The relationship between education and vulnerability to poverty in South Africa

Nicholas Ngepah[1] · Tsholofelo Makgalemele[1] · Charles Shaaba Saba[1]

Abstract

This study examines the relationship between education and poverty vulnerability in South Africa by using the disaggregated approach. The study uses the National Income Dynamics Study dataset captured from 2008 to 2017 of a balanced panel consisting of 10,800 South African households. We applied batteries of econometric techniques. The results reveal a significant negative relationship between education and vulnerability to poverty especially to those whose future consumption lies below the food poverty line. Therefore, policymakers should use education as a tool to reduce poverty vulnerability by increasing its quality so that students can compete fairly in the global labor market.

1 Introduction

South Africa (SA) has approximately 49.2% of the adult population living below the upper bound poverty line (UBPL) (StatsSA 2020). South Africa still suffers from chronic poverty because economic activity is not generating growth at the commensurate level required to address poverty (Ngepah 2017). Moreover, the

✉ Charles Shaaba Saba
  sabacharlesshaaba@yahoo.com; charless@uj.ac.za

  Nicholas Ngepah
 nngepah@uj.ac.za; nnnbal@yahoo.fr

  Tsholofelo Makgalemele
  makgalemeletsholofelo@gmail.com

[1] School of Economics and Econometrics, College of Business and Economics, University of Johannesburg, Auckland Park Kingsway Campus, Auckland Park, PO Box 524, Johannesburg, South Africa

high inequality curbs growth's ability to decrease the rate at which poverty is reduced. In effect, an exposure to a shock will result in more people falling into poverty in the future (Ngepah 2017). Thus, chronic poverty remains persistent, however, at a higher rate, making it difficult for SA to meet the Millennial Development Goal of eradicating poverty.

The present households' inability to have coping mechanisms in place such as insurance or savings in case of an adverse shock will increase vulnerability to poverty (Haq 2015; Swain and Floro 2012). However, to acquire these coping mechanisms households need to generate an income via educational investment. This necessitates the importance of using education as a catalyst for reducing vulnerability to poverty. In SA every learner is entitled to basic education as it is compulsory to attend school from grade 1 to 9. Even though this is the case, access to quality education is not available to all pupils as it is determined by affordability (Gumede and Biyase 2016). Thus, this can affect the probability of being accepted in universities and ultimately lowers the chances of being absorbed into the labor market. Statistically, South Africa has 33,171 uneducated citizens while 56,750 people have a primary education, 66,770 attained a secondary education whereas only 16,491 South Africans obtained tertiary education (National Income Dynamics Study (NIDS), 2017). Based on these statistics and the below Fig. 1, it is evident that majority of the population have a secondary education while only a hand full of people have a tertiary education.

Education is a function of economic growth through human capital and a fundamental determinant of poverty. This notion is justified by the contribution of Haughton and Khandker (2009) which highlights that the lowest level of education (primary education) slightly increases the standard of living for the impoverished population. Hence, significant policy implications have stemmed from studies which suggest that advancing human capabilities increases the rate of return
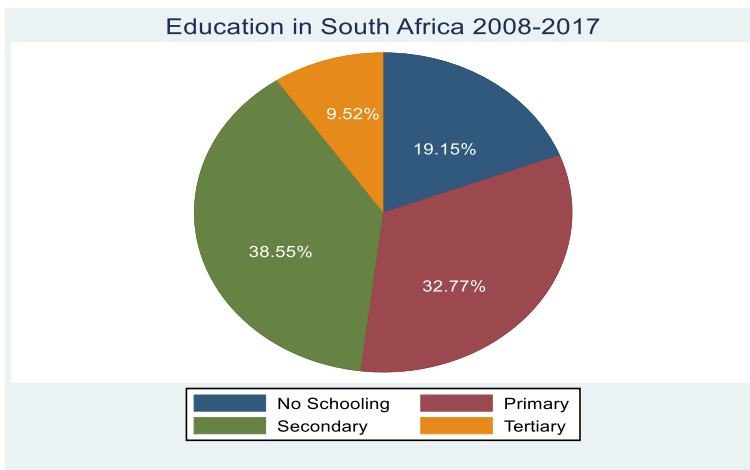
Fig. 1 Proportion of levels of education in South Africa

to educational investment via a rise in personal income (inter alia: Gumede and Biyase 2016; Haughton and Khandker 2009).

There is an extensive body of literature that study education and poverty alleviation but only a handful of studies focus on vulnerability to poverty (inter alia: Brown and James 2020; Dzidza et al. 2018; Kawulur et al. 2019). In developing countries, studies show that more than 50% of the group which is vulnerable to poverty is currently non-poor (Haughton and Khandker 2009). Hence, the population that is currently poor is not an accurate indicator of those who will be poor in the next period (Haq 2015). Thus, this marks the importance of studying poverty vulnerability.

In the light of recent events, the COVID-19 pandemic has left many South Africans who were previously non-poor vulnerable to transitioning into poverty. This highlights the importance of evaluating the effect that education can have on reducing vulnerability to poverty. The most efficient way to do that is by using the disaggregated approach as it captures complicated complex patterns and disparities in the education–vulnerability to poverty nexus that transpire within South Africa. Over and above that, this approach exposes the differing contexts in which this nexus occurs (Ajuruchukwu and Sanelise 2016). Moreover, this approach provides enough information required to formulate initiatives that target the most vulnerable (Hong and Sullivan 2013).

In most cases, studies follow the aggregation approach in analyzing the relationship between the two variables to deduce an overview about the nexus and thus disregarding the details (Christiaensen and Boisvert 2000; Pritchett et al. 2000; Yang and Guo 2020). However, in a country that suffers from chronic poverty, it is advisable to dismantle the issue of vulnerability to poverty in order to establish the root cause of the problem. This can be done effectively by using the disaggregated approach, whereby vulnerability to poverty is analyzed across all poverty lines thus allowing the assessment of poverty variability (Ajuruchukwu and Sanelise 2016).

Furthermore, this approach helps determine how much of the population falls under the most vulnerable group so that resources can be allocated effectively and efficiently. In addition to the above-mentioned benefits of disaggregating both variables, researchers can establish the impact and magnitude each level of education has on the group that is at most risk (Pritchett et al. 2000). Hence, this study is integral because it intends to fill this gap found in the literature.

This study differs from the one conducted by Yang and Guo (2020) as it simultaneously disaggregates education and vulnerability to poverty. As per policy implications, the objective of government strategies should be such that they are prospective for the advancement of economic development. However, most poverty-reduction policies emanate from research that has focused on poverty alleviation which is ex-post in its nature (Yang and Guo 2020). This in essence does not account for the population that might transit into poverty or will remain stuck in the poverty trap in the future. Hence, this necessitates studying vulnerability to poverty as it is ex-ante in its approach and allows for the formation of preventative policies.

This study is structured to review the existing literature in Sect. 2 on the subject matter. Subsequently, Sect. 3 discusses the methodology whereas Sect. 4 provides a discussion of the empirical results. Lastly, the fifth chapter will conclude the topic at hand with relevant policies.

2 Literature review

The method used to estimate vulnerability is determined by how vulnerability to poverty is defined. Poverty vulnerability is measured using the following approaches: vulnerability as the expected poverty (VEP), vulnerability as uninsured exposure to risk (VER), and vulnerability as the low expected utility (VEU) (Hoddinott and Quisumbing 2010). Notwithstanding the importance of conceptualizing vulnerability to poverty, it is equally imperative to understand education's role in eradicating poverty vulnerability. This section of the study intends to review the existing empirical literature in a quest to identify and fill in the gap by making a significant contribution to the body of literature.

