

Article

How Does Education Quality Affect Economic Growth?

Łukasz Goczek ¹, Ewa Witkowska ² and Bartosz Witkowski ^{3,*}

¹ Faculty of Economic Sciences, University of Warsaw, 00-241 Warsaw, Poland; lgoczek@wne.uw.edu.pl

² Institute of Psychology, The Maria Grzegorzewska University, 02-353 Warsaw, Poland; ewitkowska@aps.edu.pl

³ Institute of Econometrics, SGH Warsaw School of Economics, 02-554 Warsaw, Poland

* Correspondence: bwitko@sgh.waw.pl; Tel.: +48-22-564-9256

Abstract: In a seminal article, Hanushek and Woessmann explained economic growth as a function of the quality of education. While they did not find evidence of the importance of years of schooling, they argued for the relevance of cognitive skills and a basic literacy ratio for economic growth. However, this result was based on cross-country data limited to 23 observations. In this study, we extended and modified their approach based on the results of PISA (Programme for International Student Assessment) tests to explain the GDP changes over the last 50 years. Using panel data, we considered the possible lag that characterizes this relationship, used statistical methods to address the risk of reversed causality of economic performance affecting the quality of education, and extended the model by the inclusion of other potential growth factors. The results, which also included several robustness checks, confirmed the relevance of earlier education quality as a significant growth factor. Our results suggest the significance of educational skills for GDP growth, which might be treated as a confirmation of the importance of quality primary and secondary education for economic development. We showed that our results are robust to changes in the order of lags and confirmed the validity of the conclusion with the use of specification-robust Bayesian model averaging.

Keywords: economics of education; educational policy; international education/studies; economic growth; PISA test



Citation: Goczek, Ł.; Witkowska, E.; Witkowski, B. How Does Education Quality Affect Economic Growth? *Sustainability* **2021**, *13*, 6437. <https://doi.org/10.3390/su13116437>

Academic Editor: Franklin G. Mixon

Received: 30 April 2021

Accepted: 2 June 2021

Published: 5 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Education is one of the most important social institutions. It is faced with difficult tasks and is widely regarded as one of the main development mechanisms. Therefore, the relation between the quality of education and economic performance is always of topmost importance. The conviction that education has an important impact on economic growth and sustainable development is commonly shared [1–4]. In this spirit, the endogenous growth model emphasizes the role of human capital in shaping and understanding the economic growth of a country or region within a country [5]. Hess [3] argues that the relationship between human capital development and economic growth is a key factor for a well-functioning healthy economy. Although there is much discussion of its importance, politicians, however, in numerous countries and at different times, seem to treat education as a child of a lesser God whenever the education funding is decided upon. The aim of this study was thus to verify to what extent the quality of education can be considered a key driver of economic growth from the cross-country perspective.

The existence of several cross-country rankings of universities as well as the popularity of studies and internships abroad at the university level seem to confirm that not just experts but even the wider audience understand and appreciate the significance of the quality of tertiary education. At the same time, there seems to be less certainty about the importance of the quality of the primary and secondary education, on which we focus in this article. In this field, the research of Hanushek and Woessmann [6] has raised significant interest. They confirmed the intuitive importance of high-quality education and provided clear

policy implications: the better the quality of education is, the higher the expected economic growth will be. Given that numerous previous research confirms that educational results are mostly associated with the share of educational expenses in GDP [7,8], this stimulates governments to increase investment in the educational sector. Not just because people expect better education for their children but also because there exist confirmed returns, economies with better education quality grow faster. One can consider Hanushek and Woessmann's research as an important vote in the discussion: not only is it the quantity of education (measured by financial loadings or average years of schooling) but also—or even mostly—it is the quality that matters for economic growth.

Hanushek and Woessmann [6] used average GDP growth data from the years 1960–2010 to estimate their model. They then estimated a linearized regression model of GDP growth as a function of PISA results, which were supposed to reflect the quality of education. While their results were in line with expectations, certain methodological problems arose. Therefore, we extended and modified their approach based on the results of PISA tests to explain the GDP changes over the last 50 years. An important problem that needed to be faced by Hanushek and Woessmann was the lack of a sufficient number of observations: the PISA tests that they used as the measure of education quality had a relatively short history and were performed in relatively few countries. Since then, there have been further waves of PISA tests and more countries have performed them. However, we further extended the sample by using additional sets of educational and macroeconomic data in order to extrapolate backward the results of PISA scores, which was the first step of the research. In the second step, we used lagged PISA scores as regressors in the augmented Solow growth model with the quality of education factors. The methods used in this article were aimed at eliminating the six main potential critiques of Hanushek and Woessmann's [1] results, in order to reconfirm and provide a more reliant empirical basis of the policy conclusions, and furthermore, to extend and modify their approach:

1. A small sample. The original article was based on data covering 23 countries that performed PISA tests. We substantially extended this dataset in both time (using updated data) and geographic dimensions and based the model on the panel instead of a cross-sectional sample.
2. The lag length. PISA tests are used to measure the skills of large groups of teenagers aged 14 or 15. However, the main transmission mechanism between the quality of education and economic growth is based on the influence of education quality on labor skills, which further influences the state of the economy. It is thus clear that the human capital quality of 15-year-olds shall not influence the GDP growth sooner than in at least a few years when they start their first job—and it is quite likely that most of the relationship will only be observable after some further few years. We thus properly lagged PISA scores used in the function of the measure of education quality.
3. The dynamics of the model. The dynamic nature of growth and the use of cross-country panel data made it essential to apply such modern estimation tools as the Blundell and Bond system GMM estimator [9], rather than the simple OLS estimator used by most. This was mostly due to the existence of GDP convergence across countries, which should be considered (or at least tested).
4. Omitted variable bias. The quality of education is only one of the numerous growth factors (and relevant independent variables in the case of the other economic indicators of interest). Failing to include other growth factors results in an increased risk of omitted variable bias. To avoid this problem, we proposed a complex Barro-type regression including not only the PISA scores but also other typical growth factors.
5. The risk of inverse causality and endogeneity issues. The economic variables used in the model were expected to depend on the education quality; however, the relation in the opposite direction also exists, making the education quality endogenous. This is a problem in most regression model frameworks, even considering the fact that the educational results shall be appropriately lagged in the constructed model. This, however, can be tackled by proper treatment of the variables, which is possible in the

- Blundell and Bond method of estimation [9], which we used. Additionally, the use of this particular estimator made it automatic to check for causality and its direction.
- Types of skills. Most research emphasizes the possible role of cognitive skills, whereas it is not only a sensible influence on the future socio/macroeconomic indicators on the national level. PISA test results allow the evaluation of competencies acquired by students of different types (mathematical literacy, reading literacy, and scientific literacy). Whereas most studies concentrate on the acquired skills in the field of science, we separately analyzed different types of competence and their impact on economic growth. This part of the study may constitute a contribution to the adequate profiling of educational expenses.

The article is structured as follows. In Section 2, we survey the relevant literature and provide the background of the topic. In Section 3, we present a proper theoretical model that provides a theoretical background for the role of education quality for the GDP growth and the empirical model as well as the econometric approach used in the study. Section 4 presents the main empirical findings, while Section 5 concludes. All the computations were performed with the use of Stata 16 software (StataCorp LLC, College Station, TX, USA).

2. Literature Background

Hanushek and Woessmann [1] emphasized that although the role of human capital in growth is regarded as obvious in theoretical discussions, the results of the empirical analysis are unclear. According to them, this mixed evidence appears to reflect measurement issues and the skills–growth relationship becomes unequivocal when the quality of schools and the varied sources of skills are taken into account. Nevertheless, as stated by Morris and Oldroyd [10], the quality of human resources impacts economic growth by increasing employees, income, and welfare; therefore, human capital is to be seen as a key element accelerating economic growth. Hanushek and Woessmann [1] argued that economic growth is strongly affected by the knowledge capital of workers in both developed and developing countries alike. In today’s rapidly developing knowledge economy, complex skills and competencies of employees are of special importance for economic growth [11]. In their study using Indonesian annual data for 35 years, Widarni and Bawono [2] examined the long-term relationship between economic growth and its determinants, and they also ascertained the impact of human resources and technology on economic growth both in the long and short term. The authors concluded that human capital is effective in promoting long-run economic growth, whereas effective technology drives economic growth in the long and short run [2].

