires au primaire: résultats des évaluations nationales au Burkina Faso [Cross-sectional analysis of the effect of family factors on learning achievements in primary school: results of national assessments in Burkina Faso]. *Evaluer-Journal international de Recherche en Education et Formation, 2*(3), 11-31. https://journal.admee.org/index.php/ejiref/article/view/99
Dudaitė, J. (2016). Impact of socio-economic home environment on student learning achievement. *Independent Journal of Management & Production, 7*(3), 854-871. https://www.redalyc.org/pdf/4495/449546894007.pdf
Dziechciarz-Duda, M., & Król, A. (2013). On the non-monetary benefits of tertiary education. *Econometrics, 3*(41), 78-94. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2442761
Firpo, S., Fortin, N. M., & Lemieux, T. (2007a). *Decomposing wage distributions using recentered influence function regressions*. Working Paper. Department of Economics, University of British Columbia. https://users.nber.org/~confer/2007/si2007/LS/fortin.pdf
Firpo, S., Fortin, N. M., & Lemieux, T. (2007b). *Unconditional quantile regressions*. NBER Technical Working Paper No. 339. Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/papers/t0339
Gubbins, V., & Otero, G. (2016). Effect of the parental involvement style perceived by elementary school students at home on language and mathematics performance in chilean schools. *Educational Studies, 42*(2), 121-136. https://doi.org/10.1080/03055698.2016.1148586
Hanushek, E. A. (2014). Is location fate? Distributional aspects of schooling. In G. K. Ingram & D. A. Kenyon (Eds.), *Education, Land, and Location* (pp. 25-61). Cambridge, MA: Lincoln Institute of Land Policy. http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%202014%20IsLocationFate%20%281%29.pdf
Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic*

Literature, 46(3), 607-668. https://doi.org/10.1257/jel.46.3.607

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge. https://psycnet.apa.org/record/2010-10998-000

Ikeda, M., & Emma, G. (2014). Grade repetition: A comparative study of academic and non-academic consequences. *OECD Journal: Economic Studies, 2013*(1), 269-315. https://doi.org/10.1787/eco_studies-2013-5k3w65mx3hnx

Ishiguro, K. (2018). Determinants of learning achievements: Empirical analysis of seven schools in Cambodian primary school. *Journal of Education and Learning, 7*(2), 174-183. http://doi.org/10.5539/jel.v7n2p174

Jann, B. (2008). The blinder-oaxaca decomposition for linear regression models. *The Stata Journal, 8*(4), 453-479. https://doi.org/10.1177/1536867X0800800401

Kaboré, A. (2019, January 7). *Cantine scolaire au Burkina Faso: évolution d'une politique à la croisée des chemins [School canteen in Burkina Faso: evolution of a policy at a crossroads]*. WakatSera. https://www.wakatsera.com/cantine-scolaire-au-burkina-faso-evolution-dune-politique-a-la-croisee-des-chemins/

Kim, D. H., & Lin, S. C. (2017). Natural resources and economic development: New panel evidence. *Environmental and Resource Economics. 66*, 363-391. https://doi.org/10.1007/s10640-015-9954-5

Konstantopoulos, S., Li, W., Miller, S., & van der Ploeg, A. (2019). Using quantile regressionto estimate intervention effects beyond the mean. *Educational and Psychological Measurement, 79*(5), 883-919. https://doi.org/10.1177/0013164419837321

Kouraogo, P., & Dianda, A. Y. (2008). Education in Burkina Faso at Horizon 2025. *Journal of International Cooperation in Education, 11*(1), 23-38. https://cice.hiroshima-u.ac.jp/wp-content/uploads/publications/Journal11-1/11-1-3.pdf

Krishnaratne, S., White, H., & Carpenter, E. (2013). *Quality education for all children? What works in education in developing countries*. Working Paper 20. New Delhi: International Initiative for Impact Evaluation. https://ideas.repec.org/p/ris/iiierp/0000_000.html

Loayza, N., & Rigolini, J. (2016). The local impact of mining on poverty and inequality: Evidence from the commodity boom in Peru. *World Development, 84*, 219-234. https://doi.org/10.1016/j.worlddev.2016.03.005

Lortie, D. C. (2009). *School principal managing in public*. Chicago: The University of Chicago Press. https://doi.org/10.7208/chicago/9780226493503.001.0001

Lounkaew, K. (2013). Explaining urban-rural differences in educational achievement in Thailand: Evidence from PISA literacy data. *Economics of Education Review, 37*, 213-225. https://doi.org/10.1016/j.econedurev.2013.09.003

Maiga, A. (2015). *De l'or a l'argent: Analyse de la vie d'orpailleur au Burkina Faso [From gold to money: An analysis of the life of an artisanal gold miner in Burkina Faso]*. Revue DEZAN. pp. 251-267.