VER technique is defined as a measure of household's ability to be resilient against an adverse shock, instead of being an aggregate indicator of poverty vulnerability. Celidoni and Procidano (2015) contend that this theory is backward-looking and fails to consider the probability of negative covariate and idiosyncratic shocks and the household's risk appetite. Therefore, this approach contributed more to the consumption smoothing literature than vulnerability to poverty theory. According to the VEU method poverty vulnerability is when households are considered vulnerable if they have increased consumption/income variability in the past (Ligon and Schechter 2002). Contrary to the VER approach, this measure accounts for risk appetites; therefore, it is best suited to measure welfare and is disintegrated into covariate-risk, idiosyncratic risk, and the poverty component (Klasen and Waibel 2013). Although the VEU method is empirically backward-looking it can detect the probability and intensity of an adverse shock. This approach has its merits; however, it is not fixated on poverty vulnerability and therefore adds very minimal value to poverty literature (Calvo and Dercon 2013).

Unlike the approaches discussed previously, the VEP technique uses the poverty lines to measure vulnerability to poverty. Ravallion (1988) set precedence for the conceptualization of this approach. However, Holzmann and Jorgensen (1999) modified it by defining vulnerability to poverty as the probability of falling below the poverty line or staying trapped in poverty. Although this procedure is forward-looking it does not incorporate risk sensitivity and severity. Fortunately, this can be resolved by using the Foster–Greer–Thorbecke (FGT) poverty measures (Hoddinott and Quisumbing 2010). But the consequence thereof is the implication of increased risk aversion of households which is empirically inaccurate.

There is still no consensus in measuring poverty vulnerability. Nevertheless, vulnerability to poverty as a theory is a function of various determinants. Our determinant of interest is education and its substantial impact on vulnerability to poverty. Human capital theory links the nexus in question (Becker 1962). The premise of this neoclassical theory states that the marginal productivity of labor and thus increased earnings is determined by education (Marginson 2019). The screening theory also uses education as an indicator of a person's capabilities. These theories are similar in that they perceive education as a value adding medium to organizations, such that the current cost is superseded by future income and higher production.

The empirical literature on poverty analysis has shown that determinants of vulnerability to poverty are country specific (Haughton and Khandker 2009). Hence, this necessitates an analysis of the education and vulnerability to poverty nexus in South Africa (Omotoso et al. 2020). However, regardless of the type of approaches used the nexus remains negative.

The majority of the studies on poverty vulnerability only focus on the different estimation techniques (inter alia: Pham et al. 2021; Aikaeli et al. 2021). However, the studies that do control for education as a determinant of vulnerability to poverty usually take an aggregate approach (Pham et al. 2021; Guo et al. 2019). Cruces et al. (2011) evaluates Chile, Nicaragua, and Peru's vulnerability to poverty by using synthetic panels to construct vulnerability measures. Not only does the study use synthetic panels to estimate the future household income but it also examines the intergenerational poverty mobility via growth incidence curves (GIC) and transition matrices. The authors find that more years spent on attaining an education by the household head tend to reduce vulnerability to poverty by 12.4% in Chile, 9.3% in Nicaragua, and 8.8% in Peru.

Another study that examines the education–vulnerability to poverty nexus is one compiled by Haq (2015) and it used cross-sectional data of 2010. Moreover, it investigates the impact of exposure to shocks on households and the capability of alleviating such impacts in Pakistan. The study uses the multinomial logit model to find that an educated household head has a 2.8% probability of experiencing an income shock and the fixed effects model (FEM) yielded a 1.1% reduction in poverty vulnerability.

Yang and Guo (2020) did a study that specifically analyzes the link between universal basic education and vulnerability to poverty in rural China from the capability perspective. They applied the VEP approach on the China Family Panel Studies (CFPS) dataset of 2010. The authors find that the popularization of basic education significantly reduces vulnerability to structural poverty. Furthermore, the results show that compulsory basic education improves the capabilities to enhancing permanent income. However, this study fails to analyze all levels of education and their effect on the vulnerability to poverty across all poverty lines.

This study aims to fill this gap by using the VEP approach to construct the vulnerability measure in making use of the NIDS panel data of South Africa for the period 2008–2017. To calculate the expected consumption and variance the FGLS technique will be implemented. These estimates will then be used to compute the vulnerability measure which is equal to the difference between the poverty line and the expected consumption divided by the standard deviation. This study seeks to use the same approach found in Yang and Guo's (2020) study to construct the vulnerability measure. Thereafter, the pooled OLS model exploited in Christiaensen and Boisvert's (2000) study, the random effects model (REM) employed in Donou-adonsou et al.'s (2021) study and the Hausman Taylor (HT) model will be used to regress vulnerability to poverty on its determinants. This study seeks to contribute to the literature by simultaneously disintegrating education and vulnerability to poverty since to the best of my knowledge no study has investigated this important aspect.

This study will assist the government in targeting and creating policy cohesion across all levels of education ensuring that the educational gap is reduced, and households do not lack in the quality and accessibility of education.

3 Research methodology

3.1 Theoretical framework of measuring vulnerability to poverty

The determinants of poverty are also classified under the following dimensions: regional, community, household, and individual traits (Haughton and Khandker 2009). Therefore, poverty can be expressed as a function of demographic, economic and social characteristics:

$$\text{poverty} = f\left(\text{education}, \text{employmentstatus}, \text{geography}, \text{age}, \text{age}^2, \text{race}, \text{gender}, .., n\right)$$

The assessment of vulnerability to poverty necessitates the use of a proxy for welfare, which is typically represented by income/consumption. Measuring income per capita in developing countries is difficult because the majority of it is generated by self-employment in the informal sector. However, consumption expenditure per capita is easy to report and quantify because it is a direct metric (Haughton and Khandker 2009). Despite some drawbacks, many academics support the use of consumption expenditure per capita in developing nations. This is because it depicts the ongoing standard of living, adjusts inconsistencies to current well-being through time, and is considered more accurate because most individuals recall their spending patterns (Haughton and Khandker 2009).

The VEP technique is an appropriate technique for examining developing countries because it not only accommodates panel data but also cross-sectional datasets. Even so, this approach on the one hand poses a measurement error issue which can be solved by making multiplicative adjustments to the predicted variance (Chaudhuri 2003). On the other hand, the VEP method suffers from heterogeneity that is controlled for through the Monte Carlo bootstrapping technique (Kamanou and Morduch 2002). Chaudhuri (2003) argues that although unobserved heterogeneity has been accounted for heteroscedasticity remains problematic. However, by conducting both the nonparametric and parametric bootstrapping procedure will significantly reduce this econometric issue.

Imai et al. (2011) as well as Christiaensen and Subbarao (2005) conceptually augmented the VEP approach by including the household's risk mitigating capacity. Other scholars agree with this notion, however, advocate for the importance of taking risk appetite into consideration (See, for example, Chaudhuri 2003; Calvo and Dercon 2013 and Ligon and Schechter 2002).