Education plays an important role in accelerating the process of passage from a pre-scientific to a scientific mode of thinking. This change most probably accounts for the Flynn effect [12]—the worldwide improvement of populations’ intelligence test results. An effective shift from the manner of thinking in which the learner is given the principles that shall be used in the problem-solving thinking to the scientific mode of thinking based on the use of the higher levels of general and immediate resolution, results in detachment from the concrete and becoming more flexible and results in more efficiency in the way of complex problem solving, which seems crucial for today’s knowledge-based economies. In the knowledge-based environment with rapidly advancing technological development, there is a continuous need for the life-long learning of employees. Companies have to manage employees in such a way as to improve workability so that they can achieve higher levels of performance [13], but previously developed cognitive competencies form a solid base for this process.

On the microeconomic level, the accumulation of human capital improves labor productivity and increases wages [14]. A well-educated workforce is also essential for the creation and diffusion of technology. There is plenty of microeconomic evidence on the robust relationship between education and earnings, stemming from the widely known Mincer equation [15], which uses years of schooling as a key determinant of wages, and several growth-accounting studies, which adjust the workforce for improvements in

educational attainment (operationalization of the human capital factor). This seems as natural as the statement by Castriota that pointed out that empirical research has found a positive impact of education on the level of happiness [16].

The argument for macroeconomic effects relies on this microeconomic evidence. Improvement in education could have an aggregate impact on economic growth through two different channels. On the one hand, extended education may improve the productivity of the workforce as it improves workers' individual skills. On the other hand, as adopted [17,18], education can be treated as an independent factor in the growth process, which can augment labor, physical capital, and total factor productivity (TFP). The relationship with TFP reflects the view that an educated workforce is more able to implement new technologies and generate ideas for improving efficiency. Both considered mechanisms provide justifications for the expected positive correlation between educational attainment and economic growth, which stems from the aggregation of the effects on an individual level. A number of studies, including those of Mankiw et al. [17,18], have found a significant positive association between cross-national differences in the initial endowment level of education and subsequent rates of growth.

However, numerous studies that examined the relationship between the dynamics of the educational factor (measured by the years of schooling) and the dynamics of GDP failed to find a significant association [19]. Benos and Zotou [20] described, in a recent comprehensive meta-regression analysis of 57 studies with 989 estimates, the presence of a substantial upward publication selection bias in the empirical education–economic growth literature and the absence of a representative authentic growth effect of education. The addition of openness, public spending, and health variables in specification involve a lower estimated education impact on growth. The authors concluded, “most importantly, educational attainment, commonly used in empirical studies, is a crude measure of human capital, since it measures education quantity, while education quality varies widely across countries and time periods.”

There are three potential reasons for the failure to replicate the microeconomic results on the aggregate level. Firstly, the global returns to schooling that can be observed in the aggregated data may be much lower than the individual returns, which can be observed in the microdata, mostly because of the averaging processes, which can be viewed as an extended version of the statistical Simpson's paradox. Secondly, there is a high risk of measurement errors, which, after aggregation, are an even greater problem and are more difficult to discover and eliminate. Thirdly, cross-national differences in years of schooling and their dynamics or other schooling indices (such as the fraction of society with tertiary education) may not account for differences in the quality of education. Thus, educational attainment, commonly used in empirical studies, is a crude measure of human capital, since it measures education quantity, whereas education quality differs extensively across countries and time [20], which is not taken into consideration if the above-described measures are used. Therefore, as argued by Hanushek and Woessmann [6], adding the quality of education may alleviate these concerns. Schoellman [21] adjusts years of schooling for quality to explain cross-country output per worker differences. He finds that cross-country differences in education quality are roughly as important as cross-country differences in quantity. This increases the total contribution of education from 10% to 20% of output per worker differences.

Hanushek and Woessmann [1] have apprised that in developing countries the discussions on development policy often simplifies and distorts this truth by focusing most attention on ensuring school enrollment for everyone and losing sight of the importance of the quality of education. In their latest work [1], they reported findings indicating that knowledge, rather than just the length of time spent in school, is the factor accounting for economic growth. In international comparisons of a broad group of countries, the average number of years of schooling [7] and school attainment [22] were treated as a good proxy for the quality of education [7]. This measure, however, is biased as it does not embrace the differences in the quality of education across countries over time.

A powerful tool that allows for the quantification of the educational attainment quality is the use of PISA tests. Education systems might be assessed in terms of context, specific inputs, social or institutional processes, and outputs or outcomes [23]. For each of these categories, specific indicators can be designed. Context indicators focus on providing information on such contextual factors that affect learning as student characteristics, socio-economic conditions, cultural aspects, the status of the teaching profession, and local community issues. Input indicators measure the deployment and use of resources to facilitate learning, such as financial, material, and human resources in the education system. Process indicators describe to which extent specific educational processes are conducted in practice, and output indicators measure whether the program objectives were attained by revealing the level of competencies, progression and completion rates, and employer satisfaction. Output indicators may be obtained through national examinations, international assessments such as PISA or TIMSS (Trends in International Mathematics and Science Study) surveys, and systematic field observations belonging to this category [23].

PISA is a worldwide study created in 1997 by the Organization for Economic Cooperation and Development (OECD) and intended to develop tools allowing for the evaluation of the quality of education in member and non-member countries [1]. The governments of the member states have committed to regularly monitor the effectiveness of their education within the international framework developed by the Decomposition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo) Committee established by OECD. The main objective was to create a new basis for political dialogue and cooperation in defining and implementing, in an innovative way, educational goals aimed at developing key skills in adulthood [1]. Starting from the year 2000, every three years, the PISA Programme assesses the competencies of students who have reached the age of fifteen in three areas considered key for modern life and today's job market: reading, mathematics, and science.

The specificity of the PISA is based not only on its international momentum but also on measuring literacy in the three assessed areas of literacy. It is believed to measure the assimilation of knowledge and the mastery of skills necessary for students in their adult life, on the labor market, and for their fully free functioning in society. The PISA research agenda is also unique in that it broke with national curriculum-based international research traditions, introducing a problematic approach detached from the core curriculum, providing the results and knowledge that allow better planning of effective teaching. The PISA framework was carefully prepared in the process of long-term consultations of international experts and solicitously documented in separate publications [24].

Thanks to the fact that the PISA test is repeated every three years, it is possible to capture the dynamics of changes in the education system. In each edition of PISA, one of the areas (reading, mathematics, and science) Reading is the leading subject matter. The subject considered to be the leader at a given moment of the study takes half the time devoted to solving the test skills testing tasks, and the remaining two take 25% each. The student has two intervals of 60 min, with a break of 10 min, for the entire test, with the tasks being mixed up and the student being unaware of the domain a specific task is assigned to [25].

The basic units for the PISA test are countries. It was assumed that within each country, students will be randomly selected in schools—between 5000 and 10,000 students are selected from at least 150 schools per country [26]. However, actual empirical samples range widely from around 3500 to over 30,000 students in some countries. The sampling procedure thus limits the possibility of comparing the results to comparisons between countries and at the level of individual countries, and it is not possible to analyze the level of individual school classes [27]. In the selection of the sample of students in the PISA study, two-stage stratified sampling with systematic randomization is used. The first stage of sampling is the drawing of schools, and the second is the drawing of students from previously selected schools. To be able to generate unbiased estimators of the national parameters, the weighting for the individual trials is determined; replicated weights are used, which allows the estimation of unbiased standard errors of measurements for all

national parameters [28]. The probability of a particular school being selected is proportional to the number of its students eligible for the PISA survey. Thanks to the use of the stratified sampling procedure, it is possible to improve the accuracy of estimates by partially controlling the inter-school variance [25].

As a result of implementing the lifelong learning (LLL) perspective, the tasks in PISA tests assessing life-relevant competencies are not related to national curricula. Therefore, one consistent theoretical rationale allowing a combination of the competencies of adolescents and adults is to make the program foundation reflect human development and to establish a framework for the development of adequate tools to measure it [25]. The concept of lifelong learning is based, on the one hand, on the assumption through the psychological mechanism of nonspecific transfer [29] that, when relying on competencies acquired during school education, a person will be able to implement them and, as a result, cope in adult life, whereas, on the other hand, neither school nor studies will be able to equip us with all the specific competencies needed in life.