Michaelowa, K. (2002). *Teacher job satisfaction, studentachievement, and the cost of primary education in Francophone Sub-Saharan Africa*. HWWADiscussion Paper, No. 188, Hamburg Institute of International Economics (HWWA), Hamburg. https://www.ipz.uzh.ch/dam/jcr:ffffffff-ff66-65b5-ffff-ffffa53a09bc/_7.pdf

MoE (Ministere de l'Education Nationale et le L'alphabetisation). (2015). *Orpaillage et descolarisation dans les regions du centre-nord, du Sahel et du sud-ouest [Artisanal gold mining and schooling in the center-north, Sahel and south-west regions*]. Rapport D'etude: Ouagadougou, Burkina Faso: Ministere de l'Education Nationale et le L'alphabetisation.

Munawaroh. (2017). The influence of teaching methods and learning environment to the student's learning achievement of craft and entrepreneurship subjects at vocational high school. *International Journal of Environmental & Science Education, 12*(4), 665-678. https://files.eric.ed.gov/fulltext/EJ1144859.pdf

Myhr, A., Lillefjell, M., Espnes, G. A., & Halvorsen, T. (2017). Do family and neighbourhood matter in secondary school completion? A multilevel study of determinants and their interactions in a life-course perspective. *PLoS ONE, 12*(2), 1-21. https://doi.org/10.1371/journal.pone.0172281

Neumann, C. G., Murphy, S. P., Gewa, C., Grillenberger, M., & Bwibo, N. (2007). Meat supplementation improves growth, cognitive, and behavioral outcomes in Kenyan children. *The Journal of Nutrition, 137*(4), 1119-1123. https://doi.org/10.1093/jn/137.4.1119

Nieuwenhuis, J., & Hooimeijer, P. (2016). The association between neighbourhoods and educational achievement, a systematic review and meta-analysis. *Journal of Housing and the Built Environment, 31*(2), 321-347. https://doi.org/10.1007/s10901-015-9460-7

Oaxaca, R. L. (1973). Male-female wage differentials in urban labor markets. *International Economic Review, 14*(3), 693-709. https://www.jstor.org/stable/2525981

Ouattara, F., & Sore, Z. (2016). Des stratégies des parents d'élèves pour un enseignement primaire de qualité [Parents' strategies for quality primary education]. *African Education Development Issues, 7*, 29-49.

Ouma, D. H., Ting, Z., & Pesha, J. C. (2017). Analysis of the socio-economic factors that contribute to children school dropout in artisanal small-scale gold mining communities of Tanzania. *Journal of Education and Practice, 8*(14),