Ligon and Schechter (2002) introduced the VEU method to incorporate risk aversion and disintegrate vulnerability to poverty into uninsured risk and poverty. Similar to these researchers, Calvo and Dercon (2013) contributed to this concept yet proposed different types of utility functions. Furthermore, the VEU and VEP approach are both suitable for examining household vulnerability to poverty at an aggregate level as opposed to measuring individual vulnerability (Mahanta and Das 2015). As with other techniques this approach falls short in a few ways. First, when assessing developing countries, it becomes a monumental task to estimate risk preferences at an individual level. Second, in the absence of repeated cross-sectional or panel datasets calculating vulnerability becomes a challenging task. Finally, presuming that

the form of the utility function is determined by the observed risk response then risk aversion will be overestimated (Christiaensen and Subbarao 2005).

Similar to the VEU methodology, the VER technique can predict both the covariate and idiosyncratic shocks that increase welfare loss. However, it cannot generate a headline vulnerability to poverty estimate (Hoddinott and Quisumbing 2010). Moreover, Calvo and Dercon (2013) critically analyzed the VER's inability to provide the likelihood of adverse shocks occurring but rather concentrates on assessing the response to negative shocks. In addition to this critique Christiaensen and Subbarao (2005) argue that the VER approach does not deem the poor as vulnerable even though a minor adverse shock can substantially reduce consumption.

3.2 Computing vulnerability using the VEP approach

This study will adopt the VEP approach which is a method that is commonly used in measuring vulnerability to poverty (Celidoni 2015; Chaudhuri 2003; Christiaensen and Boisvert 2000; Deressa 2013). The model is specified as follows:

$$\text{vul}_{h,t} = \text{Pr}\left(y_{h,t+1} \leq z\right) \tag{1}$$

whereby $\text{vul}_{h,t}$ is the vulnerability to poverty of household h at time t, while $y_{h,t+1}$ is the consumption per capita of household h at time $(t+1)$, and z is the poverty line. Therefore, the model represents the likelihood that the household consumption of h is below the poverty line at time $(t+1)$. However, to quantify this, we first need to estimate the future consumption which is represented by the below stochastic process:

$$\ln y_{h,t} = \propto Z_{h,t} + \varepsilon_{h,t} \tag{2}$$

whereby $Z_{h,t}$ is the vector of determinants that explain the variations in the log of consumption per capita represented by $lny_{h,t}$. This vector of determinants consists of the head of household's education level, gender, geography, employment status, household size, no landownership, age, and race. There is empirical evidence in the literature that supports the economic theory which stipulates that vulnerability to poverty has an inverse relationship with education, race, geography, no landownership, and employment status. However, posits that it positively relates to household size. With regard to gender, females are more likely to be vulnerable to poverty than their male counterpart (See, for example, Christiaensen and Subbarao 2005; Christiaensen and Boisvert 2000; Cruces et al. 2011; Haq 2015; Pritchett et al. 2000). These variables have been chosen because existing empirical evidence emphasizes the significant role they play in explaining variations in consumption per capita (See, for example, Celidoni 2015; Chaudhuri 2003; Deressa 2013; Günther and Harttgen 2009; Yang and Guo 2020). The $\varepsilon_{h,t}$ is the error term that represents idiosyncratic shocks with a mean of zero. If the covariance of the error terms is unknown and there is some heteroscedasticity/autocorrelation, FGLS is an appropriate estimation technique for large samples in comparison with the OLS (Missiakoulis 2020). However, the opposite is true for a small sample size as the FGLS yields inconsistent

estimates in the presence of heterogeneity (Gujarati and Porter 2009). The FGLS follows a two-stage estimation procedure that first uses the OLS to estimate residuals that form the error covariance matrix estimator. Thereafter, this consistent estimator can be used to implement generalized least squares (GLS). This study follows the FGLS suggested by Chaudhuri (2003), to yield the estimate of the expected log consumption and its variance, $\hat{\alpha}$ and $\hat{\beta}$, respectively. This can be expressed as:

$$\hat{E} = E(\ln y_h | Z_h) = Z_h \hat{\alpha} \tag{3}$$

$$\hat{V} = \text{var}(\ln y_h | Z_h) = \sigma_{e,h}^2 = Z_h \hat{\beta} \tag{4}$$

Therefore, the estimated vulnerability to poverty is specified as:

$$\widehat{\text{vul}}_h = \widehat{\text{Pr}}(\ln y_h \le \ln z) = \Phi\left(\frac{\ln z - Z_h \hat{\alpha}}{\sqrt{Z_h \hat{\beta}}}\right) \tag{5}$$

The estimated vulnerability to poverty $(\widehat{\text{vul}}_h)$ follows a cumulative density of standard normal distribution denoted by $\Phi(\cdot)$. Moreover, to compute this probability we use the national food, lower bound, and upper bound poverty line set at R585, R840, and R1268 per month, respectively. These poverty lines were calculated in April 2020 by Statistics South Africa and have been adjusted for inflation (StatsSA 2020).

3.3 Methods used to estimate the determinants

Considering that the calculation of vulnerability to poverty is now known it is possible to regress it on its determinants by using the pooled OLS model, random effects model (REM), and the Hausman Taylor (HT) model to analyze the relationship in question. This paper uses these models because they are best suited for the NIDS dataset. Furthermore, when dealing with a panel dataset these models serve as estimation tools that control for unobserved time-invariant attributes which could be correlated with regressors (Gujarati and Porter 2009). We rely more on the HT model because its merits while other models serve as baseline models.

The pooled OLS model expresses constant coefficients by pooling data and using the OLS estimation technique which minimizes error terms. According to Wooldridge (2018), this technique is implemented when using a varied sample for each period of the panel data and is specified as follows:

$$\widehat{\text{vul}}_{h,t} = \beta + \gamma X_{h,t} + a_h + \mu_{h,t} \tag{6}$$

whereby vulnerability to poverty $\widehat{vul}_{h,t}$ is a function of the intercept β, vector of regressors $X_{h,t}$, unobserved heterogeneity a_h, and the error term $\mu_{h,t}$. The pooled OLS method assumes that this error term is identically and independently distributed with a zero mean and a homoscedastic variance, $\mu_{h,t} \sim iid\left(0, \sigma_\mu^2\right)$. The sec-

ond assumption postulates that the regression coefficients are the same across all households. The last key assumption stipulates that time-invariant attributes absorbed in the error term should not be correlated to explanatory variables, $\text{cov}(\mu_{h,t}, X_{h,t}) = 0$ (Gujarati and Porter 2009).

The REM shares the same assumptions with the FEM of homoscedasticity $[\text{var}(v_h) = E(a_h + \mu_h^2 | X_h) = \sigma_v^2]$ and strict exogeneity $[E(\mu_{h,t} | X_{h,t}, a_h) = 0]$ (Wooldridge, 2018). The rationale of the REM is that the differences across households are random as opposed to fixed. Moreover, it employs the quasi-demeaning transformation which can be specified as:

$$\text{vul}_{h,t} - \overline{\theta \text{vul}}_h = \gamma(X_{h,t} - \overline{\theta X}_h) + \mu_{h,t} - \theta\overline{\mu}_h \tag{7}$$

and can also be written as:

$$vul_{h,t} = \delta X_{h,t} + v_{h,t} \tag{8}$$

$$v_{h,t} = a_h + \mu_{h,t} \tag{9}$$

Equation (10) illustrates that the vulnerability to poverty ($vul_{h,t}$) depends on a vector of determinants ($X_{h,t}$) and the composite error term ($v_{h,t}$). This composite error term is a function of the unobserved heterogeneity (a_h) and the idiosyncratic error term ($\mu_{h,t}$).