Sellar and Lingard [30] state that the key to the increase in the importance of education for the OECD is the growing importance of PISA tests' results and the OECD's impact on global-scale education along with the implementation of the human capital theory in which the simultaneous process of the "economization" of education policy and the "education" of economic policy is inherent. Currently, the OECD conceptualizes education, skills, and competencies as essential in a world of knowledge-based economic policies. The necessity to equip today's students with lifelong learning tools results from the demographic trends. The worldwide decline in fertility rates along with increasing life expectancy leads to population aging. Therefore, economic growth and stability depend on workers' ability to be present in the labor market and maintain high productivity over an extended time. As the proportion of young working people in societies declines in the coming years, it will be increasingly important for education systems to remove barriers that prevent some of today's students from reaching their full potential in the future. Never before have equality of opportunities and economic efficiency been so closely related [31].

Authors in [6,32] have argued that the cross-country comparative approach provides several advantages over national studies. In these studies, differences in skills measured by international tests such as PISA are believed to be strongly related to individual labor-market outcomes and, consequently, to cross-country variance in the economic growth of particular countries [33–36]. The existence of such relationships can be attributed to two, possibly simultaneous processes: since the results of PISA tests measure the abilities of the young generation, this can, on the one hand, be treated as a proxy for the quality of the future labor force and, as such, may influence the future performance of the economies, measured by the GDP dynamics. However, on the other hand, the results of the PISA tests themselves might be a function of economic development [37]. Most of the surveyed studies used contemporary PISA results to explain the GDP growth over a period of 50 years. While the country's economic performance and the quality of education are most likely associated, a natural question regards the direction of this relationship. Either it might be one of the above-described processes that yields the observed relation, or it might be both of them. Since the PISA results are generally attained for 15-year-olds, it seems rational to believe that the members of today's PISA cohort shall begin to have a greater impact on the country's economy at least 10 years later, after they graduate or reach higher positions in companies. Consequently, the PISA results as a proxy for quality of education might be believed to perform as a relevant GDP driver only if they are appropriately lagged. On the other hand, the relation between the current country's economic performance and PISA results with no lags most likely reflects the reversed relation: they are the better performing economies that generally invest more in high-quality education, which in a short time horizon may result in higher PISA scores. We thus believe that Hanushek and Woessmann's [6] results should not be used to prove that the quality of education matters for GDP growth: they do confirm the existence of a relation, however, the lack of the lag mechanism in the proposed model suggests that they rather confirmed the existence of

an inversed relation and that they did not shed sufficient light on the influence of high quality education on economic performance. This allowed us to argue that the problem of endogeneity should be addressed.

While some research on the influence of educational quality measured by PISA scores on economic growth exists, most of the applied work was based on cross-sectional regressions or correlations. However, a few more methodologically advanced studies, mostly by Hanushek and Woessman, should also be mentioned. In their study [38], they address a problem of causality by trying to gradually eliminate the influence of potentially relevant features on growth to eliminate the potential spuriousness of the regression. More importantly, they emphasized that the education-quality–GDP growth relation is not contemporaneous, and they used properly lagged education measures, which enforced them to use the results of tests performed before the PISA was introduced (such as TIMSS). While [6] mentioned the problem of causality and attempted to perform a series of robustness checks to discover the direction of the existing relation, others have not always done so. For example, authors of [34] simply regressed growth on test scores. The latter researchers investigated the link between test scores (mathematics and science) and cross-country income differences. The authors posed the question of whether test scores are good indicators of workforce quality. The obtained results suggested that, while other variables that are typical in cross-country economic growth studies are properly controlled for, the strong link between test scores and cross-country income differences cannot be confirmed. However, they discovered 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 is possible that either the skills as they are measured in PISA tests are not key factors to the future human capital quality or that it is the quality of tertiary education that has the greatest importance.

Summarizing the results of the study by Cheung and Chan [33], who considered the relation between the PISA scores in mathematics and reading, the gender structure of agriculture, service and industry employment, the number of R&D-involved researchers, and GDP per capita, have asserted that, in the analyzed group of 32 countries, PISA scores are significantly related to employment in different economic activities across countries. At the same time, the PISA science and reading scores positively predicted both male and female employment in the industrial and service sector, whereas the PISA mathematics score positively predicted the number of researchers in the field of research and development but also predicted female and male employment in the agricultural sector negatively. The authors concluded that better academic performance might be crucial for countries to stay competitive in a knowledge-based global economic environment. Yet again, this research was based on a simple regression and correlation analysis, disregarding the perils of endogeneity, country-specific effects, causality, or spurious regression. It thus seems risky to conclude the direction of the discovered relations.

The second issue related to [6] is that the authors did not attempt to construct a more theoretically based GDP growth model, such as the augmented Solow growth model. Instead, they estimated a regression in which the results of PISA tests (interpreted as a proxy for education quality) were almost the only regressor. The risk behind such a procedure is that the quality of education might also be correlated with other GDP-relevant growth factors, such as the total government expenditures as a percentage of GDP. Their omission results in omitted variable bias and, consequently, inconsistency in the applied estimators. One can interpret the estimates of parameters by the PISA scores presented by Hanushek and Woessmann as reflecting the influence of the quality of education and everything else which is correlated with the quality of education on economic growth.

In Hanushek and Woessmann's study [36], the authors developed a new metric for the distribution of educational achievement across countries that allows for tracking

the cognitive skill distribution within countries and over time. Cross-country growth regressions generated a close and stable relationship between educational achievement and GDP growth. In a series of approaches that address the issue of causality, they limited the range of plausible interpretations of this strong cognitive-skills–growth relationship. The results indicated that school policy can be an important instrument to spur growth. The shares of basic literates and high performers have independent relationships with growth, the latter being larger in less-developed countries. In their study, they used cross-national data and applied the two stage least squares (2SLS) estimator while addressing the potential endogeneity of educational achievements, which reflects the possible two-sided relationship and thus the endogeneity of this variable. The authors of [6] addressed the problem of causality in yet one more way, namely by using the difference instead of the level relation. While the latter approaches are a huge step forward, the failure to use panel data imposes a limited number of observations, which is crucial, especially in the 2SLS framework given that, while the essential assumptions are fulfilled, the 2SLS (and the more-developed GMM) estimators are consistent but, in general, not unbiased in a finite sample. Still, most research is based just on cross-sectional data with only a few researchers applying instrumental variables regression.

We found no trace of dynamic panel data models analysis in the literature that makes use of the PISA and similar test measurements of human capital, although the application of some form of GMM dynamic panel model would seem natural. The worries about the data-generating process, which relates education and GDP growth and its potential regressors, can be summarized in the following list:

1. The process is dynamic, with current realizations of the dependent variable influenced by past ones. Particularly, while the GDP growth processes are analyzed, the economically sound existence of beta-convergence [39] requires the use of the GMM approach if the analysis is based on a panel of countries [40,41]. The latter seems essential because of the otherwise very small sample size if only a time series or a cross-section of countries are used.
2. Furthermore, even in a larger group of countries, there may be arbitrarily distributed fixed individual effects in the dynamic so that the dependent variable consistently changes faster for some observational units than others. This argues against cross-section regressions, which must essentially assume fixed effects away, and in favor of a panel set-up, where variation over time can be used to identify parameters.
3. Some regressors (the macroeconomic GDP growth factors in particular) may be endogenous.
4. Measurement errors can be substantial.
5. The idiosyncratic disturbances (apart from the fixed effects) may have individual-specific patterns of heteroskedasticity and serial correlation.
6. Good instruments may not be available outside the immediate dataset.

One prominent way to address these problems is the use of a first-differenced generalized method of moment estimators applied to dynamic panel data models. A commonly employed estimation procedure to estimate the parameters in such a model with unobserved individual-specific heterogeneity is to transform it into first differences. Sequential moment conditions are then used where lagged levels of the variables are instruments for the endogenous differences and the parameters estimated by GMM. The relevant estimator was originally developed in [30] and further in [14], where the so-called system GMM estimator used in this study is discussed. The potential for obtaining consistent parameter estimates even in the presence of the above problems is a considerable strength of the GMM approach in the context of empirical growth research.

3. Materials and Methods

In order to understand the underlying process, consider a simplified version of the Lucas model [42] as in [43], with discrete-time and successive generations of individuals whose lives are divided into two periods. The model is described in Section 3.1. In Section 3.2, we provide the details of the modeling strategy used in the study.

