

Empirical Paper

Łukasz Goczek^{1,*}, Bartosz Witkowski², Ewa Witkowska³

Does an increase in education quality cause developing countries to catch up?

<https://doi.org/10.2478/ijme-2022-0028>

Received: October 6, 2022; accepted: November 21, 2022

Abstract: We investigate whether increasing the education quality causes increases in economic growth allowing poorer countries to catch up. To this end, we extend Nelson-Phelps's classic paper by introducing differences in education quality (proxied by students' performance on the Program for International Student Assessment [PISA] test) in a leader-follower type of growth model with knowledge diffusion. We use students' performance in a standardized international PISA test to measure education quality's impact on economic growth using a panel Vector Error Correction allowing for cross-correlation in the co-integration analysis in a set of all countries observed over the years 1975–2018. Additionally, we consider the possible reverse causality that characterizes economic development and the quality of education. Our results confirm the relevance of education quality as a significant causal factor of economic growth. However, further analyses indicate that the discussed relation is long-run, which may have important implications for policymakers.

Keywords: economic growth, economics of education, educational policy, international education/studies

JEL Classification: C4; C5; I2; O1

1 Introduction

For decades the differences in human capital have been considered to be one of the major factors that explain why some countries are richer than others. While neoclassical and endogenous growth theory identified and analyzed numerous important factors affecting these output differences such as foreign trade, government consumption, geography, and political institutions to mention just a few [Barro, 1991; Acemoglu et al., 2001; Jones, 2014; Cieřlik and Goczek, 2018], one of these important determinants has been the human capital which up to a point would be associated with education-related indicators frequently understood as average educational attainment. Along these policy lines, the vast majority of countries have targeted the Organization for Economic Co-operation and Development (OECD) leaders in their efforts to improve their educational attainment in the past decades. However, some countries have consistently lagged in output per capita despite having introduced an education reform aimed at increasing average years of schooling. This brings about the idea that the lion's share of the income gap between poor and rich countries can be attributed not just to the length or population enrolment of education but also to its quality. Recent works by Erosa et al. [2022], Hanushek and Woessmann [2008, 2012], Manuelli and Seshadri [2014], Tamura [2001], and Schoellman [2012] confirm that the role of quality of education may be underappreciated as opposed to average attainment or years of schooling.

*Corresponding author: Łukasz Goczek, Faculty of Economic Sciences, University of Warsaw, Warsaw, Poland.

E-mail: lgoczek@wne.uw.edu.pl

Bartosz Witkowski, Collegium of Economic Analysis, SGH Warsaw School of Economics, Warsaw, Poland.

Ewa Witkowska, The Institute of Psychology, The Maria Grzegorzewska University, Warsaw, Poland.

The main objective of this paper is to identify the role of the quality of education in economic growth. We extend the classical Nelson and Phelps [1966] approach by introducing differences in education quality in a leader-follower type of growth model with knowledge diffusion. As in Hanushek and Woessmann [2012, 2015], we use students' performance in a standardized international test (the Program for International Student Assessment [PISA] test) to approximate the quality of education in a set of all countries observed over the years 1975–2018. This issue is important also from a political point of view: a confirmed crucial role of education quality may encourage governments to rethink their policy of financing education. This, however, raises a question about the stability of the education quality and economic growth nexus: we intend to verify whether the short and long-run effects are equal and what the pattern otherwise is. Moreover, we consider the possible reverse causality that characterizes economic development and the quality of education. Finally, in most papers, any form of co-variation of education and GDP is assumed to be the confirmation of an existing impact of the education factor on economic growth. This, however, is not at all obvious and we set as our target in identifying the direction of this relation.

This paper is arranged as follows. In Section 2, we briefly outline the data and procedures used in this paper to allow for the reproduction of procedures by an independent researcher. We review the relevant theory in Section 3, presenting adequate literature discoveries and a theoretical model that provides a theoretical background for the role of education quality in GDP growth. In Section 4 we discuss the empirical model and the econometric approach used in this paper. Section 5 presents the main empirical findings and their discussion and briefly concludes.

2 Material and methods

To verify the research hypothesis, we use two main sources of data. The quality of education is assessed with PISA test results which are freely available from the OECD website. These data are extrapolated backward with the use of linear panel data Generalized Least Squares approach allowing for autocorrelation and heteroskedasticity of the errors as described by Goczek et al. [2021]. The macroeconomic country data used in the calculation are taken from the World Bank (WB), International Money Fund (IMF), and the OECD. Table 1 presents the descriptive statistics.

The main part of the research is based on the vector error correction model (VECM; Keblowski, 2016, for instance) explaining GDP as a function of different factors, including the education quality. Its application requires the processes to be co-integrated, thus we test for the series cross-correlation with Pesaran's [2004] test, and having found cross-dependence we use the appropriate stationarity Im, Pesaran, and Shin's (CIPS) test (2003). Next, we use Westerlund's [2007] co-integration test to confirm the existence of the co-integrating relation. The existence of co-integration allows for the estimation of the considered VECM which we carry out with the use of the dynamic fixed effects (DFEs) estimator. As the last step, we test for causality by extending the classical Granger's method with Dumitrescu and Hurlin's [2012] individual coefficients approach. All the estimations are carried out with Stata with the use of either the standard procedures or add-ins available freely for download and easy to find with Stata's search engine.

Table 1. Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
logGDP	4,942	8.942929	1.233535	5.508054	11.82894
OPEN	7,538	80.11214	51.46048	0.0209992	531.7374
FDI	5,535	1.86752	12.44171	-202.8239	247.9023
PISAAVE	4,154	420.173	50.2345	312.482	550.1743
PISASCIE	4,154	424.9734	48.47972	323.4689	550.6454
PISAMAT	4,154	419.5068	51.97938	319.7289	566.1837
PISAREAD	4,154	416.0389	51.054	294.2483	543.4544

3 Theoretical background

3.1 Literature review

Human capital can be defined as “the set of intangible resources embedded in the labor factor which have improved its productivity” [Goldin, 2016]. These resources are related to knowledge and skills acquired through education and experience [Schultz, 1961; Becker, 1962]. There exists a large body of microeconomic evidence of a strong relationship between human capital and earnings, stemming from the classical Mincer model which uses years of schooling as a key determinant of wages. Mincer’s model implies that a change in a country’s average level of schooling should be the fundamental factor of economic growth [Krueger and Lindahl, 2001]. Owing to this finding, several studies adjust the workforce for improvements in human capital measured by their educational attainment. However, on the macroeconomic level, numerous studies have failed to find a significant relationship between economic growth and educational attainment [Bosworth and Collins, 2003]. Elbakidze and Jin [2015] go even one step further and show that the increased secondary attainment resulted mostly in an increased amount of violence rather than educational effects.