Since the FEM omits time-invariant variables thus impeding the analysis of the nexus in question this study will not make use of it. Instead, the HT model will be conducted not only to curb this issue but also model unobserved heterogeneity. Therefore, the HT model is structured as a mixture of the FEM and REM. Given that this study uses panel data implies that it is ideal for dealing with endogeneity problems. However, those who are vulnerable to poverty might allow their vulnerability to influence their educational and labor-market decisions. Furthermore, the same underlying determinants of educational investment and labor market can impact vulnerability. This would suggest that there is significant endogeneity bias which the HT model accounts for. This model uses the IV regression technique to solve endogeneity bias whereby the exogenous regressors serve as instruments. The HT model is specified as the following:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \delta_1 Z_{1it} + \delta_2 Z_{2it} + \mu_i + v_{it} \tag{10}$$

whereby X_{1it} is a $1xk_1$ vector of exogenous, time-varying variables assumed to be uncorrelated with $\mu_i + v_{it}$. While X_{2it} is a $1xk_1$ vector of endogenous, time-varying variables are assumed to be correlated with μ_i but not with v_{it}. Z_{1i} is a $1xg_1$ vector of exogenous, time-invariant variables assumed to be uncorrelated with $\mu_i + v_{it}$. And Z_{2i} is a $1xg_2$ vector of endogenous, time-invariant variables assumed to be correlated with μ_i but not with v_{it}.

In reference to Table 1 that describes the dependent variables as well as regressors our model can be specified as follows:

Pooled OLS:

Table 1 Description of the dependent and independent variables.

Variables	Type data source	Theoretical expectation	Description
Dependent variable			
Vulnerability to poverty (vul)	Normalized NIDS,2008–2017		It ranges from 0 to 1
Explanatory variables			
Household size (hhsize)	Continuous NIDS,2008–2017	Positive	Number of members in the household
Education (edu)	Dummy Authors	Negative	1 = Primary, 0 = Otherwise
Primary education	Dummy Authors	Negative	2 = Secondary, 0 = Otherwise
Secondary education	Dummy Authors	Negative	3 = Tertiary, 0 = Otherwise
Tertiary education			
Race			
Africans	Dummy Authors	Negative	1 = Africans, 0 = Otherwise
Coloured	Dummy Authors	Negative	1 = Coloured, 0 = Otherwise
Asian/Indian	Dummy Authors	Negative	1 = Asian/Indian, 0 = Otherwise
White	Dummy Authors	Negative	1 = White, 0 = Otherwise
Gender (gen)	Dummy Authors	Negative	1 = Male, 0 = Otherwise
Male	Dummy Authors	Positive	2 = Female, 0 = Otherwise
Female			
Employment status (employstat)	Dummy Authors	Positive	1 = Unemployed, 0 = Otherwise
Unemployed	Dummy Authors	Negative	2 = Employed, 0 = Otherwise
Employed	Dummy Authors	Positive	1 = Youth, 0 = Otherwise
Age	Dummy Authors	Negative	2 = Middle age, 0 = Otherwise
Youth	Dummy Authors	Positive	3 = Pensioner, 0 = Otherwise
Middle age	Dummy Authors	Positive	1 = Youth Squared, 0 = Otherwise
Pensioner	Dummy Authors	Negative	2 = Middle age squared, 0 = Otherwise
Age squared	Dummy Authors	Positive	3 = Pensioner squared, 0 = Otherwise
Youth squared	Dummy Authors	Negative	1 = if you own land, 2 = No landownership
Middle age squared			
Pensioner squared			
Landownership			
Geography (geo)			

Table 1 (continued)

Variables	Type data source	Theoretical expectation	Description
Traditional areas	Dummy Authors	Positive	1 = Household in traditional areas, 0 = Otherwise
Urban	Dummy Authors	Negative	2 = Household in urban areas, 0 = Otherwise
Education_Race (edu_race)	Dummy Authors	Positive	Interaction term
Primary education_African	Dummy Authors	Positive	1 = Primary_African, 0 = Otherwise
Primary education_Coloured	Dummy Authors	Positive	2 = Primary_Coloured, 0 = Otherwise
Primary education_Asian/Indian	Dummy Authors	Positive	3 = Primary_Asian/Indian, 0 = Otherwise
Primary education_White	Dummy Authors	Positive	4 = Primary_White, 0 = Otherwise
Secondary education_African	Dummy Authors	Positive	1 = Secondary_African, 0 = Otherwise
Secondary education_Coloured	Dummy Authors	Positive	2 = Secondary _Coloured, 0 = Otherwise
Secondary education_Asian/Indian	Dummy Authors	Positive	3 = Secondary _Asian/Indian, 0 = Otherwise
Secondary education_White	Dummy Authors	Positive	4 = Secondary _White, 0 = Otherwise
Tertiary education_African	Dummy Authors	Positive	1 = Tertiary_African, 0 = Otherwise
Tertiary education_Coloured	Dummy Authors	Positive	2 = Tertiary _Coloured, 0 = Otherwise
Tertiary education_Asian/Indian	Dummy Authors	Positive	3 = Tertiary _Asian/Indian, 0 = Otherwise
Tertiary education_White	Dummy Authors	Positive	4 = Tertiary _White, 0 = Otherwise

Source: NIDS dataset

$$\widehat{\text{vul}_{h,t}} = \alpha + \beta_1\text{edu}_{h,t} + \beta_2\text{gen}_{h,t} + \beta_3\text{race}_{h,t} + \beta_4\text{geo}_{h,t} + \beta_5\text{employstat}_{h,t} + \beta_6\text{hhsize}_{h,t}$$
$$+ \beta_7\text{age}_{h,t} + \beta_8\text{age2}_{h,t} + \beta_9\text{no\_landownership}_{h,t} + \beta_{10}\text{edu\_race}_{h,t} + \mu_{h,t} \tag{11}$$

3.4 REM

$$\widehat{\text{vul}_{h,t}} = \alpha + \beta_1\text{edu}_{h,t} + \beta_2\text{gen}_{h,t} + \beta_3\text{race}_{h,t} + \beta_4\text{geo}_{h,t} + \beta_5\text{employstat}_{h,t} + \beta_6\text{hhsize}_{h,t}$$
$$+ \beta_7\text{age}_{h,t} + \beta_8\text{age2}_{h,t} + \beta_9\text{no\_landownership}_{h,t} + +\beta_{10}\text{edu\_race}_{h,t} + \mu_{h,t} \tag{12}$$

3.5 HT model

$$\widehat{\text{vul}_{h,t}} = \alpha + \beta_1\text{edu}_{h,t} + \beta_2\text{gen}_{h,t} + \beta_3\text{race}_{h,t} + \beta_4\text{geo}_{h,t} + \beta_5\text{employstat}_{h,t}$$
$$+ \beta_6\text{hhsize}_{h,t} + \beta_7\text{age}_{h,t} + \beta_8\text{age2}_{h,t} + \beta_9\text{no\_landownership}_{h,t} \tag{13}$$
$$+ +\beta_{10}\text{edu\_race}_{h,t} + \mu_{h,t}$$

The exogenous variables represented in the above equation are gender, race, geography, employment status, household size, age, and no landownership, except education as it is an endogenous variable.