3.1. Theoretical Model

In her first period of life, an agent chooses how to share her time between production and human capital accumulation, which is education. Let us denote the fraction of her time currently allocated to production, then $(1 - u)$ is the current schooling time and δ is the quality of schooling. Therefore, human capital accumulates according to:

$$H_2 - H_1 = \delta(1 - u)H_1, \quad (1)$$

where H_1 (and H_2) is the agent's stock of human capital in period 1 (and period 2, respectively). Let the per capita production in any of the two periods y_t , where $t = 1, 2$, be a function of human capital H_t :

$$y_t = k_t^\alpha (uH_t)^{1-\alpha}, \quad (2)$$

where k_t denotes the physical capital stock in period t ($t = 1, 2$), which evolves over time according to the differential equation as in the Solow or Ramsey model and equals $\dot{k}_t = y_t - c_t$ with c_t representing current consumption in period t ($t = 1, 2$) and α , the parameter of the model, is the elasticity of growth with respect to capital.

Let u^* maximize the agent's intertemporal constant relative risk aversion utility (as in [44]), which spells out how the fraction of time devoted to schooling in the first period $(1 - u)$ affects the accumulation of human capital subject to Equation (1), assuming all the other parameters are known:

$$\max_u \left\{ \frac{(H_1 u)^{1-\sigma}}{1-\sigma} + \beta \frac{(H_2)^{1-\sigma}}{1-\sigma} \right\}, \quad (3)$$

where σ is the risk aversion coefficient that represents the agent's willingness to smooth consumption over time, and β is the subjective discount factor that can be interpreted as the rate of time preference (clearly, $\beta < 1$), which is inversely linked with the agent's rate of time preference r , namely $\beta = \frac{1}{1+r}$. The first-order condition for this maximization is $u^{-\sigma} - \beta\delta[1 + \delta(1 - u)]^{-\sigma} = 0$, which yields

$$u^* = \frac{1}{\delta + (\beta\delta)^{\frac{1}{\sigma}}} \quad (4)$$

Thus, the optimal level of time dedicated to allocated to production is decreasing in δ and β . As a result, the equilibrium economic growth $g = \delta(1 - u^*)$ increases in the quality of schooling δ and decreases in the rate of time preference. This clearly confirms the positive role of education quality as a positive stimulus of economic growth.

3.2. Modeling Strategy

The main concern associated with the use of PISA results as a measure of the quality of schooling δ and the growth factor in empirical growth models is that the PISA tests have only been performed over approximately the last 20 years, and, during most of this period, the number of countries that participated in the tests was anything but large. Assuming that using the results of the test as a growth factor requires lagging them by at least 10 years and given that PISA tests are performed every 3 years, the number of available true observations is limited to, at most, just five per participating country. Such a low number of observations makes it difficult to estimate an appropriate growth model. Its statistical properties would be poor, especially considering the necessity to use the initial two waves of observations as instruments in the GMM estimation procedure, which further shrinks the dataset. In consequence, even interpolating the PISA scores to attain a series of annual observations does not solve the problem of a low number of available observations, which is also due to the low number of countries that participate in the PISA tests: while the group is relatively numerous right now, it has not been throughout the PISA's history. In order to overcome this problem, we applied the following strategy.

Firstly, we used exponential smoothing to interpolate the PISA scores in the years when the tests were not performed in a given country. Secondly, we used backward-extrapolated PISA results series for each of the countries in the sample. While numerous methods of extrapolation are available, we based our rationale on previous research, which suggests that the quality of education is mostly associated with government investment in the educational sector [7,22,45,46]; although, one cannot place an equality sign between the quality and the quantity in this respect. In consequence, as the first step, we used the panel data on the group of countries that participated in PISA tests in order to estimate:

$$PISA_{it}^s = \beta_0 + c_ef_i + \beta_1 GEE_{i,t-p} + \varepsilon_{it}, \quad (5)$$

where $PISA_{it}^s$ is the general score in the skill s measured in the PISA test (mathematics, science, or reading) in country i and year t ; c_ef_i is the country effect for country i , $GEE_{i,t-p}$ are the government expenditures on education as a % of GDP in country i lagged by p years β_0 and β_1 are the parameters of the model, and ε_{it} is the error term for country i in year t .

In the model (5), we included the government educational expenses as the main determinant of the result of the PISA test measuring the given type of skills, and the separate equations were estimated for each type of skills that included mathematical, scientific, and reading skills. However, the past government expenditures, rather than current expenditures, had the greatest influence on the quality of education today. Nevertheless, it remains unclear how much time passes before the results of educational investment can be seen. In modeling terms, how much should the government expenditures be lagged in the model to reflect the effect of additional investments in the skills of students? We thus applied a mixed solution and included three lags of government educational expenditures: 3-year-, 6-year-, and 9-year-lagged expenditures in the model to capture the potential middle, middle-long, and long time horizon, which converts the structure of the model into:

$$PISA_{it}^s = \beta_0 + c_ef_i + \beta_1 GEE_{i,t-3} + \beta_2 GEE_{i,t-6} + \beta_3 GEE_{i,t-9} + \varepsilon_{it} \quad (6)$$

The individual effect of a given country was included so as to reflect each country's unobserved characteristics in the result of which similar spending might not yield the same results of the test scores because of a different starting point, which was due to, among others, diverse cultural and physical capital in the sector of education as well as dissimilar schedules of the educational system across countries.

Finally, we allowed the error term to be non-spherical. This reflected the belief that because of the quite strong stability of PISA results over time, the autocorrelation of the error term is quite likely. In addition, in the case of less-developed countries with lower expenditures on education, the variance of the PISA results is usually greater, which suggests heteroscedasticity of the error term. In the assumptions of the model, we allowed for both.

The panel-data-based model (6) was estimated with the use of the GLS estimator that takes account of autocorrelation and heteroscedasticity and was further used to extrapolate the results of the PISA test backwards. These can be viewed as a form of backward-prediction: given the estimated country effect and the known government educational expenditures, those can be substituted into the model to answer what results of the PISA test would have been likely in the past if those tests had been performed at that time. Naturally, these results, which are the conditional expected values of PISA tests (should those have been performed) can be interpreted as a proxy for the quality of education given the country's real educational expenditures in the past and its estimated individual characteristics. Furthermore, we predicted the theoretical PISA scores for the countries that did not participate in the tests, in the same way, assuming that the individual effects in their case co-respond to the average c_ef_i for the countries that were included in the sample used to estimate model (6). While it is truly the dynamics and not the level of PISA score that matters in the further discussed growth model, the procedure of extrapolating PISA

scores outside the group of countries that have ever participated in the test might raise certain doubts. Thus, we also provided the results restricted to the sample of countries used to estimate model (6) as a part of a robustness analysis.

Using the theoretical results of PISA tests in the past years solves the problem of unavailability of the potential GDP growth factor, which is the proxy for the quality of education lagged by at least a decade. As a result, the remaining step was the estimation of an appropriate GDP growth model. We based the analysis on the model proposed in Section 3 of this article while considering the empirics. We drew from the vast literature devoted to the GDP growth modeling and used the augmented Solow model, which is operationalized in the form of the Barro regression [39]. Importantly, and contrary to Hanushek and Woessmann, we constructed 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. Those involve the human capital factors (life expectancy, population change) and physical capital factors (level of investment, government expenditures, the openness of the economy), while the predicted backward PISA results (for each of the skills—mathematical, reading, and science) were included as an additional human capital factor. The typical Barro regression, which assumes (or at least allows for) the existence of beta-type convergence of the GDP between countries, adapted to the panel data environment, can be written as:

$$\Delta GDP_{it}/GDP_{i,t-1} = \gamma \ln GDP_{i,t-1} + x'_{it}\theta + u_i + \mu_{it}, \quad (7)$$

where GDP_{it} is the GDP of country i in year t in real prices, x'_{it} is the vector of GDP growth factors of country i in year t , including the appropriately backward predictions of PISA scores for the different skills, u_i is the country's individual effect, μ_{it} is the error term for country i in year t , and γ and θ are the parameters of the model.

Given the typically low short-term changes of the GDP and the relative change of GDP in the country i and year t , the $\Delta GDP_{it}/GDP_{i,t-1}$, for computational reasons, was approximated with:

$$\Delta GDP_{it}/GDP_{i,t-1} = \ln GDP_{it} - \ln GDP_{i,t-1} \quad (8)$$

yielding:

$$\Delta \ln GDP_{it} = \gamma \ln GDP_{i,t-1} + x'_{it}\theta + u_i + \mu_{it} \quad (9)$$

Obviously, (8) is equivalent to the estimable form:

$$\ln GDP_{it} = (\gamma + 1)\ln GDP_{i,t-1} + x'_{it}\theta + u_i + \mu_{it} \quad (10)$$

which allows for solving the endogeneity problem that arises in (8).