While on the microeconomic level the accumulation of human capital improves labor productivity and increases the returns to capital and a well-educated workforce is also essential for the creation and diffusion of technology, the issue is the distinction between the quantity (an example of which could be the educational attainment) and the quality of education. Stern [2000] documents one example of the increased quantity not necessarily being followed by the desired economic effects. This could be attributed to the measurement problem. In a comprehensive study, Benos and Zotou [2014] show that educational attainment, frequently used in empirical studies, is an inaccurate measure of human capital since it measures education quantity, which is more or less stable, while education quality varies widely across countries and periods. To investigate this difference, Schoellman [2012] adjusts years of schooling by considering the education quality to explain the cross-country output differences per worker. He finds that the cross-country differences in education quality are roughly as important as cross-country differences in its quantity. This increases the total contribution of education differences from 10% to 20% of output. In consequence, as long as it can be measured, the quality of education seems a better measure to describe the quality of human capital in the macroeconomic production function than the years of schooling or educational attainment. In this paper, we follow the approach of Hanushek and Woessmann [2012] and use the OECD’s PISA results to describe each country’s quality of education in the primary and early secondary stages. The PISA measures 15-year-olds’ ability to solve real-life-like types of problems with their reading, mathematics, and science knowledge and skills. The main strength of this cyclical program that has lasted over the last two decades is the fact that in each participating country a sample of at least a few thousand students is drawn. Those take a test whose results can be compared across different countries. The fact that the analysis is repeated every three years with an increasing number of participating countries allows for the cross-country dynamic comparison of educational achievements measured by the test results.

There exists some research on the impact of educational quality measured by PISA scores on economic growth, however, most of it is based on cross-sectional regressions or correlations [Goczek et al., 2021]. Breton [2011] challenges Hanushek and Woessmann’s [2008] contention that the quality and not the quantity of schooling determines a nation’s rate of economic growth by arguing this cross-sectional method is flawed. Nevertheless, a few methodologically more advanced papers, mostly by Hanushek and Woessman, should be also mentioned. In their paper, Hanushek and Woessmann [2012] address the problem of causality by trying to eliminate gradually the influence of potentially relevant features on growth. This is aimed at avoiding the potential spuriousness of the regression. The authors argue that the education quality-GDP growth relation is not instantaneous and they use lagged education measures, which enforces them to use the results of tests performed even before PISA was introduced (such as Trends in International Mathematics and Science Study TIMSS). In line with this, OECD [2010] mentions the problem of causality and performs some robustness checks to determine the direction of the existing relation, others do not always do so – for example, Chen and Luoh [2010] simply regress growth on test scores. They examined the relationship

between test scores (mathematics and science) and cross-country income differences. The authors pose a question of whether test scores are good indicators of workforce quality. The obtained results suggest that the strong link between test scores and cross-country income differences cannot be confirmed. Instead, they discover that the per capita number of researchers involved in R&D as well as the per capita number of scientific and technical journal articles could better account for the cross-country income differences. This might be due to improper lagging of the test scores. However, it might also mean that the general proxy for the country's education quality (which PISA intends to measure at the mid-teen level) is not sufficient. Indeed, it can be construed that the skills as they are measured in PISA tests are not key factors to the future human capital quality. Campbell and Üngör [2020] develop a national accounting exercise with human capital to account for schooling, cognitive skills, and health. We present group comparisons, finding for some regions values as large as 40–50%. The group comparisons of education quality indicate, for some regions, values as large as 40–50%. Moreover, Balart et al. [2018] decompose PISA test scores into a cognitive- and non-cognitive components to find that both components are positively associated with economic growth. We believe that the data could be decomposed further due to increased data availability. In addition, one of the reasons for the lack of statistical significance of education in the convergence process could be the presence of omitted control variables like openness as indicated in the meta-analysis of the literature by Benos and Zotou [2014], which could be seen as a significant critique of many previous studies.

Hanushek and Woessmann [2015] present evidence that when controlling for cognitive skills acquired during the education process (a measure of education quality), the initial value of years of schooling does not have any effect on the average annual growth rate of GDP per capita in the period 1960–2010. The authors address the problem of causality in yet one more way, namely instead of analyzing the level-level relation, they regress deviation on deviation. Still, most research is based just on cross-sectional data with few researchers applying instrumental variables regression. We found no trace of dynamic panel data model analysis in the literature that makes use of PISA (and similar) test measurement of human capital.

We argue that the possibility of controlling for endogeneity is not appropriately undertaken in the surveyed articles. To our knowledge, no co-integration analysis is present in the literature. The latter seems to be particularly important when time-series (or panel) data from different countries are used. While it is commonly known that the GDP series are almost always non-stationary, the human capital quality measured by the results of such tests as PISA is likely to share the same property due to the existence of the so-called “Flynn effect.” The latter expression is used to name the regular and systematic worldwide increase of mean results in standardized intelligence test performance, observed since the beginning of the 20th century [Flynn, 1984, 1987]. Based on the analysis of data from 14 different countries a considerable increase in intelligence test scores of 5 to 25 points on the deviation intelligence quotient scale for each generation has been demonstrated. The rise of scores is usually found to be greater on tests of fluid intelligence, such as Raven Progressive Matrices [Brouwers et al., 2009], than on tests of crystallized intelligence, especially nonverbal IQ tests. However, a considerable increase is recorded also for tests that measure crystallized intelligence [Flynn, 1987] such as the Similarities Scale in Wechsler's Intelligence Test. Given the above – and also the inspection of empirical data – one can expect the test results to be generally improving over time. The co-integration analysis would then seem a reasonable tool to avoid the risk of attaining spurious regression or reverse causality in which GDP increase can lead to better education quality as countries can invest in better education to make their income higher. Although there exists much more research whose authors try to attain the PISA score determinants as well as PISA score implications, there seems to be space for a lot more, especially on the country aggregated data, as well as the potential non-stationarity of the data, which suggests the use of co-integration techniques, not present in the literature in this context.

3.2 Theoretical model

The model which is employed investigates the relationship between education quality and growth in an open economy version of the leader-follower model of Nelson and Phelps [1966] that allows for conditional convergence in growth rates between countries but not in the levels of income. In the steady state, the leader and the follower exhibit the same growth rate. The convergence of the follower to the leader occurs

through the international diffusion of knowledge through international trade with the leader economy. In our setting increasing education quality allows for a quicker diffusion of technology.

The model builds on earlier studies by Cieřlik and Goczek [2018] as well as Cieřlik and Tarselewska [2011] who employed the leader-follower model in a different context of privatization and corruption. In this paper, we extend this setting to show that education quality can impact the steady-state level of income and the rate of growth in the follower economy during its transition to the steady-state but not its steady-state level of growth. In the steady state, the growth in the follower economy matches the growth in the leader economy.

We set out by considering an autarky case with no international knowledge diffusion. We then introduce knowledge diffusion enabled by international openness in the next step. Let the technological leader be denoted as object 1 and the follower as object 2. The aggregate production functions can be then written as:

$$Y_i = A_i K_i^\alpha L_i^{1-\alpha} \quad \text{for } i = 1, 2 \quad (1)$$

where: $0 < \alpha < 1$.