3.6 Econometric issues of estimation methods

Even though the pooled OLS is not a popular choice in the literature due to its susceptibility to violating the classical linear regression model assumptions, it remains a good starting point for panel data analysis (See, for example, Bergh and Nilsson 2014; Donou-adonsou et al. 2021; Vuko and Cular 2014). As it can still yield consistent and unbiased estimates even in the presence of time-invariant attributes, however, the REM estimates are more efficient (Vuko and Cular 2014). Moreover, in the absence of heteroscedasticity and autocorrelation, the pooled OLS estimates are efficient and have consistent standard errors thus allowing researchers to draw inferences. Yet, in an event that heteroscedasticity and autocorrelation are not controlled for then, the unobserved heterogeneity will generate biased and inconsistent estimates when using the pooled OLS model (Wooldridge 2018). This occurs by grouping different households with their respective unobserved household-specific characteristics in a pooled OLS estimation technique which conceals these unobserved time-invariant attributes in the error term that could be correlated to the regressor. Essentially it is the econometric issue of endogeneity that produces biased and inconsistent estimates. Fortunately, as a remedy, the REM will be used to control for unobserved heterogeneity.

Unfortunately, the FEM eliminates time-invariant variables thus in such a scenario it is best to use the REM (Gujarati and Porter 2009). Unlike the FEM, the REM automatically corrects for heteroscedasticity.

Even though the REM generates better estimates, they still suffer from endogeneity and hence the HT model proves to be the best choice. The HT model has combined features of the REM and FEM as well as controls for endogeneity bias. Moreover, exogenous included variables and their means across time are employed as efficient instruments, which eliminates ambiguity in terms of instrument selection.

3.7 Data

This study applies the disaggregated approach of examining the education–vulnerability to poverty nexus by using the National Income Dynamics Study (NIDS) panel data. Based on individual and household data availability we conducted a study that focuses on a period between 2008 and 2017, whereby the same 39,400 individuals and 10,800 households were interviewed since its inception 12 years ago. Moreover, NIDS is the fundamental national household panel study conducted in South Africa that collects data every two years.

This study makes use of secondary data that is structured as longitudinal panel data which controls for omitted variable bias even though it poses the risk of suffering from heteroscedasticity and autocorrelation (Wooldridge 2018). Moreover, to ensure that the data used will produce the accurate results we cleaned it by removing missing observations, those who did not know the answer provided in the questionnaire, and those who refused to answer.

4 Empirical results

Table 2 provides the descriptive statistics of profiling vulnerability to poverty according to education. In South Africa, a person with a primary education on average has an 82% probability of becoming poor in the future based on the upper bound poverty line, a 62.1% probability according to the lower bound poverty line, and a 40.6% chance based on the food poverty line. These probabilities in comparison with those who have a secondary and tertiary education are higher across all poverty lines. Thus, the general trend suggests that the more education an individual obtains the lower the probability of falling into poverty in the next period regardless of which poverty line a person falls under. Furthermore, these results show that vulnerability to poverty is an issue in SA as the majority of the population will experience poverty in the future; however, this kind of poverty will not be severe.

Table 2 Descriptive statistics of profiling vulnerability to poverty according to education.

Education Level	Observations	Mean (vul_upl)	Mean (vul_lpl)	Mean (vul_fpl)
No schooling	97,693	0.945	0.838	0.662
Primary	97,693	0.820	0.621	0.406
Secondary	97,693	0.662	0.441	0.251
Tertiary	97,693	0.316	0.152	0.066

Source: NIDS data (Wave 1—5) estimation by the author

The below table shows the estimation results using the pooled OLS model (which minimizes the sum of squared error terms) based on three poverty lines. Given that these models have been corrected for heteroscedasticity and autocorrelation it allows us to interpret the results and deduce inferences about the sample and thus the population at large. The second column illustrates the pooled OLS model used to regress vulnerability to poverty of the poor whose future consumption lies below R1268 (upper bound poverty line) on its determinants. The nexus between education and vulnerability to poverty is negative. For example, a person whose consumption is below the upper bound poverty line who has a primary education relative to one with no education will reduce vulnerability to poverty by 1.2%. However, poorer people with the same level of education will reduce poverty vulnerability even further by 5.6% and 15.2%, respectively. This is similar to the Omotoso (2019) study.

Race indicates that the racial group that was less previously disadvantaged has a much higher probability of not falling into poverty in the future relative to Africans. Coloured people can reduce vulnerability to poverty by 2.4%, whereas white people will reduce it by a substantial 91.7%. However, each race reduces poverty vulnerability further the more their consumption level is below a more stringent poverty line. In SA being a female increases the probability of falling into poverty in the future by 3.7% compared to males for women whose consumption falls under the food poverty line. These results are similar to the Nortjie (2017) study. The same positive relationship exists for those who do not own land, have large household sizes and are young. The middle aged and pensioners have a negative relationship with vulnerability to poverty, but the squared versions do not exhibit any results at all. There is a negative nexus between the education–race interaction term and poverty vulnerability. For example, an Asian/Indian person with a secondary education will reduce poverty vulnerability by 37.8% further than an African person with no schooling (based on Poor ≤ R1268 poverty line). Geography has a negative relationship with poverty vulnerability. For instance, based on the food poverty line those who live in the urban area will reduce poverty vulnerability by 8.9% and farm dwellers will reduce it by 9.8%. This is contrary to the theoretical expectation which stipulates that those who live in urban areas should reduce the probability of falling into poverty in the future because urban areas are more developed and thus have more opportunities of employment and starting a business (Yang and Guo 2020).

Furthermore, most of the regressors in Table 3 are statistically significant at a 1% significance level and all three models have high R^2 implying that 87.5%, 82.7%, and 85.4%, respectively, of variations in the vulnerability to poverty are explained by these regressors. Moreover, the RMSE is low (0.058, 0.089, and 0.111) but cannot be analyzed in isolation but rather it needs to be compared to those of the REM. The pooled OLS models make for a good starting point; however, the REM (see Table 4) are more appropriate for controlling the unobserved heterogeneity.

In Table 4 we can see that all regressors are statistically significant and the coefficients are larger than those reported in the pooled OLS. However, the standard errors are slightly lower. What is different between the two tables is that the no landownership in Table 4 is statistically significant under the upper and lower bound poverty line only whereas in Table 3 it is significant under the upper and food poverty line. By caution, we conducted the LM test to verify that the REM

Table 3 Estimations using the Pooled OLS model based on three poverty lines

Variables	Poor ≤ R1268	Poor ≤ R840	Poor ≤ R585
Primary	− 0.012***	− 0.056***	− 0.152***
	(0.001)	(0.003)	(0.004)
Secondary	− 0.043***	− 0.147***	− 0.317***
	(0.002)	(0.004)	(0.005)
Tertiary	− 0.125***	− 0.327***	− 0.537***
	(0.005)	(0.007)	(0.006)
Coloured	− 0.024**	− 0.155***	− 0.351***
	(0.011)	(0.028)	(0.027)
Asian/Indian	− 0.122**	− 0.431***	− 0.703***
	(0.049)	(0.111)	(0.059)
White	− 0.917***	− 0.772***	− 0.546***
	(0.003)	(0.005)	(0.006)
Primary_Coloured	− 0.077***	− 0.136***	− 0.056*
	(0.016)	(0.033)	(0.029)
Primary_Asian/Indian	− 0.378***	− 0.297**	0.030
	(0.074)	(0.116)	(0.060)
Primary_White	− 0.034***	− 0.079***	− 0.071***
	(0.002)	(0.004)	(0.006)
Secondary_Coloured	− 0.129***	− 0.171***	− 0.035
	(0.015)	(0.031)	(0.028)
Secondary_Asian/Indian	− 0.464***	− 0.284**	0.111*
	(0.057)	(0.112)	(0.061)
Secondary_White	0.005	0.024*	0.096***
	(0.005)	(0.014)	(0.025)
Tertiary_Coloured	− 0.435***	− 0.279***	0.134***
	(0.033)	(0.032)	(0.033)
Tertiary_Asian/Indian	− 0.618***	− 0.114	0.460***
	(0.057)	(0.112)	(0.067)
Tertiary_White	0.096***	0.254***	0.402***
	(0.006)	(0.011)	(0.015)
Female	0.012***	0.027***	0.037***
	(0.001)	(0.002)	(0.002)
Employed	− 0.033***	− 0.094***	− 0.154***
	(0.001)	(0.002)	(0.003)
Urban	− 0.018***	− 0.054***	− 0.089***
	(0.002)	(0.003)	(0.003)
Farm	− 0.023***	− 0.059***	− 0.098***
	(0.003)	(0.004)	(0.005)
Household size	0.004***	0.014***	0.032***
	(0.000)	(0.000)	(0.001)
No landownership	0.002*	0.000	− 0.004*
	(0.001)	(0.002)	(0.002)