71-78. https://files.eric.ed.gov/fulltext/EJ1143926.pdf
Owusu, E. E., & Dwomoh, G. (2012). The Impact of Illegal Mining on the Ghanaian Youth: Evidence. *Research on Humanities and Social Sciences, 2*(6), 2225-0484. https://www.iiste.org/Journals/index.php/RHSS/article/viewFile/2383/2382
Paré-Kaboré, A. (2012). Disparities in primary teaching and pedagogical innovation in Burkina Faso. *Revue Internationale D'éducation de Sèvres, 59*, 71-82. https://doi.org/10.4000/ries.2274
Pelinescu, E. (2015). The impact of human capital on economic growth. *Procedia Economics and Finance, 22*, 184-190. https://doi.org/10.1016/S2212-5671(15)00258-0
Ramos, R., Duque, J. C., & Nieto, S. (2016). Decomposing the rural-urban differential in student achievement in colombia using PISA microdata. *Estudios de Economía Aplicada, 34*(2), 379-411. https://www.redalyc.org/pdf/301/30146038005.pdf
Rau, T., Urzua, S., & Reyes, L. (2015). Early exposure to hazardous waste and academic achievement: Evidence from a case of environmental negligence. *Journal of the Association of Environmental and Resource Economists, 2*(4), 527-563. https://doi.org/10.1086/683112
Sanfo, M. B. J. (2021). Connecting family, school, gold mining community and primary school students' reading achievements in Burkina Faso-A three-level hierarchical linear model analysis. *International Journal of Educational Development, 84*, 102442. https://doi.org/10.1016/j.ijer.2020.101531
Sanfo, M. B. J. (2019). Cross-sectional analysis of the determinants of education demand in small-scale gold mining communities in Burkina Faso. *Journal of Education and Practice, 10*(12), 119-131. https://core.ac.uk/download/pdf/234642568.pdf
Sanfo, M. J. (2020a). A three-level hierarchical linear model analysis of the effect of school principals' factors on primary school students' learning achievements in Burkina Faso. *International Journal of Educational Research, 100*, 101531. https://doi.org/10.1016/j.ijer.2020.101531
Sanfo, M. J. (2020b). Leaving no place behind: Community participation and primary school students' learning achievements in Burkina Faso's small-scale gold mining communities. *International Journal of Educational Research Open, 1*, 100010. https://doi.org/10.1016/j.ijedro.2020.100010
Santos, R. J. (2018). Blessing and curse. The gold boom and local development in Colombia. *World Development, 106*, 337-355. https://doi.org/10.1016/j.worlddev.2018.02.016
Scheerens, J. (2016). School effectiveness research and the meaning of the most important effectiveness-enhancing conditions. *Educational effectiveness and ineffectiveness*. Dordrecht: Springer. https://doi.org/10.1007/978-94-017-7459-8_5
Sikora, J., Evans, M. R., & Kelley, J. (2019). Scholarly culture: How books in adolescence enhance adult literacy, numeracy and technology skills in 31 societies. *Social Science Research, 77*, 1-15. https://doi.org/10.1016/j.ssresearch.2018.10.003
Soré, Z. (2015). *Massification scolaire, rapport au savoir et qualité de l'enseignement primaire a Rambo (province du Yatenga, Burkina Faso)* [School massification, relationship to knowledge and quality of primary education in Rambo (Yatenga Province, Burkina Faso)]. Doctoral dissertation, University of Ouagadougou, Ouagadougou.
Soré, Z., & Maiga, A. (2015). Descendre dans le trou plutot que g'aller en classe: Les sites auriferes artisanaux et les deperditions scolaires au Burkina Faso [Going down the pit instead of going to to school: Artisanal gold sites and school dropouts in Burkina Faso]. *Annuals of University of Ouagadougou, 20*(A).
Strand, S. (2014). Ethnicity, gender, social class and achievement gaps at age 16: Intersectionality and 'getting it' for the white working class. *Research Papers in Education, 29*(2), 131-171. https://doi.org/10.1080/02671522.2013.767370
Sullivan, K., Perry, L. B., & McConney, A. (2013). How do school resources and academic performance differ across Australia's rural, regional and metropolitan communities? *Australian Educational Researcher, 40*, 353-372. https://doi.org/10.1007/s13384-013-0100-5
Tawil, S., Akkari, A., & Macedo, B. (2012). *Beyond the conceptual maze: the notion of quality in education*. UNESCO Education Research and Foresight Occasional Papers. https://learningportal.iiep.unesco.org/en/library/beyond-the-conceptual-maze-the-notion-of-quality-in-education
Thomson, S. (2018). Achievement at school and socioeconomic background-An educational perspective. *npj Science of Learning, 3*(5), 1-2. https://doi.org/10.1038/s41539-018-0022-0
van der Ploeg, F. (2011). Natural resources: Curse or blessing? *Journal of Economic Literature. 49*(2), 366-420. https://doi.org/10.1257/jel.49.2.366
Woessmann, L. (2003). Schooling resources, educational institutions, and student performance: The international evidence. *Oxford Bulletin of Economics and Statistics, 65*(2), 117-170. https://doi.org/10.1111/1468-0084.00045

Wokadala, J. (2016). Examining effects of school and teacher inputs on productive efficiency of Ugandan primary schools: A stochastic forntier analysis. *Advances in Social Sciences Research Journal, 3*(11), 93-105. https://doi.org/10.14738/assrj.311.2311.

World Bank. (2018). *World development report 2018: Learning to realize education's promise.* Washington, DC: World Bank. https://www.worldbank.org/en/publication/wdr2018

Zabaleta, B. M. (2011). The impact of child labor on schooling outcomes in Nicaragua. *Economics of Education Review, 30*(6), 1527-1539. https://doi.org/10.1016/j.econedurev.2011.08.008

Zabsonré, A., Agbo, M., & Some, J. (2018). Gold exploitation and socioeconomic outcomes: The case of burkina faso. *World Development, 109*, 206-221. https://doi.org/10.1016/j.worlddev.2018.04.021

APPENDIX

Table 6. Variables used to construct latent variables

Latent variables	Variables used
health	Health problems
	mosq_net = student sleeps in mosquito net
	ear_pb = student has hearing problems
	sight_pb = student has sight problems
	glas_wear = student wears glasses