Model (9) has been discussed and estimated frequently in the economic growth literature. We followed the mainstream literature and applied the common Blundell and Bond estimator [9], and it is worth noticing that the number of considered cross-sectional units (countries) in our case was indeed large enough to apply this GMM-based technique. One important implication is that proper instrumentalization allowed for treating the considered growth factors as potentially endogenous, which reflected their possible immediate two-way relation with the GDP growth. An exception to that were the lagged PISA scores, which surely do not react immediately to the GDP dynamics. The attained estimates allowed for drawing conclusions regarding the relevance of the particular PISA test results, measuring the impact of various aspects of education quality on growth.

4. Results

The PISA score model (6) can be estimated with the use of 638 observations, and we were restricted by the availability of true or interpolated PISA scores for the particular types of skills in this respect and the macroeconomic statistics regarding the adequately lagged government expenditures on education (the data provided in [7] are mostly used for that purpose). Table 1 provides a complete list of country–year observations used to estimate Equation (6).

Table 1. Country–year observations used to estimate the PISA-score Equation (6).

Country	Earliest/Latest Available Observation	<i>n</i>	Country	Earliest/Latest Available Observation	<i>n</i>
Albania	2003–2015	11	Latvia	2005–2015	9
Argentina	2000–2015	14	Lithuania	2006–2015	8
Australia	2000–2015	14	Luxembourg	2000–2004	5
Austria	2000–2015	14	Macao	2007–2015	7
Azerbaijan	2006–2009	4	Malaysia	2009–2012	4
Belgium	2000–2012	13	Malta	2009–2015	2
Chile	2000–2015	14	Mexico	2000–2012	13
Colombia	2006–2015	8	Moldova	2009–2015	2
Costa Rica	2009–2015	5	Netherlands	2003–2015	11
Croatia	2011–2012	2	New Zealand	2000–2015	14
Czech Republic	2001–2015	13	Norway	2000–2015	14
Denmark	2000–2012	13	Peru	2000–2015	14
Estonia	2006–2015	8	Poland	2000–2015	14
Finland	2000–2015	14	Portugal	2000–2012	13
France	2000–2015	14	Qatar	2006–2015	8
Georgia	2009–2015	2	Romania	2005–2015	9
Germany	2002–2015	12	Russia	2009–2015	5
Greece	2000–2008	9	Singapore	2009–2015	5
Hong Kong	2000–2015	14	Slovak Republic	2003–2015	11
Hungary	2000–2012	13	Slovenia	2006–2015	8
Iceland	2000–2102	13	Spain	2000–2015	14
Indonesia	2000–2015	14	Sweden	2000–2015	14
Ireland	2000–2102	13	Switzerland	2000–2015	14
Israel	2000–2015	14	Thailand	2000–2015	14
Italy	2000–2102	13	Tunisia	2003–2015	11
Japan	2000–2015	14	Turkey	2003–2009	7
Kazakhstan	2009–2012	4	UK	2000–2015	14
Korea	2000–2015	14	USA	2000–2012	13
Kyrgyz Republic	2006–2009	4	Uruguay	2003–2012	10

A large number of observations allowed for secure GLS estimation of Equation (6) and reduced the risk of attaining inadequate values of the standard errors as suggested by Beck and Katz [47]. The results of the estimation are provided in Table 2.

Table 2. Estimates of model (6) for different set of skills measured in PISA tests.

Variable	Reading	Mathematics	Science
GEE(-3)	1.077 ***	1.079 **	0.247
GEE(-6)	2.846 ***	0.927 **	1.089 ***
GEE(-9)	0.215	0.681 *	0.484 *
Country-specific effect		-skipped-	
Constant	367.7 ***	379.6 ***	390.3 ***
<i>N</i>	638	638	638
<i>R</i> ²	0.938	0.953	0.957
ρ	0.7003 ***	0.7349 ***	0.7405 ***
chi ² for $u_i = \text{cons}$	14,332.68 ***	7101.90 ***	19,308.63 ***

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; particular columns contain the estimates of model (6) for the PISA score in reading, mathematics, and science; GEE(-3), GEE(-6), and GEE(-9) represent government expenditures on education lagged by 3, 6, and 9 years, respectively; country-specific effects have were skipped because of space limitations and are available from the authors on request; n is the total number of observations; ρ is the estimated AR(1) coefficient (the p value referring to the $H_0: \rho = 0$); chi² for $u_i = \text{cons}$ provides the test statistic for $H_0: u_i = \text{cons}, i = 1, \dots, N$.

Three earlier values of the government expenditures of education have been included in the model: 3-, 6-, and 9-year-long lags, which were supposed to capture the potential middle-, long-, and very long-term effects. Clearly, in the case of each of the three types of skills measured in the PISA tests, there existed a notable influence of earlier government expenditures on education on the results measured in the tests, and, while each of the lags mattered at least for some of the aspects, they were the 6-year-lagged inputs that demonstrated the most doubtless relevance. The results confirmed the positive influence of the educational expenditures on the attained results as a whole and (together with the country-specific effects) allowed for the backward prediction of hypothetical PISA scores if those had been performed in the earlier years. The updated data on government educational expenditures provided in [48] is once more used for that purpose.

The latter allows for the estimation of the augmented Solow GDP growth model (9) with education quality approximated with PISA scores. A total of 1731 observations in the period 1994–2015 were used, and the sample included 111 countries throughout the world (conditional upon availability of the considered growth factors). It was far from obvious by how much the extrapolated PISA scores should be lagged, which is equivalent to asking at what age the graduates enter the labor market to the extent that allows their skills to have a key influence on the economy's functioning. We used the 5-, 10-, and 15-year lags of the predicted PISA scores (named reading(-15), mathematics(-15), and science(-15), respectively) and concluded that the 15-year-lagged scores seemed most relevant, thus further exploiting these results (presented in Table 3). Moreover, there were numerous possibilities regarding the rest of the operationalization of x_{it} —Sala-i-Martin et al. [29] mentioned that there are approximately 600 variables used as such in the empirical models present in the literature. We used four important human and physical capital growth factors that were found significant in most of the empirical research: *inv* (gross capital formation as % of the GDP), *gov* (total government expenditures as % of the GDP), *life_exp* (life expectancy at birth), and *open* (value of trade as % of the GDP) in order to explain the logarithm of the GDP in constant prices (yet contradictory conclusions regarding the direction of the influence of the government expenditures were drawn in different studies). The $\ln\text{GDP}(-1)$ represents the logarithm of the one-year-lagged logarithm of the GDP, as in model (9); thus, the estimate of the parameter on $\ln\text{GDP}(-1)$ was the estimate of the $\gamma + 1$ and would need to be subtracted from 1 in order to obtain the annual speed of the beta convergence.

Table 3. Estimates of the augmented Solow GDP growth model (9) with education quality.

	(1) Reading	(2) Mathematics	(3) Science
lnGDP(-1)	0.9801828 ***	0.9805469 ***	0.9802120 ***
inv	0.0016506 ***	0.0016418 ***	0.0016490 ***
gov	−0.0007520	−0.0007466	−0.0007275 **
life_exp	0.0017133 ***	0.0017415 ***	0.0017211 ***
open	0.0000764 ***	0.0000791 ***	0.0000783 **
reading(-15)	0.0001130 ***		
mathematics(-15)		0.0000765 ***	
science(-15)			0.0001091 **
Constant	0.0016247	0.0111554	0.0008775
<i>n</i>	1731	1731	1731
Arellano–Bond AR(1)	−2.7035 (0.0069)	−2.7037 (0.0069)	−2.7038 (0.0068)
Arellano–Bond AR(2)	−0.49696 (0.6192)	−0.49529 (0.6204)	−0.49610 (0.6198)

*** $p < 0.01$, ** $p < 0.05$; models (1), (2), and (3) include the reading, mathematical, and scientific skills as regressors, respectively; n is the total number of observations; number of countries in the dataset = 111 world countries; Arellano–Bond AR(1) is the test statistic (p value) of the Arellano–Bond test of autocorrelation of order 1 (order 2 in the case of Arellano–Bond AR(2)).