Table 3 (continued)

Variables	Poor ≤ R1268	Poor ≤ R840	Poor ≤ R585
Youth	0.024***	0.039**	0.007
	(0.005)	(0.016)	(0.023)
Middle age	−0.014**	−0.060***	−0.157***
	(0.005)	(0.016)	(0.023)
Pensioner	−0.021***	−0.084***	−0.215***
	(0.006)	(0.016)	(0.023)
Youth squared	–	–	–
Middle age squared	–	–	–
Pensioner squared	–	–	–
Constant	0.988***	0.978***	0.942***
	(0.006)	(0.016)	(0.024)
RMSE	0.058	0.089	0.111
Observations	12,420	12,420	12,420
R-squared	0.875	0.827	0.854

is indeed the most suitable model. Thereafter, we conducted the heteroscedasticity test and subsequently controlled for it. Although we conducted the LM test to choose the most appropriate model the root mean squares error (RMSE) is a better measure. It is evident that based on this measure the REM is definitely the most appropriate model because it has the lowest RMSE results (for example, 0.029, 0.052, and 0.037) across all poverty lines which implies a better goodness of fit. However, because this model still suffers from endogeneity, we make use of the HT model to correct for such an econometric issue. The HT results can be found in Table 5.

This study relies heavily on the HT model because of its merits, whereby other previous models serve as baseline models. The relationship of interest is one between vulnerability to poverty and education. An individual with a primary school education relative to one with no education will reduce their vulnerability to poverty by 1%, ceteris paribus, based on Poor ≤ R1268. However, in comparison with secondary and tertiary education the individual is expected to reduce their vulnerability to poverty much further. Thus, this highlights a particular pattern that the more education a person has leads to a higher reduction of their vulnerability to poverty.

A second trend that is evident is that as the individual's future consumption falls below a lower poverty line, any level of education will have a bigger impact in reducing their vulnerability to poverty compared to the less restrictive poverty line. For example, those who are vulnerable to poverty in the next period based on the food poverty line (Poor ≤ R585) will be able to reduce that probability by 25% with a secondary education compared to a 2.8% reduction in poverty vulnerability based on the upper bound poverty line (Poor ≤ R1268). According to economic theory, the reason why education can reduce vulnerability to poverty is because the development of cognitive capabilities allows individuals to be creative (Yang and Guo

Table 4 Estimations using the random effects model based on three poverty lines. *Source*: NIDS data (Wave 1—5) estimation by the author, robust standard errors in parentheses

Variables	Poor ≤ R1268	Poor ≤ R840	Poor ≤ R585
Primary	− 0.013***	− 0.060***	− 0.155***
	(0.001)	(0.003)	(0.004)
Secondary	− 0.042***	− 0.145***	− 0.313***
	(0.002)	(0.004)	(0.005)
Tertiary	− 0.118***	− 0.318***	− 0.531***
	(0.005)	(0.007)	(0.006)
Coloured	− 0.020*	− 0.145***	− 0.340***
	(0.011)	(0.028)	(0.027)
Asian/Indian	− 0.115**	− 0.399***	− 0.680***
	(0.058)	(0.122)	(0.060)
White	− 0.936***	− 0.847***	− 0.674***
	(0.003)	(0.009)	(0.014)
Primary_Coloured	− 0.072***	− 0.130***	− 0.053*
	(0.016)	(0.032)	(0.029)
Primary_Asian/Indian	− 0.373***	− 0.306**	0.021
	(0.085)	(0.128)	(0.061)
Primary_White	− 0.009***	0.008	0.066***
	(0.003)	(0.009)	(0.015)
Secondary_Coloured	− 0.133***	− 0.178***	− 0.043
	(0.015)	(0.030)	(0.028)
Secondary_Asian/Indian	− 0.480***	− 0.312**	0.099
	(0.064)	(0.123)	(0.062)
Secondary_White	0.029***	0.108***	0.230***
	(0.002)	(0.005)	(0.010)
Tertiary_Coloured	− 0.447***	− 0.287***	0.128***
	(0.033)	(0.032)	(0.032)
Tertiary_Asian/Indian	− 0.604***	− 0.144	0.429***
	(0.066)	(0.123)	(0.067)
Tertiary_White	0.109***	0.320***	0.520***
	(0.006)	(0.012)	(0.018)
Female	0.014***	0.029***	0.039***
	(0.001)	(0.002)	(0.002)
Employed	− 0.029***	− 0.086***	− 0.147***
	(0.001)	(0.002)	(0.003)
Urban	− 0.020***	− 0.060***	− 0.093***
	(0.002)	(0.003)	(0.003)
Farm	− 0.028***	− 0.071***	− 0.110***
	(0.004)	(0.005)	(0.005)
Household size	0.005***	0.015***	0.033***
	(0.000)	(0.000)	(0.001)

Table 4 (continued)

Variables	Poor ≤ R1268	Poor ≤ R840	Poor ≤ R585
No landownership	0.004***	0.004**	−0.002
	(0.001)	(0.002)	(0.002)
Youth	0.020***	0.036***	0.008
	(0.004)	(0.011)	(0.020)
Middle age	−0.017***	−0.061***	−0.154***
	(0.004)	(0.011)	(0.021)
Pensioner	−0.024***	−0.084***	−0.209***
	(0.004)	(0.011)	(0.021)
Youth squared	–	–	–
Middle age squared	–	–	–
Pensioner squared	–	–	–
Constant	0.986***	0.970***	0.928***
	(0.004)	(0.011)	(0.021)
RMSE	0.029	0.052	0.037
Observations	12,420	12,420	12,420
Number of pid	8,724	8,724	8,724

***$p<0.01$, **$p<0.05$, *$p<0.1$

2020). By so doing people can establish innovative ideas for generating income, wealth, and thus increase consumption.