In the case of autarky, there is no trade and investment so no international knowledge diffusion takes place as each country has to come up with all varieties of capital goods alone. In this case, the behavior of the follower equals the behavior of the leader. The critical difference lies in the costs of inventing a new variety of capital goods in the i -th country (η_i , $i = 1, 2$) expressed in terms of GDP, which are higher in the former than in the latter of the considered countries, i.e. $\eta_2 > \eta_1$.

In each country, the output Y_i depends on inputs of capital (K_i) and labor (L_i), as well as the country-level productivity parameter A_i . This parameter is usually assumed to capture cross-country differences in national policies, such as infrastructure maintenance, political stability, etc. Nelson and Phelps [1966] already thought of growth as being generated by productivity-improving technology adaptations, which development rate would depend upon the stock of human capital. Accordingly, in our analytical framework, we assume that productivity can be positively influenced by human capital. In our case it is understood as the education quality:

$$A_i = f(\text{Education Quality}_i) \quad (2)$$

We expect that better education makes it easier to innovate. Thus, the variable A_i represents the technology level at time t . Both labor input L_i and technology level A_i are assumed to grow at constant rates. This formulation of technology is referred to as labor augmenting. We should think of technology as making labor more “effective” so the effective labor input equals $A(t)L(t)$. This allows us to focus on the specific ratio of capital to effective labor. This assumption means that the meaning of education quality is relative: the ratio of productivities in the analyzed countries is proportional to the ratio of their education quality.

The capital K_i can be defined as a set of varieties of capital goods using constant elasticity of substitution between different varieties:

$$K_i = \left[\sum_{j=1}^{N_i} x_{ij}^\alpha \right]^{1/\alpha} \quad \text{for } i = 1, 2, \quad (3)$$

where x_{ij} denotes the input of the j -th variety of the capital goods in the i -th country, N_i is the number of varieties that exist in the i -th country, and α is the share of the input.

In the beginning, economies are closed to both international trade and FDI. This imposes that the national output equals domestic expenditures which are distributed between consumption (c_i) and production of varieties of the capital good (x_{ij}). A simplifying assumption is that each unit of c_i or x_{ij} requires one unit of Y , and that there is no entry cost into the R&D sector. A free-entry assumption requires that the present value of innovation profits should be equal to its fixed cost. The assumption regarding the fixed cost of innovation implies that the interest rate (r_i) in the i -th economy is likewise constant and equal to the ratio of profits (π_i) over the cost of innovation (η_i).

In both countries, both innovating and imitating firms face a demand for a variety of capital goods, which comes from the profit maximization problem of the final goods producers. The demand is given by:

$$x_i = L_i A_i^{1/(1-\alpha)} p_i^{1/(1-\alpha)} \quad (4)$$

To make the innovation profitable, patents are introduced. The patents yield monopoly power to the owner – the innovating representative firm. As the firms are homogenous, profit maximization causes prices to be the same across all varieties. Thus, the profit maximization of the innovator $\pi_i = (p_i - 1)\pi x_i(p_i)$ concerning the price of j -th variety of the capital good which it produces yields a monopoly price $p_{ij} = p = 1/\alpha > 1$. It is straightforward to see that the total quantity of each variety supplied is constant in each country as well:

$$x_{ij} = x_i = L_i A_i^{1/(1-\alpha)} \alpha^{2/(1-\alpha)} \quad (5)$$

Consumers in both countries are assumed to be Ramsey-type with infinite time horizons who maximize standard intertemporal utility function at time zero:

$$U_i = \int_0^\infty e^{-\rho t} \frac{c_i^{1-\theta} - 1}{1-\theta} dt \quad \text{for } i = 1, 2, \quad (6)$$

where $\rho > 0$ is the constant rate of time preference and $\theta > 0$ is the constant magnitude of the elasticity of the marginal utility of consumption that is assumed to be the same across countries like in most research (which, based on the equimarginal principle, can be motivated easily assuming constant price relations of the goods in the market).

Standard maximization of utility, subject to the intertemporal budget constraint yields the rate of consumption growth:

$$\frac{\dot{c}_i}{c_i} = \frac{r_i - \rho}{\theta} \quad \text{for } i = 1, 2. \quad (7)$$

The growth rate of consumption is constant since the interest rate in each country is constant. Therefore, the growth rates of the number of varieties N and output Y in the steady state must be the same as the growth rate of c . In the closed economy, the steady-state rate of growth is higher in the leader country than in the follower country because the leader has lower innovation costs than the follower.

Let us now consider a case of an open economy. In this case, a firm in the follower country can invent a new variety of goods or imitate a product that is already known in the leader country. However, it is cheaper to imitate pre-existing goods from the leader country because the imitation cost (v_2) is lower than the innovation cost (η_2). For this reason, in the open economy equilibrium, the companies in the follower's economy do not innovate. It is assumed that the imitation cost decreases with the number of varieties manufactured by multinational firms (N_1) as well as on the number of tradable varieties (m) relative to the total number of varieties produced in the follower economy (N_2) (knowledge spillovers, since they are imitated). The follower country can only imitate some of the unimitated subsets of N_1 varieties that are established in the leader country. Thus, it is assumed that as N_2 grows relative to N_1 , the imitation cost goes up. This means that the easiest-to-copy varieties are imitated first but on the other hand, imitation costs get higher with more complex varieties. As a result, we observe a catching-up effect as the follower countries with lower N_2/N_1 ratios experience lower imitation costs and tend to grow faster. This intuition is described as the imitation cost function:

$$v = \eta_2 \left(\frac{N_2}{N_1} \right)^\sigma \left(\frac{n_1}{N_2} \right)^{-\varepsilon} \left(\frac{m}{N_2} \right)^{-\delta} \quad (8)$$

where $\rho, \varepsilon, \delta > 0$ are constant parameters.

In the steady state, the growth rate of per capita income in the follower country must equal the growth rate in the leader country. The equality of consumer preferences in both countries causes the international interest rate to equalize. Therefore, the steady-state ratio of per capita incomes $(y_2/y_1)^*$ is:

$$\left(\frac{y_2}{y_1}\right)^* = \left[\left(\frac{A_2}{A_1}\right)^{(1+\sigma)} \left(\frac{L_2}{L_1}\right)^{(1-\alpha)} \left(\frac{\eta_1}{\eta_2}\right) \left(\frac{n_1}{N_2}\right)^\varepsilon \left(\frac{m}{N_2}\right)^\delta \right]^{1/\sigma}. \quad (9)$$

Eq. (8) demonstrates that the steady-state per capita income of both countries is not equal and a constant income gap may continue, depending on, among others, the national productivity parameters as well as external openness. In this model, education quality does not have permanent effects on the rate of growth in the steady state as the rate of growth in the follower country converges in the long run to the rate of growth of the leader country. Nevertheless, these parameters determine the level of income in the steady state and affect the rate of growth during the transition to the steady state. In particular, a higher level of education quality translates into quicker transitional growth, having controlled for country characteristics that impact their steady-state levels of income, namely their international openness to trade and FDI.