It should first be noticed that all of the macroeconomic variables demonstrated the influence that should be expected on the theoretical basis: the level of investment, life expectancy, and openness of the economy had a clear positive influence on the GDP growth rate, whereas the influence of the government expenditures remains unclear. The estimate of the parameter by the lagged GDP confirmed the existence of strong convergence processes with a rate of slightly below 2% per year, which confirmed the mainstream literature result. At the same time, the results of the Arellano–Bond test for autocorrelation of order two in the differenced equation that corresponds to the order one autocorrelation in the level equation provided no reason to reject the null hypothesis of no autocorrelation, which was essential for the consistency of the estimator.

Most importantly, each of the PISA scores, interpreted as a proxy for the quality of different skills taught at school, proved statistically significant for the GDP growth at a 5% significance level (1% in the case of reading and mathematics skills). Since in order to assess this result we used 15-year-long lags, this meant using the quality of education of the 30-year-olds at the time when they were approximately 15 years old using the GDP growth factor. While this result was also largely confirmed if 10-year-long lags were used (which meant including the quality of education of today's 25-year-olds at the time when they were 15), it remained most clear in the presented group (the results of the estimation with 10-year lags of the PISA scores are available on request from the authors). No greater differences could be seen while comparing the influence of each of the separate PISA indicators: while the exact estimate of the parameter was not the same, the difference between the estimated coefficients for reading, mathematical, and science literacy lagged by 15 years was not statistically significant (assuming any reasonable level of significance, such as 5%).

While all the discussed results confirmed the intuitive and were confirmed by the provided theoretical model relevance of the quality of education for the GDP growth, we ran a series of robustness checks.

Firstly, it may raise doubts about whether computing and exploiting the theoretical PISA scores in the group of countries that never actually ran the test does not require stronger homogeneity of the population used to estimate the growth model (9) than indeed took place. Thus, we restricted the estimation sample to the countries used to estimate model (6), whose list is provided in Table 1. We also restricted the list of lags of the regressors used as instruments to just the two most recent usable values so as to avoid the risk of instrument proliferation [49]. The estimates of the respective models are provided in Table 4.

Table 4. Estimates of the augmented Solow GDP growth model (9) with education quality in the sample of countries that performed PISA tests.

	(1) Reading	(2) Mathematics	(3) Science
lnGDP(-1)	0.9954993 ***	1.0046150 ***	1.0025350 ***
lnGDP(-2)	−0.1691006 ***	−0.1765280 ***	−0.1754891 ***
lnGDP(-3)	0.1568341 ***	0.1551946 ***	0.1562269 ***
Inv	0.0030705 ***	0.0029663 ***	0.0030006 ***
Gov	−0.0020520 **	−0.0019407 **	−0.0019221 **
life_exp	−0.0027801 **	−0.0026226 **	−0.0027087 **
Open	0.0000928 ***	0.0000793 ***	0.0000873 **
reading(-15)	0.0002229 ***		
mathematics(-15)		0.0001747 ***	
science(-15)			0.0002083 ***
Constant	0.2615308 **	0.2730802 ***	0.2610700 ***
N	408	408	408
Arellano–Bond AR(1)	−3.593 (0.0003)	−3.6714 (0.0002)	−3.6722 (0.0002)
Arellano–Bond AR(2)	−1.0928 (0.2745)	−1.0499 (0.2938)	−1.0414 (0.2977)

*** $p < 0.01$, ** $p < 0.05$; models (1), (2), and (3) include the reading, mathematical, and scientific skills lagged by 15 years as regressors, respectively, while all models include the natural logarithm of the GDP lagged by 1, 2, and 3 years; n is the total number of observations; number of countries in the dataset = 42 countries participating in the PISA tests (listed in Table 1); Arellano–Bond AR(1) is the test statistic (p value) of the Arellano–Bond test of autocorrelation of order 1 (order 2 in the case of Arellano–Bond AR(2)); 2nd and 3rd lag of the dependent variable were added to avoid autocorrelation, which was present otherwise.

Secondly, we included the fixed time effects, which helped capture the global economic cycle peaks and troughs and omitted the government expenditures, whose influence was unambiguous in the light of both the theory and the presented empirical results. The estimates are provided in Table 5.

Table 5. Estimates of the augmented Solow GDP growth model (9) with education quality with time effects.

	(1) Reading	(2) Mathematics	(3) Science
lnGDP(-1)	0.9683281 ***	0.9688890 ***	0.9685778 ***
Inv	0.0021138 ***	0.0021080 ***	0.0021265 ***
life_exp	0.0030568 ***	0.0030578 ***	0.0030005 ***
Open	0.0000162	0.0000174	0.0000146
reading(-15)	0.0001878 **		
mathematics(-15)		0.0001514 *	
science(-15)			0.0002033 **
Constant	−0.0371567	−0.0277849	−0.0445571
N	1733	1733	1733
Wald, sign. of the time effects	361.70 (0.0000)	361.11 (0.0000)	357.79 (0.0000)
Arellano–Bond AR(1)	−2.662 (0.0078)	−2.6617 (0.0076)	−2.6605 (0.0078)
Arellano–Bond AR(2)	0.01578 (0.9874)	0.01789 (0.9857)	0.02021 (0.9839)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; models (1), (2), and (3) include the reading, mathematical, and scientific skills as regressors, respectively; n is the total number of observations; number of countries in the dataset = 111 world countries; instruments lagged by up to three periods were used; Arellano–Bond AR(1) is the test statistic (p value) of the Arellano–Bond test of autocorrelation of order 1 (order 2 in the case of Arellano–Bond AR(2)); Wald, sign. of the time effects is the test statistic (p value) for the Wald test of the joint significance of the time effects.

Thirdly, we ran the complex robustness analysis in the spirit of Bayesian model averaging (BMA). While the classical article [50] discusses its version for the growth models estimated with the use of the OLS estimator (which usually would apply to a cross-section of countries), we applied the approach of Prochniak and Witkowski [51], assuming that the beta convergence does exist and verifying the relevance of the particular growth factors. We thus began (separately for each of the considered PISA skills) with estimating the autoregressive model (9) with all the possible subsets of the considered

macroeconomic growth factors {inv, gov, life_exp, open} and the PISA skills in one of the areas. We made no prior assumptions regarding the number of the relevant growth factors in the true relation, and the number of the considered regressors was not large. Thus, it allowed us to assume that the prior probability of the relevance of each of the possible models was the same. Under that assumption, we followed [51] and computed the averaged estimates of the parameters γ and θ as well as the p-values for the t-test of significance, treating the posterior probabilities of the relevance for each of the models as weights. The BMA-averaged estimates of γ and θ and the p-values obtained in the main sample of 111 countries are provided in Table 6. Additionally, we completed it with an approach based on Leamer's [52] extreme bound analysis (EBA) and provided the range of the estimate of particular parameters in all the estimated models as well as the fraction of the models that confirm the hypothesis of the given variable's significance.

Table 6. BMA and EBA estimation results for the augmented Solow GDP growth model (9) with education quality.

	(1) Reading	(2) Mathematics	(3) Science
Bayesian Model Averaging Estimates (<i>p</i>-Value for the Significance Test Given in Brackets)			
lnGDP(-1)	0.9704477 (0.0000000)	0.9694659 (0.0000000)	0.9684520 (0.0000000)
Inv	0.0017909 (0.0000000)	0.0017853 (0.0000000)	0.0017889 (0.0000000)
Gov	−0.0005426 (0.2030227)	−0.0005388 (0.2075655)	−0.0005249 (0.2185428)
life_exp	0.0027987 (0.0000000)	0.0028204 (0.0000000)	0.0028047 (0.0000000)
Open	0.0002220 (0.0017697)	0.0002186 (0.0023417)	0.0002213 (0.0018750)
reading(-15)	0.0005646 (0.0060577)		
mathematics(-15)		0.0006020 (0.0069853)	
science(-15)			0.0006809 (0.0077673)
Extreme Bound Analysis—Type Statistics for the PISA Score Variables (Reading, Mathematics and Science)			
% confirming relevance at 1%	56.3%	56.3%	56.3%
% confirming relevance at 5%	75.0%	68.8%	62.5%
% confirming relevance at 10%	81.3%	75.0%	75.0%
lowest parameter estimate	0.0000144	−0.0000272	−0.0000244
highest parameter estimate	0.0006132	0.0006581	0.0007419
% negative estimates	0.0%	12.5%	6.3%
<i>N</i>	1731	1731	1731

Models (1), (2), and (3) include the reading, mathematical, and scientific skills as regressors, respectively; *n* is the total number of observations; number of countries in the dataset = 111 world countries; the averaged parameter estimates and *p* values are provided in the first (BMA) part of the table; the fraction of models in which PISA scores regarding the given skills are significant at the 1%, 5%, and 10% levels with the lowest and the highest obtained estimate of the parameter on the PISA score variable and the fraction of models in which the estimate of the parameter on the PISA score are negative are provided in the second (EBA) part of the table.