When analyzing other socio-demographic variables it is evident that gender, no landownership, and household size have a positive relationship with vulnerability to poverty whereas employment status, geography, the middle aged, pensioners, and race have a negative nexus. For instance, being a female will increase the probability of falling into poverty in the future by 1.3% relative to the male counterpart based on the upper bound poverty line (Poor ≤ R1268). This occurs because females are paid far less than males (Geiler and Renneboog 2015). When examining race, it is apparent that being a Coloured person will reduce vulnerability to poverty further by 14.9% relative to an African person. Moreover, being Asian/Indian and White will decrease vulnerability to poverty by a greater magnitude than both Africans and Coloureds. This is because they were previously disadvantaged (Gumede and Biyase 2016). The relationship between age and vulnerability to poverty depends on which age group is analyzed. For instance, the middle-aged group has a lower probability of reducing vulnerability to poverty than pensioners because most of the middle-aged group incurs a lot of debt. Therefore, age follows an inverted u-shape (Sekhampu 2013). Moreover, people who do not own land are vulnerable to falling into poverty in the future because there is no land to generate an income from. The interaction term between education and race shows a negative relation with poverty vulnerability. For instance, a Coloured person with a primary education will reduce poverty vulnerability by 5.3% further than an African person with no schooling (based on Poor ≤ R1268 poverty line). Given the social and economic ills perpetuated in the apartheid regime, this social construct plays a vital role in explaining

Table 5 Estimations using the HT model based on three poverty lines. *Source*: NIDS data (Wave 1—5) estimation by the author, robust standard errors in parentheses

Variables	Poor ≤ R1268	Poor ≤ R840	Poor ≤ R585
Primary	− 0.010**	− 0.049***	− 0.127***
	(0.004)	(0.007)	(0.010)
Secondary	− 0.028***	− 0.102***	− 0.250***
	(0.005)	(0.008)	(0.010)
Tertiary	− 0.090***	− 0.243***	− 0.444***
	(0.005)	(0.009)	(0.011)
Coloured	− 0.149***	− 0.299***	− 0.367***
	(0.003)	(0.005)	(0.006)
Asian/Indian	− 0.583***	− 0.682***	− 0.558***
	(0.007)	(0.011)	(0.013)
White	− 0.875***	− 0.649***	− 0.307***
	(0.007)	(0.011)	(0.013)
Primary_Coloured	− 0.053*	− 0.055*	− 0.061*
	(0.029)	(0.031)	(0.009)
Primary_Asian/Indian	0.021	0.034	0.051
	(0.061)	(0.016)	(0.041)
Primary_White	0.066***	0.071***	0.091***
	(0.011)	(0.015)	(0.010)
Secondary_Coloured	− 0.073	− 0.022	− 0.055
	(0.028)	(0.028)	(0.028)
Secondary_Asian/Indian	− 0.099	− 0.132	− 0.332
	(0.042)	(0.069)	(0.07)
Secondary_White	0.230***	0.413***	0.583***
	(0.010)	(0.010)	(0.010)
Tertiary_Coloured	− 0.128***	− 0.129***	− 0.141***
	(0.032)	(0.002)	(0.018)
Tertiary_Asian/Indian	− 0.429***	− 0.521***	− 0.563***
	(0.067)	(0.017)	(0.064)
Tertiary_White	0.520***	0.554***	0.592***
	(0.012)	(0.018)	(0.014)
Female	0.013***	0.027***	0.038***
	(0.001)	(0.002)	(0.003)
Employed	− 0.030***	− 0.087***	− 0.147***
	(0.001)	(0.002)	(0.002)
Urban	− 0.024***	− 0.069***	− 0.103***
	(0.002)	(0.003)	(0.003)
Farm	− 0.024***	− 0.066***	− 0.108***
	(0.003)	(0.004)	(0.005)
Household size	0.004***	0.014***	0.032***
	(0.000)	(0.000)	(0.000)

Table 5 (continued)

Variables	Poor ≤ R1268	Poor ≤ R840	Poor ≤ R585
No landownership	0.003***	0.003**	−0.003
	(0.001)	(0.001)	(0.002)
Youth	0.012	0.022	0.003
	(0.010)	(0.016)	(0.021)
Middle age	−0.018*	−0.058***	−0.139***
	(0.010)	(0.016)	(0.021)
Pensioner	−0.021**	−0.068***	−0.178***
	(0.010)	(0.017)	(0.022)
Youth squared	–	–	–
Middle age squared	–	–	–
Pensioner squared	–	–	–
Constant	0.928***	0.928***	0.928***
	(0.021)	(0.021)	(0.021)
Observations	12,420	12,420	12,420
Number of pid	8724	8724	8724

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

why certain races are more impoverished than others (Baiyegunhi et al. 2010). Following the available literature these relationships are in agreement with economic theory (Christiaensen and Boisvert 2000; Günther and Harttgen 2009).

However, when it comes to geography, it is expected that a person living in an urban area will reduce their vulnerability to poverty (i.e., 10.3% based on Poor ≤ R585) by a greater margin than those living on farms (i.e., 10.8% based on Poor ≤ R585) because there are better job opportunities in urban areas. Yet based on these results this is not the case (Deressa 2013). Nkrumah et al. (2021) found the same results and argue that because those in urban areas rent as opposed to own their properties unlike people living on farms. Thus, this allows farm owners to generate income from their land, and hence, they can reduce vulnerability to poverty at a greater margin than urban dwellers. Sekhampu (2013) on the other end argues that in the process of land reforms most previously disadvantaged South Africans opted for monetary compensation instead of land redistribution. Moreover, the standard errors presented in the above results are very low which suggests that there is no misspecification of the models.

5 Conclusion

This study intended to contribute to the growing body of literature that analyzes the link between education and vulnerability to poverty by applying the disaggregating procedure in South Africa. Most of the studies that have investigated the nexus in question have only disaggregated education which has led to the failure of also being able to analyze poverty variability (Christiaensen and Subbarao

2005; Cruces et al. 2011; Haq 2015). Therefore, in quest of filling the above-mentioned gap this paper adopted the VEP approach, pooled OLS, REM, and HT estimation techniques to assess the education–vulnerability to poverty nexus. As far as we know no study has simultaneously disaggregated education and vulnerability to poverty.

In conducting a study that assessed the link between education and vulnerability for nine years in South Africa we found that there is a statistically significant negative relationship between education and vulnerability to poverty across all levels of education. For instance, an individual with a primary school education will reduce their vulnerability to poverty by 1%. However, in comparison with secondary and tertiary education, the individual reduced their vulnerability to poverty much further. Thus, this highlights a particular pattern that the more education a person has led to a higher reduction of their vulnerability to poverty. These findings agree with some studies (inter alia: Christiaensen and Subbarao 2005; Cruces et al. 2011; Haq 2015; Pritchett et al. 2000). One study that does agree is conducted by Deressa (2013) who argued that a household head with a primary education decreased vulnerability to poverty by 13.6% whereas with a secondary education it was reduced by 49.5%.

A second trend that is evident is that as the individual's future consumption falls below a lower poverty line, any level of education will have a bigger impact in reducing their vulnerability to poverty compared to the less restrictive poverty line. For example, those who are vulnerable to poverty in the next period based on the food poverty line (Poor ≤ R585) will be able to reduce that probability by 25% with a secondary education compared to a 2.8% reduction in poverty vulnerability based on the upper bound poverty line (Poor ≤ R1268). Unfortunately, there is no study that we can compare to our results on poverty variability within the context of different levels of education.

Based on the two findings it is evident that education is a reliable instrument for reducing poverty vulnerability. Therefore, we recommend the provision of quality education as it will address the issue of unemployment and reduce the gap between skills available and skills required. However, the most impoverished group must be prioritized. Furthermore, the 30% pass requirement needs to be increased to 50%, qualified teachers must be employed and reeducated continuously, and the government must implement a futuristic curriculum that is integrated throughout the three levels of education. The government ought to have an agency that holds authorities account for misusing educational investment thus ensuring the development of schools' infrastructure.