Therefore, the main hypothesis is that education quality has a positive impact on economic growth, having controlled for other effects. As a result, the hypotheses based on the formal theoretical model above that may be tested empirically are:

Hypothesis 1: Education quality increases economic growth in the short run

Hypothesis 2: Increases in education quality have a permanent effect on growth

Hypothesis 3: A higher level of external openness (higher Foreign Direct Investment and Trade) results in more growth

Hypothesis 4: Initial level of GDP per capita is negatively related to growth (i.e., conditional convergence in GDP levels among follower countries occurs).

4 Empirical verification

We base the analysis on the model proposed in Section 3.2 of the article. Importantly and contrary to Hanushek and Woessmann [2010], we construct a model in which we explain the GDP change as a function of not only the quality of education but also the necessary and relevant economic characteristics captured in the theoretical GDP growth model. Furthermore, a key issue in economics is the distinction between the short run and the long run while the most frequently used estimators were derived under the assumption of short- and long-run homogeneity. In such models, the estimated coefficients are held constant across all units in the sample; therefore, the coefficients are the same for each analyzed country over time. However, Pesaran [2007] note that this situation does not have to be in line with reality, especially in the short run. For this reason, the commonly used estimators may not be consistent, and long-run coefficients may be biased. The details of this methodology are explained in detail by Blackburne and Frank [2007]. The methodology fits the needs of our theoretical model and data sample characteristics, since it allows us to decompose the short and the long run, allows for non-stationarity of the variables, panel cross dependence, and last, not least allows us to tackle the short country dimension as compared to the time dimension of the panel.

Therefore, the theoretical results from the formal model suggest the use of the VECM and cointegration analysis approach, which we provide within the panel framework:

$$\begin{aligned} \Delta \log GDP_{i,t} = & \gamma_i \Delta \log GDP_{i,t-1} + \theta_1 \Delta OPEN_{i,t-1} + \theta_2 \Delta FDI_{i,t-1} + \theta_3 (\Delta L_s PISA_{i,t-1}) \\ & + \varphi_i (\log GDP_{i,t-1} - \beta_1 OPEN_{i,t} - \beta_2 FDI_{i,t} - \beta_3 (L_s PISA_{i,t})) + \alpha_i + \varepsilon_{i,t} \end{aligned} \quad (10)$$

where:

$\log GDP_{it}$ – natural logarithm of the gross domestic product of country i in year t (Source: World Bank, International Comparison Program database),

$OPEN_{it}$ – the sum of exports and imports to GDP of country i in year t (Source: World Bank national accounts data, and OECD National Accounts data files),

FDI_{it} – the share of foreign direct investment of country i in year t (Source: International Monetary Fund, Balance of Payments database),

$L_s PISA_{i,t}$ – the average score on the considered item of the PISA test of country i in year t lagged by s years, where s is a presumed constant >0 (Source: OECD Programme for International Student Assessment),

θ_i – short-run dynamics coefficients,

γ_i – short-run business cycle dynamics coefficients,

φ_i – country-specific error correction coefficient,

α_i – country-specific fixed effect,

β_k – long-run cointegration coefficients, $k = 1, 2, 3, \dots, n$,

ε_{it} – error term.

The cointegration approach allows us to tackle reverse causality where a GDP increase can lead to better education attainment as countries can invest in better education. This obvious risk of endogeneity raises doubts regarding the causality inference as it may lead to giving too much weight to the impact of education on GDP and may overestimate its true impact. Our method provides consistent coefficients (despite the possible presence of endogeneity) because it includes lags of both dependent and independent variables which are included in the estimation which allows us to interpret the results as Granger-type causality.

The PISA scores in different types of skills and their average across all the considered skill groups are considered separately. Thus, in empirical models, the PISA variable in Eq. (10) is replaced with PISAAVE (global PISA score), PISAREAD (PISA score for the reading skills), PISAMATH (PISA score for the mathematical skills), and PISASCIE (PISA score for the science skills) respectively in separately estimated versions of Eq. (10).

It should be observed that the PISA test results in the model are taken from the period lagged by s years ($L_s PISA_{i,t} = PISA_{i,t-s}$). The reason for the use of lags was discussed in Section 1: we can use the results of the PISA tests to estimate the quality of education, however, the results are based on the cohort of 15-year-olds at the time of the survey. At least 10 years elapse before the main part of the cohort enters the labor market. Since it is the quality of the human capital that constitutes the labor force, not the whole nation, that needs to be considered as input in the growth model, this requires the PISA scores to be lagged by at least 10 years, which we do. It should though be noticed that in a framework of other growth models the results provided by Goczek et al. [2021] suggest consistency of the results attained in the models with PISA scores being lagged by 10–15 years, which is also the case here).

Testing for cointegration is essential before the estimation of Eq. (9). Determining the issue of non-stationarity of the PISA results time series hinges, however, on the issue of cross-sectional dependency between individual countries which is expected to take place because of many reasons. The most obvious is the openness of the economies that results in the global trends of the education processes as well as the global existence of the Flynn effect itself. In light of these arguments, the results of the contemporaneous correlation across the countries' cross-sectional dependence (CD) test developed by Pesaran [2004] that confirm the existence of cross-country correlation of the particular PISA series would not be a surprise.

This dependency may be troublesome for the first-generation unit root tests used to determine the issue of stationarity of the time series that are used in Eq. (10). The first-generation unit root tests assume cross-sectional independence between observational units. This is restrictive and unrealistic in the majority of macroeconomic applications, like the study of convergence where co-movements of business cycles between economies are obvious. The second-generation unit root tests strive to overcome this issue, relaxing the assumption of cross-sectional independence and tackling potential heterogeneity. The most commonly known second-generation test is the cross-sectional Im, Pesaran, Shin (2003) (CIPS) test. Instead of the standard Augmented Dickey Fuller regression, CIPS tests are based on the cross-sectional ADF regression (CADF), which adds lagged cross-sectional means of individuals to control for effects of the common factors,

Table 2. Pesaran CD test results

	PISAAVE	PISAREAD	PISAMAT	PISASCIE
Breusch-Pagan LM	47,950.26 <0.001	47,579.72 <0.001	46,412.74 <0.001	48,879.8 <0.001
Pesaran scaled LM	563.4106 <0.001	558.7497 <0.001	544.0704 <0.001	575.1033 <0.001
Pesaran CD	202.8806 <0.001	202.8742 <0.001	197.609 <0.001	205.7417 <0.001

CD, Pesaran's cross-sectional dependence test statistics; LM, Lagrange multiplier test statistics.