The results obtained in a series of robustness checks fully confirmed the validity of the baseline analysis. The significance of the quality of education was stable and exhibited no sensitivity to structural changes of the model, such as the inclusion of time effects, the inclusion of additional earlier lags of the GDP enforced by the autocorrelation in the initial form of the model, or manipulation of the classical growth factors (both by eliminating the government expenditures whose impact cannot be fully confirmed, as well as in light of the BMA based on every possible subset of the considered growth factors).

The changes in the technical details of the estimation approach (manipulating the lag length of the set of applied instruments) as well as the changes in the sample used to estimate the growth model (limiting the sample to the countries that have performed PISA tests) did not result in any changes in the qualitative conclusions regarding the relevance of the PISA scores. The validity of these results was reassured given that other considered GDP growth factors exhibited a stable influence that complies with the mainstream literature: while there is an open discussion regarding the impact of the government expenditures on the economic growth, other growth factors in most models have exhibited the expected properties (an exception to that is that the unconfirmed influence of the economic openness is just one of the robustness checks), and the estimated rate of convergence complies with most of the existing research.

The only doubts may be raised by the results of the EBA-type analysis, given that the formal confirmation of the relevance of PISA scores was observed in 75.0–81.3% of the cases (at the 10% significance level) for different skills equations and that there exist single cases in which the estimate of the parameter on the PISA score variable is slightly lower than zero. However, it must be emphasized that Leamer's approach was mostly criticized because of its being too stringent: in most empirical analyses there are hardly any (or no) variables that are unambiguously found robustly significant in the EBA sense. This is particularly important when the very sensitive Blundell–Bond approach is used: indeed, all the “doubtful” results are obtained in the models that included the government expenditures among the regressors. Still, even in its presence, the much more widely acknowledged BMA procedures raise no doubts about the relevance of the PISA scores for the GDP growth processes. As a result, we believe that the robustness checks provided very strong confirmation of the results provided by the baseline model.

5. Conclusions

It is common knowledge that education is important, and there is overwhelming evidence that better education gives great returns to individuals. Oreopoulos and Salvanes [53] suggest that better education might lead individuals to make better decisions about health, marriage, and parenting style. It is also believed that schooling improves patience, making individuals more goal-oriented and less likely to engage in risky behavior. Yet, at the macroeconomic level, there is little empirical evidence that better education in a given country translates into better economic results. This can be mostly attributed to the quantitative, not qualitative, data on education available across countries. While Hanushek and Woessmann [6] published evidence of the importance of education quality as a factor of GDP growth, their results might cause some doubts, mostly because they investigated the contemporaneous relationship between the two—or, to be more accurate, the correlation between the 1960–2010 GDP growth and the available PISA scores (from the beginning of the current century). As a result, their research confirms the existence of the relation itself; however, its direction is most likely opposite to the claim of the authors—or at least it is difficult to identify the direction of the relation. A possible solution would be to use properly lagged PISA results and include them in the GDP growth regression. An obvious problem is the lack of the lagged PISA results given that the tests started at the beginning of the current century. As a partial solution to this problem, we suggested estimating a model that explains the results of PISA as a function of educational expenditures. The model would then be used to provide backward predictions of PISA scores, and the final model of GDP growth could be estimated with the use of adequately lagged PISA scores in the role of input.

The resulting model seems to be rational from the economic and the educational point of view. Although the results are in line with earlier claims by the cited authors, it provides stronger evidence for the relevance of the quality of education as it is statistically more robust, and the properties of the applied estimator are generally better. As a result, applying the alternative technique should be viewed as an important value-add of this research. Justifying the value of a good education is of double importance. In practical

terms, it is an important voice in the discussion and provides an additional argument for directing a stream of investment in education, which is particularly important while governments might start searching for savings if global markets fall into the recession phase of the economic cycle.

One might wonder why the different types of skills have such a similar influence on the performance of the economies. Certainly, a few reasons could be given. Firstly, each of them should be viewed as a proxy for the general quality of education in a given school, district, or even country—probably more than the indicator of the level of teaching of a particular class. Secondly, the final PISA scores in different areas are in some cases constructed with the use of their values (or actually, their components) in various areas. As a result, the distributions of different PISA scores are not fully independent, which can be partly reflected in the above-described similarities between the three different models presented in Table 2. In any case, it can be believed that the methodology discussed here should be viewed as a formal confirmation that expenditures on education, which result in higher education quality, just pay back. The return is not immediate, but some years after the graduates enter the labor market, the quality of the education that they attained at the age of 15 begins to matter.

Several elements are crucial for the properly constructed model and for trustworthy conclusions. An important question is how long after graduation the employees have the greatest impact on the total productivity in the economy and contribute the most to the GDP growth. While answering this question would suggest the adequate lag length for the models analyzed in the study, it is not simple to do so. It seems rational to assume that the employees of crucial significance should obtain better remuneration. This process, however, exhibits little stability: while in 1975 they were the 29-year-olds who had the highest average wages, recently, the peak is observed in the cohort of 40-year-olds. That could be attributed to the increasing professionalization and an increasing role of knowledge and experience in the labor market, which suggests that the significance of the quality of education has increased and might be expected to further increase in the future. However, such a result makes it more challenging to properly lag the regressors in the model equations. Still, the robustness analysis partly described in this paper and delivered by Witkowska and Witkowski [54] leaves no doubts: while the results are observed most clearly with lags of 15 years, which means considering the education quality of today's 30-year-olds, the figures are very similar for the 25- as well as the 40-year-olds.

The transmission channel in the analyzed phenomenon is interesting. While in the theoretical model we concentrated on the economic aspects of education quality, those are not limited to these. Authors in [4,55] have pointed out that, on the individual level, expenditures on education as well health will develop adequate competencies and improve the state of health so that the productivity and income of that person will increase in the future. These two factors, education and health, have an impact on human productivity, which has an impact on production, and with an increase in production, economic growth will also increase. Therefore, education and health, which are important components of human capital, have an impact on economic growth. A study on economic growth in Korea and Japan by Han and Lee [56] provides empirical arguments that there is strong cointegration between health services and education in improving the quality of human resources and economic growth.

Yet another transmission channel to be taken into consideration is the democracy–education nexus. While most of the highest-developed countries in the world are adult democracies, there is a clear relationship between education and democracy across countries [57,58]; however, the reason for this remains unclear. In their study [59], they proposed the explanation hinging on the connection between education and the costs and benefits of political engagement. Schools not only educate but also socialize young people, and political involvement is a form of the latter. There is numerous evidence showing a positive connection between education and civic engagement. Ref. [59] models education as raising the benefits of political action when individuals choose to support a more or less democratic

regime. In this model, democratic regimes offer weak incentives to a wide base of potential supporters, whereas dictatorships offer strong incentives to a narrower base. Education increases the society-wide support for democracy because democracy relies on people with high participation benefits for its support. The authors showed that better-educated nations are more likely both to protect democracy and to undertake effective efforts to prevent coups. The performed analysis additionally raised two broader questions. First, whereas the model itself focused on the effects of education on participation, the analysis applied to 32 all social glues that encourage collective action; so, perhaps the analysis suggests a solution to Olson's free-rider problem in all organizations, and not just in political regimes—namely, human capital or other kinds of social glue as a motivation to participate. Secondly, the results shed a light on the problem of why some dictators invest in education that might be a threat to them.

One of the possible answers is that many dictators face an external threat and, therefore, must grow their economies and their armies (including investing in human capital) to counter these threats even if this raises the risk of democratization. A second answer is that, even with a lack of external threats, dictators might benefit from economic growth, and, therefore, they might promote education to become richer. A third idea is that all dictators face significant ouster risks and that it is much better for the dictator's life for him to be replaced by democracy in an educated country than by another dictator in an uneducated one.