This study was not able to examine the nexus in question of developing countries in comparison with developed states so as to evaluate whether the level of a country's development affects the relationship between vulnerability to poverty and education. Thus, it is appropriate to suggest that further research is required to enhance the contribution of the existing literature.

Declarations

Conflict of interest The author(s) declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

Consent for Publication This manuscript is an original work and has been done by the author(s). NN, TM, and CSS are aware of its content and approve its submission. It is also important to mention that the manuscript has not been published elsewhere in part or in entirety and is not under consideration by another journal. The author(s) have given consent for this article to be submitted for publication in *Economic Change and Restructuring*.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the author(s).

References

Aikaeli J, Garcés-Urzainqui D, Mdadila K (2021) Understanding poverty dynamics and vulnerability in Tanzania: 2012–2018. Rev Dev Econ 25(4):1869–1894

Ajuruchukwu OBI, Sanelise TAFA (2016) The determinants of household poverty in South Africa. Afr Public Serv Deliv Perform Rev 4(4):516–538

Baiyegunhi LJS, Fraser GCG, Darroch MAG (2010) Credit constraints and household welfare in the Eastern Cape Province, South Africa. Afr J Agric Res 5(16):2243–2252

Becker GS (1962) Investment in human capital: a theoretical analysis. J Polit Econ 70(5):9–49

Bergh A, Nilsson T (2014) Is globalization reducing absolute poverty? World Dev 62:42–61

Brown P, James D (2020) Educational expansion, poverty reduction and social mobility: reframing the debate. Int J Educ Res 100:101537

Calvo C, Dercon S (2013) Vulnerability to individual and aggregate poverty. Soc Choice Welf 41(4):721–740

Celidoni M (2015) Decomposing vulnerability to poverty. Rev Income Wealth 61(1):59–74

Chaudhuri S (2003) Assessing vulnerability to poverty: concepts, empirical methods and illustrative examples. Department of Economics, Columbia University, New York, p 56

Christiaensen LJ, Boisvert RN (2000) On measuring household food vulnerability: Case evidence from Northern Mali

Christiaensen LJ, Subbarao K (2005) Towards an understanding of household vulnerability in rural Kenya. J Afr Econ 14(4):520–558

Cruces G, Lanjouw PF, Lucchetti L et al (2011) Intra-generational mobility and repeated cross-sections: a three-country validation exercise. World Bank Policy Research Working Paper (5916)

Deressa DF (2013) Assessing households vulnerability to poverty in rural Oromiya-Ethiopia. J Econ Sustain Dev 4(8)

Donou-Adonsou F, Pradhan G, Basnet HC (2021) Human capital expenditure and income in developing countries. J Dev Areas 55(3):113–128

Dzidza PM, Jackson I, Normanyo AK et al (2018) Educational policies on access and reduction of poverty: the case of Ghana. Int J World Peace 35(2):53–83

Geiler P, Renneboog L (2015) Are female top managers really paid less? J Corp Finan 35:345–369

Gujarati DN, Porter DC (2009) Causality in economics: the Granger causality test basic econometrics, 5 international. McGraw-Hill, New York, p 652

Gumede V, Biyase M (2016) Educational reforms and curriculum transformation in post-apartheid South Africa. J Environ Econ 7(2):69–76

Günther I, Harttgen K (2009) Estimating households vulnerability to idiosyncratic and covariate shocks: a novel method applied in Madagascar. World Dev 37(7):1222–1234

Guo L, Huang J, Zhang Y (2019) Education development in China: education return, quality, and equity. Sustainability 11(13):3750

Haq R (2015) Pakistan institute of development economics, islamabad quantifying vulnerability to poverty in a developing economy. J Dev Econ 54(4):721–740

Haughton J, Khandker SR (2009) Handbook on poverty and inequality. World Bank Publications, Washington DC

Hoddinott J, Quisumbing A (2010) Methods for microeconometric risk and vulnerability assessment. In: Fuentes-Nieva R, Seck PA (eds) Risk, shocks, and human development. Palgrave Macmillan, London, pp 62–100. https://doi.org/10.1057/9780230274129_4

Holzmann R, Jorgensen S (1999) Social protection as social risk management. The World Bank, Washington DC

Hong AH, Sullivan FR (2013) Aggregate and disaggregate data in measuring school quality. Educ Tech Res Dev 57(5):613–627

Imai KS, Gaiha R, Kang W (2011) Vulnerability and poverty dynamics in Vietnam. Appl Econ 43(25):3603–3618

Kamanou G, Morduch J (2002) Measuring vulnerability to poverty (No. 2002/58). WIDER Discussion Paper

Kawulur AF, Sumual TEM, Kawulur H et al (2019) Human development and poverty reduction through economic growth. Int J Bus Manag Sci 9(1):19–31

Klasen S, Waibel H (2013) Vulnerability to poverty. In: Klasen S, Waibel H (eds). Palgrave Macmillan, London

Ligon E, Schechter L (2002) Measuring vulnerability: the director's cut (No. 2002/86). WIDER Discussion Paper

Mahanta R, Das D (2015) Vulnerability to poverty: a survey. Available at: SSRN 2621859/. Accessed 25 Oct 2021

Marginson S (2019) Limitations of human capital theory. Stud High Educ 44(2):287–301

Missiakoulis S (2020) Two and three stage least squares as Aitken estimators. J Stat Econ Methods 9(4):87–89

Ngepah N (2017) A review of theories and evidence of inclusive growth: an economic perspective for Africa. Curr Opin Environ Sustain 24:52–57

NIDS (2017) Data access. Available at: Download NIDS Wave Datasets | Register with NIDS - | NIDS (uct.ac.za). Accessed 9 Jul 2021

Nkrumah RK, Annim SK, Afful B (2021) Household social expenditure in Ghana: examining the ex-post effects and vulnerability to poverty. Soc Sci 10(2):40

Nortje MJ (2017) The effect of poverty on education in South Africa. Educor Multidiscip J 1(1):47–62

Omotoso KO, Adesina JO, Gbadegesin TF (2020) Children on the edge: estimating children's vulnerability to multidimensional poverty in post-apartheid South Africa. Child Indic Res 13(4):1155–1174

Pham ATQ, Mukhopadhaya P, Vu H (2021) Estimating poverty and vulnerability to monetary and non-monetary poverty: the case of Vietnam. Empir Econ. https://doi.org/10.1007/s00181-020-01991-4

Pritchett L, Suryahadi A, Sumarto S (2000) Quantifying vulnerability to poverty: a proposed measure, applied to Indonesia (No. 2437). World Bank Publications, Washington DC

Ravallion M (1988) Expected poverty under risk-induced welfare variability. J Econ 98(393):1171–1182

Sekhampu TJ (2013) Determinants of poverty in a South African township. J Soc Sci 34(2):145–153

StatsSA (2020) Five facts about poverty in South Africa, survey find. Available at: http://www.statssa.gov.za/?p=12075/. Accessed 12 Jul 2021

Swain RB, Floro M (2012) Assessing the effect of microfinance on vulnerability and poverty among low income households. J Dev Stud 48(5):605–618

Vuko T, Čular M (2014) Finding determinants of audit delay by pooled OLS regression analysis. Croat Oper Res Rev. https://doi.org/10.17535/crorr.2014.0030

Wooldridge J (2018) Introductory econometrics, applied discrete-choice modelling, 5th edn. South-Western Cengage learning, Mason

Yang Y, Guo X (2020) Universal basic education and the vulnerability to poverty: evidence from compulsory education in rural China. J Asia Pac Econ 25(4):611–633

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.