Note: *p*-values in parentheses.

while the computation of the test statistics and the inference follows the IPS procedure. The CIPS statistic is then computed as a group mean of *t*-statistics obtained from particular CADF equations.

Similar to panel unit root tests, panel cointegration tests strive to provide more reliable results by checking for the presence of cointegration or cross-dependency issues. The most frequently used panel cointegration tests are based on unit root testing of residuals from an Ordinary Least Squares regression, however, the major limitation of such tests is that they cannot accommodate structural breaks and account for cross-dependency which exists in the data which is confirmed by the results of the tests shown in Table 2. The panel cointegration test proposed by Westerlund [2007] copes with this issue by determining structural breaks endogenously. More importantly, however, the Westerlund test accounts for cross-sectional dependency evident in the business cycle data across countries with the use of bootstrapping to compute robust critical values. This is an important issue since the application of cointegration tests assuming independence (Kao, Pedroni) to series that are characterized by cross-sectional dependencies leads to size distortions and low power.

Identifying the existence of co-integrating relation allows for the estimation of the panel VECM Eq. (10). The most frequently used methods to estimate Eq. (10) are the dynamic ordinary least squares (DOLS) and the fully modified ordinary least squares (FMOLS). However, an important shortcoming of both methods is that they do not estimate the parameter of the short-run relations that are as important as the long-run relations. Alternative methods, such as the pooled mean group (PMG) regression, mean group (MG) regression, and the DFE approaches, are available to consider different levels of heterogeneity across countries while estimating both the short-run and the long-run effects simultaneously. Due to insufficient data, only the DFE estimation is feasible. However, the application of the DFE approach is justified in this case given that the countries available in the sample have important similarities and most of the heterogeneity between them is captured by country-specific intercepts. In consequence, the DFE model is appropriate in the considered situation.

The last issue that requires consideration in the paper is the problem of causality. The seminal research by Hanushek and Woessmann seems not to shed sufficient light on this crucial issue which, as emphasized by Altman [2020] is often the case. Thus, we adopt two approaches to causality testing in the panel data framework. The first one consists in treating the panel data as one large stacked set of data, and then performing the Granger Causality test in the standard way, except for not letting data from one cross-section enter the lagged values of data from the next cross-section. This method assumes that all coefficients are the same across all cross-sections. Since this might be doubtful, we also consider another approach adopted by Dumitrescu and Hurlin [2012], who make an opposite assumption, allowing all coefficients to be different across cross-sections.

5 Results and discussion

The main concern associated with the use of PISA results as a measure of the quality of schooling δ and growth factor in the empirical growth model is that the PISA tests have only been performed over

approximately the last 20 years in general and only a few years in a numerous group of countries. To overcome the problem which immediately stems from a low number of available observations we apply the strategy described by Goczek et al. [2021]: we use exponential smoothing to interpolate the PISA scores in the years when the tests were not performed in a given country and apply backward extrapolation of the PISA results to obtain an adequately long series. As an extrapolation tool, we use a dynamic regression model in which we explain PISA scores (for each of the three groups of skills distinguished in the PISA tests: mathematics, science, and reading) as a function of lagged education expenditures [Heyneman, 1995; Hanushek and Woessmann, 2012; Barro and Lee, 2015] and individual country effects. The countries included in the sample include Albania, Argentina, Australia, Austria, Azerbaijan, Belgium, Chile, Colombia, Costa Rica, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hong Kong, Hungary, Iceland, Indonesia, Ireland, Israel, Italy, Japan, Kazakhstan, Korea, Kyrgyz Republic, Latvia, Lithuania, Luxembourg, Macao, Malaysia, Malta, Mexico, Moldova, Netherlands, New Zealand, Norway, Peru, Poland, Portugal, Qatar, Romania, Russia, Singapore, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Thailand, Tunisia, Turkey, UK, USA, and Uruguay.

An adequate GLS estimator allowing for autocorrelation and heteroskedasticity is then applied to yield theoretically lagged PISA country-average results. Naturally, possession of true data without any forms of attrition or imputation is an ideal situation, however, this often is not the case, especially in cross-country studies. In consequence, we partly estimate the missing values. However, it should be emphasized that we use an algorithm based on the estimation of the missing values with the use of an appropriate regression equation with an individual effect, individual autocorrelation, and inclusion of educational expenditures as a regressor.

The PISA results for different skills are country correlated, as was expected. The results of Pesaran's test of CD are provided in Table 2.

The results that confirm the existence of correlation across different PISA test item scores are by no means surprising. On the one hand, to some extent, the 15-year-olds that are tested within the frame of PISA can be – for simplification – considered as members of the groups of either the more or of the less involved in studying. This division does not fully determine the results of the PISA test for each skill but has a certain effect on it. In consequence, the conditional and the unconditional mean result for one skill given the results regarding other skills are not equal. On the other hand, the cross-dependence is partly caused by the construction of the PISA test itself.

This dependency determines the use of CIPS test to determine the orders of integration of the incurred series. Table 3 summarizes the results for the variables included in the model.

The results are clear and indicate the $I(1)$ character of all the series of interest (assuming a 5% level of significance). This is crucial for further steps of analysis. Having identified the existence of cross-correlation as well as equal order of integration across the series, a proper co-integration test is required.

Table 3. CIPS unit root test results

Variable	Levels		First differences	
	Zt-bar	p-value	Zt-bar	p-value
logGDP	44.523	1.00	-1.854	0.03
OPEN	0.01	0.50	-5.525	<0.01
FDI	0.72	0.76	-16.309	<0.01
PISAAVE	1.318	0.91	-17.421	<0.01
PISAREAD	1.159	0.88	-17.024	<0.01
PISAMAT	3.338	1.00	-18.784	<0.01
PISASCIE	-1.907	0.03	-16.632	<0.01

CIPS, Cross-sectional Im, Pesaran, and Shin.

Notes: CIPS test statistics (Zt-bar) for the levels of variables (testing for stationarity of the variables) and first differences (testing for the integration of the first order); p-values in parentheses.

Table 4. Westerlund cointegration test results

Statistic	Value	Z-value	<i>p</i> -value	Robust <i>p</i> -value
PISAAVE				
Gt	−2.655	2.798	<0.01***	0.01**
Ga	−9.173	1.592	0.94	<0.01***
Pt	−15.039	2.881	<0.01***	0.01**
Pa	−8.567	1.054	0.15	<0.01***
PISAREAD				
Gt	−2.738	−3.347	<0.01***	<0.01***
Ga	−8.582	2.117	0.98	0.01**
Pt	−14.547	−2.415	0.01**	<0.01***
Pa	−8.166	−0.673	0.25	<0.01***
PISAMAT				
Gt	−2.569	−2.225	0.01**	0.04**
Ga	−8.584	2.115	0.98	<0.01***
Pt	−14.21	−2.097	0.02**	0.01**
Pa	−8.025	−0.538	0.30	<0.01***
PISASCIE				
Gt	−2.528	−1.951	0.03**	0.04**
Ga	−9.397	1.393	0.91	<0.01***
Pt	−14.082	−1.975	0.02**	0.01**
Pa	−8.61	−1.095	0.14	<0.01***

Note: Gt, Ga, Pt, Pa – subsequent Westerlund test statistics.

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

As was stated in Section 3, few tools are adequate given the dependence across different variables, however, Westerlund's approach does have this feature. Besides, the Westerlund test, which results are provided in Table 4, accounts for cross-sectional dependency by using bootstrapping to compute robust critical values. Indeed, the issue of cross-dependency has been shown to affect the results of cointegration as can be seen in the comparison between *p*-values and robust *p*-values (the two last columns of Table 4). As a result, we can conclude that the series are cointegrated and that Eq. (9) is correctly specified.

The use of robust *p*-values allows for consistent inference on the existence of co-integrating relations. This would not be the case if standard *p*-values were used or if inference was based on tests that neglect the existence of cross-dependence.

The results provided in Tables 2–4 allow us to conclude that the necessary conditions are met and the main Eq. (9) can be estimated credibly without the risk of being interpreted as spurious regression. This is an important issue: the non-stationarity of the incurred series would otherwise raise serious doubts about the validity of the results. The results of the extrapolation procedure described by Goczek et al. [2021] can also be used to predict the theoretical PISA scores in other countries which have not performed the PISA tests so far upon the assumption of them being typical in terms of their individual country effects. This approach also allows for further analysis to be based on the much broader group of 111 diverse world countries conditional upon the availability of the considered growth factors and the educational expenditures as the latter can be used to estimate the theoretical PISA test results even in the countries where the test itself has not been organized. It should be mentioned though, that all the following results have been verified also in the initial group of countries that have undergone the PISA testing and the results are consistent. To estimate Eq. (9), a total of 1,731 observations in the period 1994–2015 are used. In Eq. (9), we apply the PISA scores lagged by a decade (i.e., we use $s = 10$), which is equivalent to the age at which the graduates enter the labor market. Subsequent columns of Table 4 present the estimates of Eq. (9) with different PISA scores (average, reading, mathematics, and science, respectively).

The theoretical model displayed an output distance between the country's leader and follower. If the follower increases the quality of education, then in a steady state he catches up with the leader and grows at the same pace as the leader. This corresponds to the idea of a cointegrating relationship which is a

Table 5. Panel VECM dynamic fixed-effects estimates

		1	2	3	4
Long-run					
	OPEN	0.0043*** (0.0014)	0.0044*** (0.0014)	0.0041*** (0.0014)	0.0042*** (0.0014)
	FDI (% GDP)	-0.0029 (0.0027)	-0.0029 (0.0028)	-0.0027 (0.0026)	-0.0028 (0.0027)
	L ₁₀ PISAAVE	0.0221*** (0.0085)			
	L ₁₀ PISAREAD		0.0164** (0.0071)		
	L ₁₀ PISAMAT			0.0236*** (0.0084)	
	L ₁₀ PISASCIE				0.0258*** (0.0094)
Short-run					
	Error Correction	-0.0316*** (0.0049)	-0.0312*** (0.0050)	-0.0327*** (0.0049)	-0.0321*** (0.0049)
	ΔOPEN	0.0002** (0.0001)	0.0002** (0.0001)	0.0002** (0.0001)	0.0002** (0.0001)
	ΔFDI (% GDP)	0.0001* (0.0001)	0.0001* (0.0001)	0.0001* (0.0001)	0.0001* (0.0001)
	ΔL ₁₀ PISAAVE(-1)	-0.0007 (0.0007)			
	ΔL ₁₀ PISAREAD(-1)		-0.0008 (0.0005)		
	ΔL ₁₀ PISAMAT(-1)			-0.0001 (0.0009)	
	ΔL ₁₀ PISASCIE(-1)				-0.0008 (0.0007)
	Constant	0.0018 (0.1130)	0.0780 (0.0992)	-0.0184 (0.1136)	-0.0517 (0.1247)

VECM, vector error correction model.

Notes: ΔL₁₀X = 10-year-lagged annual difference in the variable X; (-1) = single year lag; Standard errors in parentheses

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

simultaneous long-run positive feedback loop. For cointegration identification purposes, the lagged GDP value has its coefficient normalized to 1. This relates to the long-run equilibrium (Table 5, upper panel), in which the quality of education, as captured by all classes of PISA scores, has a significantly positive impact on growth. Therefore, this confirms the main hypothesis. Therefore, a significant coefficient in the long-run cointegrating equation means that whenever the steady-state level of GDP in the follower changes due to an increase in education quality, it follows that this positive spillover impacts the leader as well. Therefore, improvement in the quality of education brings increases in the level of GDP, and the follower, as well as the leader, benefit from this improvement. The obvious rationale behind this result is that the flow of technical thought and human capital in an open economy spillover from the leader to follower and back. Success stories of catching up in terms of development require the absorption of technological expertise from abroad in ways that correspond approximately to our theoretical framework.

In contrast, we cannot confirm that the increases in education quality bring a temporary acceleration of economic growth given the result (Table 5, lower panel) that short-run increases in the quality of education do not significantly impact transitory growth to the steady state, though as indicated above, in the steady state, the quality of education has a positive effect on the level of income.

In addition, we find that external openness as captured by trade to GDP is significant and positive both in the short and in the long run. Besides, the FDI variable impacts transitional dynamics in the

short run (openness to FDI increases the speed of attaining the steady state), which is in line with the theoretical model. Besides, to test the robustness of the specification, we have carried out DOLS and FMOLS estimations that provide further evidence in support of the DFE results (omitted due to lack of space). We can interpret the significant variables as an increase in steady-state growth. Therefore, we need both crucial factors to make this growth happen. First, we need country openness in terms of trade, second, we need education quality.

Lastly, having found that the variables are first-difference stationary, an attempt can be made to identify the causal direction of the relationship between them. Information about the exact direction of the causal link enables a more nuanced discussion of the policy implications of the findings – otherwise, it is difficult to clearly state that there exists any “impact” of the quality of education on the GDP growth and formally the only thing that can be confirmed is their co-variation.

Two approaches are performed and described in Table 6. We provide the results of tests under two sets of assumptions: imposing parameter equality across countries (“common coefficients”) and allowing for their differences, as in Dumitrescu and Hurlin (“individual coefficients”).

As could be expected, the results strongly indicate a bi-directional causal link between education quality and economic growth that can be thought of as a positive feedback loop: increases in the quality of education impact the level of development, and the level of development positively increases the quality of education resulting in a positive synergy effect. This is an important result, which on the one hand confirms that the conclusions achieved in the milestone articles in the field that point to the existence of the influence of education quality on GDP growth are indeed correct in the qualitative sense, while on the other hand – suggests that the quantitative results should be taken very carefully: unless the techniques used take care of the bilateral relation of the dependent and the independent variables, the quantitative interpretation of the results may be misleading.

The notion that quality of education is important is prevalent in the educational literature and there is an overwhelming body of knowledge to support that claim. However, the methods used to obtain the results that confirm the relevance of the quality of education for the GDP growth that can be found in the literature are often excessively simple compared to the current state of the art. Most notably, they do not allow answering whether increases in education quality cause increases in economic growth allowing poorer countries to catch up. To answer this question, we propose a formal theoretical model using Nelson and Phelps [1966] classic dynamics by introducing differences in education quality (proxied by students’ performance on the PISA test) in a leader-follower type of growth model to formulate several hypotheses regarding the analyzed education quality, external openness and economic growth nexus.

A panel DFE model has been estimated to examine in both the short run and the long run in our model. The results can be interpreted on a Granger-type causality basis. Education quality as captured by various PISA skills was found to have a positive significant impact on economic growth in the long run while no significant short-run relationship between these variables was observed. External openness was found to

Table 6. Panel VECM Granger causality

Null Hypothesis:	Common coefficients		Individual coefficients		
	F-statistic	p-value	W-Stat.	Zbar-Stat.	p-value
PISAAVE does not cause Log(GDP)	127.347	0.000	9.534	15.982	0.000
LOG(GDP) does not cause PISAAVE	17.455	<0.001	4.692	3.111	0.002
PISAREAD does not cause Log(GDP)	22.825	0.000	1.482	−2.938	0.003
LOG(GDP) does not cause PISAREAD	1.740	0.157	1.704	−2.109	0.035
PISAMAT does not cause Log(GDP)	83.056	0.000	9.422	15.683	0.000
LOG(GDP) does not cause PISAMAT	11.649	<0.001	5.699	5.788	<0.001
PISASCIE does not cause Log(GDP)	214.826	0.000	11.419	20.993	0.000
LOG(GDP) does not cause PISASCIE	23.700	<0.001	7.167	9.690	<0.001

VECM, vector error correction model.

increase economic growth in the short run, while in the long run, only openness to trade was significantly positive.

The results discussed here should be viewed as a formal backing of the frequently expressed notion that expenditures on education, bring about the increased output that may finance the reform. The return is not immediate, however, but coincides with the external openness of countries – the open countries can use their increasing quality of education to catch up with the technological leaders. In this context, the countries that want to provide sustainable economic growth coping with technological leaders should allocate more resources to the quality of education. Importantly, the results are based on the quality – not the quantity of education, which is a confirmation of a key role in the way that education processes are planned, organized, and administrated. While considered in pair with numerous results from the literature that do not confirm the important role of the average educational attainment or the average years of schooling, can also be treated as an important voice in the discussion whether from the economic point of view the governments should concentrate on massive scale tertiary education, perhaps covering its costs for all the interested students or, instead, invest in the quality of the primary and secondary education. There is more evidence for the long-run relevance of the latter.

Having reviewed the field one can find out that most of the authors make use of quite simple research methods, while reverse causality and the omitted variable bias seem to be the biggest threat. Considering the importance of the research problem, it is worthwhile to adopt a more up-to-date research methodology, which would allow for validation of the long-standing argument justifying the value of high-quality education. To the best of our knowledge, no co-integration analysis is present in the literature in the field. The latter seems to be particularly important when time-series (or panel) data from different countries are used. With the use of country-level panel data, we consider the possible lag that characterizes this education-growth nexus and use statistical methods to address the risk of reversed causality of economic performance affecting the quality of education, and we extend the model by inclusion of other potential growth factors. This we believe is the main value-added aspect of this paper. Obviously, attempts to modernize the econometric approach by developing further and further technical innovations should not be viewed as a target in itself. However, as noticed by Altman [2020], most authors neglect the fact that basing the research and drawing conclusions about relations simply basing on the test of significance might be misleading: not only can the regression be spurious but also the “influence” cannot be treated as a simple consequence of the existence of correlation which seems to be the case in most papers that explore the GDP-education nexus, including the early Hanushek and Woessmann’s milestone analysis.

The main limitation of the study is data availability. While there is no problem with the availability of the macroeconomic variables used in the study, the seven waves of PISA test that have been carried out so far (2000, 2003, ..., 2015, 2018) is already a substantial sample, it still requires extrapolation to base the study with lagged PISA results on a reasonable number of observations. This problem is diminishing over time as there is a new wave every three years and the fraction of the sample that is extrapolated – not just collected – is decreasing. This of course also means the possibility to apply more estimators than just the DFEs, which gives a chance to attain more robustness checks and verify the stability of the results to an even bigger extent. On the other hand, there is a natural limitation with the measurement of the quality of education. While it is difficult to point out a better cross-country education quality measurement tool than the PISA scores, there exists a wave of criticism against PISA. The critiques raise different arguments, whose summary can be found in Zhao [2020], however, while criticizing the PISA itself, no superior alternatives are proposed by the opponents of the OECD’s initiative.

6 Conclusions

The results of our formal theoretical and empirical analysis confirm the huge relevance of earlier education quality for GDP growth. Furthermore, we have shown that the changes in education quality shall have effects on the GDP growth only on a longer time horizon. In policy terms, this is an important voice in the discussion on the role of education and provides an additional argument for directing a stream of investment toward

education. This seems particularly important since politicians may find it tempting to limit educational expenditures to assign more resources to other sectors that have suffered during the COVID-19 epidemic. This in turn may have an effect on the quality of education and further – on the rate of economic growth in the future. Thus, the true threat could be associated with the fact that the positive effects of increased education quality are not immediate, while the same is true in reverse – cutting education expenditures may hurt growth only in the long run. Lowering educational expenditures may result in negative selection in teaching professions, lower workforce motivation, or worse studying conditions – all these effects are in the long term. Since politicians are usually focused on immediate electoral success, they might be willing to reduce educational expenditures motivating this policy with a lack of instantly observable negative consequences.

Acknowledgment

This research has been conducted and the results have been obtained within a project financed by the National Science Centre, Poland—number 2015/17/B/HS4/02095.

References

- Acemoglu, D., Johnson, S., Robinson, J. (2001), The colonial origins of comparative development: an empirical investigation, *American Economic Review*, Vol. 91, No. 5, pp. 1369–1401, <https://doi.org/10.1257/aer.91.5.1369>.
- Altman, M.A. (2020), More scientific approach to applied economics: reconstructing statistical, analytical significance, and correlation analysis, *Economic Analysis and Policy*, Vol. 66, pp. 315–324.
- Balart, P., Oosterveen, M., Webbink, D. (2018), Test scores, noncognitive skills and economic growth, *Economics of Education Review*, Vol. 6, pp. 134–153, <https://doi.org/10.1016/j.econedurev.2017.12.004>.
- Barro, R. (1991), Economic growth in a cross-section of countries, *The Quarterly Journal of Economics*, Vol. 106, No. 2, pp. 407–443.
- Barro, R., Lee, J. (2015), *Education matters: global schooling gains from the 19th to the 21st century*, Oxford University Press, Oxford.
- Becker, G. (1962), Investment in human capital: a theoretical analysis, *Journal of Political Economy*, Vol. 70, pp. 9–49.
- Benos, N., Zotou, S. (2014), Education and economic growth: a meta-regression analysis, *World Development*, Vol. 64, pp. 669–689.
- Blackburne, E., Frank, M. (2007), Estimation of nonstationary heterogeneous panels, *Stata Journal*, Vol. 7, No. 2, pp. 197–208.
- Bosworth, B.P., Collins, S.M. (2003), The empirics of growth: an update, *Brookings Papers on Economic Activity*, Vol. 34, No. 2, pp. 113–206.
- Breton, T. (2011), The quality vs. the quantity of schooling: what drives economic growth? *Economics of Education Review*, Vol. 30, No. 4, pp. 765–773, <https://doi.org/10.1016/j.econedurev.2011.01.005>.
- Brouwers, S.A., Van de Vijver, F.J.R., Van Hemert, D. (2009), Variation in Raven's Progressive Matrices scores across time and place, *Learning and Individual Differences*, Vol. 19, pp. 330–338.
- Campbell, S., Üngör, M. (2020), Revisiting human capital and aggregate income differences, *Economic Modelling*, Vol. 91, pp. 43–64, <https://doi.org/10.1016/j.econmod.2020.05.016>.
- Chen, S.S., Luoh, M.C. (2010), Are mathematics and science test scores good indicators of labor-force quality? *Social Indicators Research*, Vol. 96, No. 1, pp. 133–143.
- Cieślak, A., Goczek, Ł. (2018), Control of corruption, international investment, and economic growth – Evidence from panel data, *World Development*, Vol. 103, pp. 323–335.
- Cieślak, A., Goczek, Ł. (2018), Corruption, privatisation and economic growth in post-communist countries, *Europe-Asia Studies*, Vol. 70, No. 8, pp. 1303–1325.
- Cieślak, A., Tarsalewska, M. (2011), External openness and economic growth in developing countries, *Review of Development Economics*, Vol. 15, No. 4, pp. 729–744.
- Dumitrescu, E., Hurlin, C. (2012), Testing for granger non-causality in heterogeneous panels, *Economic Modelling*, Vol. 29, No. 4, pp. 1450–1460.
- Elbakidze, L., Jin, Y. (2015), Are economic development and education improvement associated with participation in transnational terrorism? *Risk Analysis*, Vol. 35, No. 8, pp. 1520–1535, <https://doi.org/10.1111/risa.12378>.
- Erosa A, Fuster L, Kambourov G, Rogerson R., (2022). „Labor Supply and Occupational Choice,” NBER Working Papers 30492, National Bureau of Economic Research, Inc. Flynn, J. (1984). The mean IQ of American: massive gains 1932 to 1978, *Psychological Bulletin*, Vol. 95, pp. 29–51.

- Flynn, J. (1987), Massive IQ gains in 14 nations: what IQ tests really measure, *Psychological Bulletin*, Vol. 101, pp. 171–191.
- Goczek, Ł., Witkowska, E., Witkowski, B. (2021), How does education quality affect economic growth? *Sustainability*, Vol. 13, No. 11, pp. 6437. <https://doi.org/10.3390/su13116437>
- Goldin, C. (2016), Human capital, in: *Handbook of cliometrics*, Springer, Heidelberg, Germany. <https://scholar.harvard.edu/goldin/publications/human-capital>
- Hanushek, E.A., Woessmann, L. (2008), The role of cognitive skills in economic development, *Journal of Economic Literature*, Vol. 46, No. 3, pp. 607–668.
- Hanushek, E., Woessmann, L. (2012), Do better schools lead to more growth? Cognitive skills economic outcomes and causation, *Journal of Economic Growth*, Vol. 17, No. 4, pp. 267–321.
- Hanushek, E.A., Woessmann, L. (2015), *The knowledge capital of nations: education and the economics of growth*, MIT press, Cambridge, MA.
- Heyneman, S.P. (1995), Economics of education: disappointments and potential, *Prospects*, Vol. 25, No. 4, pp. 557–583.
- Im, K.S., Pesaran, M.H., Shin, Y. (2003), Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, Vol. 115, pp. 53–74.
- Jones, B. (2014), The human capital stock: a generalized approach, *American Economic Review* American Economic Association, Vol. 104, No. 11, pp. 3752–3777.
- Kebłowski, P. (2016), Canonical correlation analysis in panel vector error correction model, performance comparison, *Central European Journal of Economic Modelling and Econometrics*, Vol. 8, pp. 203–217.
- Krueger, A., Lindahl, M. (2001), Education for growth: why and for whom? *Journal of Economic Literature*, Vol. 39, No. 4, pp. 1101–1136.
- Manuelli, R., Seshadri, A. (2014), Human capital and the wealth of nations, *American Economic Review*, Vol. 104, No. 9, pp. 2736–2762.
- Nelson, R., Phelps, E.S. (1966), Investment in humans, technological diffusion, and economic growth, *American Economic Review*, Vol. 56, No. 1/2, pp. 69–75.
- OECD. (2010), The high cost of low educational performance, OECD Publishing, retrieved from www.sourceoecd.org/education/9789264077485
- Pesaran, M.H. (2004), General diagnostic tests for cross-section dependence in panels. Cambridge working papers in economics no. 435. University of Cambridge and CESifo working paper series no. 1229.
- Pesaran, M.H. (2007), A simple panel unit root test in the presence of cross-section dependence, *Journal of Applied Econometrics*, Vol. 22, pp. 265–312.
- Pesaran, M.H., Shin, Y., Smith, R. (1999), Pooled mean group estimation of dynamic heterogeneous panels, *Journal of the American Statistical Association*, Vol. 94, pp. 621–634.
- Schoellman, T. (2012), Education quality and development accounting, *Review of Economic Studies*, Vol. 79, No. 1, pp. 388–417.
- Schultz, T.W. (1961), Investment in human capital, *The American Economic Review*, Vol. 51, No. 1, pp. 1–17.
- Stern, J. (2000), Pakistan's Jihad culture, *Foreign Affairs*, Vol. 79, No. 6, pp. 115–126.
- Tamura, R. (2001), Teachers growth and convergence, *Journal of Political Economy*, Vol. 109, No. 5, pp. 1021–1059.
- Westerlund, J. (2007), Testing for error correction in panel data, *Oxford Bulletin of Economics and Statistics*, Vol. 69, pp. 709–748.
- Zhao, Y. (2020), Two decades of havoc: a synthesis of criticism against PISA, *Journal of Educational Change*, Vol. 21, pp. 245–266.