Fortunato & Panizza [57] in their study on the interaction between democracy and education and its impact on the quality of government, draw three important conclusions. Firstly, the interaction between democracy and education is always positively and significantly correlated with the quality of government. Secondly, the correlation between democracy and quality of government is statistically significant only in countries with high levels of education. Thirdly, the marginal effect of education is positive and statistically significant in countries with high levels of democracy. In their model [57], they synthesized, in one framework, the stance emphasizing the importance of political institutions as a fundamental factor explaining cross-country differences in income per capita with the stance that institutional improvements and development is driven by social and human capital. The most important empirical finding from this work is the conclusion that democratic institutions and education complement each other, but they argue that democracy leads to the election of better candidates only in the situation where the level of education is above a certain threshold. Simultaneously, amelioration of education can affect the quality of the elected officials but only if the cost of entry into politics is not prohibitive. The authors ran a set of Monte Carlo simulations to show that these results were not driven by reverse causality. By looking explicitly at the interaction between democracy and education, they demonstrated how these two variables complement each other in the selection of high-quality policymakers, which guarantees good governance.

In addition, we should bear in mind that economic growth is an important facet, but just one of many, of country development. We can expect societies with higher education quality to be more democratic and politically stable, to exhibit less violence, poverty and inequality, and to enjoy a higher quality of governance. All of these additional factors associated clearly with higher education quality can have a noteworthy positive impact on both economic growth and society's welfare, going far beyond simple economic calculation. The development of a knowledgeable population does not only contribute to economic growth itself but also might contribute to such aspects of national well-being as welfare and poverty reduction [3]. The authors in [58] also agree that the development of economic growth analysis provides a basis for the role of human capital as an important part of increasing economic growth. Wensley and Evans [60] are convincing that the higher the quality of human capital, the higher its effect on economic growth, and there are numerous studies stating that education is of particular importance for growth in developing countries [2,61–63].

The above results seem to be an important confirmation of the role of education not just for the well-being of individuals but also for the well-being of entire societies. Of course, the milestone study [6] and the earlier analysis of Hanushek and Woessmann suggested the existence of such a relation. However, we believe that this study is the first to confirm them with the use of modern econometric tools that include not just the dynamic panel data models but also the BMA approach. Its strength consists in the elimination of a vast amount of subjectivity that accompanies the construction of a single model. Instead, a number of models were analyzed and averaged, confirming the validity of the results. They seem vital, especially in the pandemic era when numerous governments will be looking for various areas in which the costs can be cut to compensate for the recent excessive expenditures on healthcare and lockdown support. The decision of where to cut costs will be challenging; however, the conclusions of this study are clear: saving on the quality of education in the middle and long time horizon will not pay off in terms of economic growth and should not be considered as a profitable solution. On the other hand, our study has natural limitations. The crucial one is the limited number of lagged PISA scores due to the relatively short history of this tool. Secondly, although the number of countries that participate is quite large today, initially, it was notably lower. These shortcomings simply require more time. Secondly, while we believe that the PISA scores are the most adequate measure of quality of education, they are not perfect either.

Providing high-quality education for humankind is of crucial importance and, as such, has been listed as one of the priorities on various global development agendas, such as the United Nations' Sustainable Development Goals (SDGs) of the 2030 Agenda for Sustainable Development [64]. Education is crucial for individual and social development given that it allows for the transmission of knowledge and facilitates the ability to understand and cope with the surrounding world in addition to inspiring innovation [65]. Good education reduces poverty and promotes prosperity.

Author Contributions: Conceptualization, Ł.G., E.W. and B.W.; methodology, Ł.G. and B.W.; validation, Ł.G. and E.W.; formal analysis, Ł.G. and B.W.; resources, Ł.G.; writing—original draft preparation, E.W., Ł.G. and B.W.; writing—review and editing, E.W. and B.W.; project administration, E.W.; funding acquisition, E.W., Ł.G. and B.W. All authors have read and agreed to the published version of the manuscript.

Funding: 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. The cost of publication has been additionally co-financed by University of Warsaw, The Maria Grzegorzewska University and SGH Warsaw School of Economics.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data used in this paper are freely and publicly available in the World Bank national accounts data, OECD National Accounts data files, and the OECD Programme for International Student Assessment (PISA).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hanushek, E.A.; Woessmann, L. Education, knowledge capital, and economic growth. In *The Economics of Education*; Bradley, S., Green, C., Eds.; Academic Press: London, UK, 2020; pp. 171–182.
2. Widarni, E.L.; Bawono, S. Human Capital, Technology, and Economic Growth: A Case Study of Indonesia. *J. Asian Financ. Econ. Bus.* **2021**, *8*, 29–35.
3. Hess, P.N. *Economic Growth and Sustainable Development*; Routledge: London, UK, 2016.
4. Rajan, S.I. *India Migration Report 2020: Kerala Model of Migration Surveys*; Taylor and Francis: London, UK, 2020.
5. Felipe, J.; Widyasanti, A.; Foster-McGregor, N.; Sumo, V. *Policies to Support the Development of Indonesia's Manufacturing Sector during 2020–2024: A Joint ADB–BAPPENAS Report*; Asian Development Bank: Jakarta, Indonesia, 2019.
6. Hanushek, E.A.; Woessmann, L. *The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes*; OECD Publishing: Paris, France, 2020.

7. Barro, R.J.; Lee, J.W. Education Matters. In *Global Schooling Gains from the 19th to the 21st Century*; Oxford University Press: Oxford, UK, 2015.
8. Awaworyi, C.S.; Ugur, M.; Yew, S.L. Government education expenditures and economic growth: A meta-analysis. *BE J. Macroecon.* **2017**, *17*, 1–17.
9. Blundell, R.; Bond, S. Initial conditions and moment restrictions in dynamic panel data models. *J. Econom.* **1998**, *87*, 115–143. [[CrossRef](#)]
10. Morris, S.; Oldroyd, J. *International Business*; John Wiley & Sons: Hoboken, NJ, USA, 2020.
11. Xu, Y.; Aiya, L. The relationship between innovative human capital and interprovincial economic growth based on panel data model and spatial econometrics. *J. Comput. Appl. Math.* **2020**, *365*, 112381. [[CrossRef](#)]
12. Flynn, J.R. *What is Intelligence? Beyond the Flynn Effect*; Cambridge University Press: Cambridge, UK, 2007.
13. Jones, R.; Jenkins, F. *Managing Money, Measurement, and Marketing in the Allied Health Professions*; CRC Press: London, UK, 2018.
14. Hanushek, E.A.; Schwerdt, G.; Wiederhold, S.; Woessmann, L. Returns to skills around the world: Evidence from PIAAC. *Eur. Econ. Rev.* **2015**, *73*, 103–130. [[CrossRef](#)]
15. Mincer, J. Investment in human capital and personal income distribution. *J. Political Econ.* **1958**, *66*, 281–302. [[CrossRef](#)]
16. Castriota, S. Education and Happiness: A Further Explanation to the Easterlin Paradox. Unpublished Paper 2006. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.607.1432> (accessed on 4 June 2021).
17. Mankiw, N.G.; Romer, D.; Weil, D.N. A contribution to the empirics of economic growth. *Q. J. Econ.* **1992**, *107*, 407–437. [[CrossRef](#)]
18. Klenow, P.; Rodriguez-Clare, A. The Neoclassical Revival in Growth Economics: Has It Gone Too Far? *NBER Macroecon. Annu.* **1997**, *12*, 73–103. [[CrossRef](#)]
19. Bosworth, B.; Collins, S. The empirics of growth: An update. *Brook. Pap. Econ. Act.* **2003**, *2*, 113–206. [[CrossRef](#)]
20. Benos, N.; Zotou, S. Education and economic growth: A meta-regression analysis. *World Dev.* **2014**, *64*, 669–689. [[CrossRef](#)]
21. Schoellman, T. Education Quality and Development Accounting. *Rev. Econ. Stud.* **2012**, *79*, 388–417. [[CrossRef](#)]
22. Hanushek, E.A.; Woessmann, L. *The Knowledge Capital of Nations: Education and the Economics of Growth*; MIT Press: Cambridge, MA, USA, 2015.
23. Scheerens, J.; Luyten, H.; Van Ravens, J. *Measuring Educational Quality by Means of Indicators. Perspectives on Educational Quality*; Springer: Dordrecht, The Netherlands, 2011; pp. 35–50.
24. OECD. *PISA Report: What It Takes to Learn*; OECD: Paris, France, 2012.
25. Federowicz, M. *Program Międzynarodowej Oceny Umiejętności Uczniów OECD PISA. Wyniki Badania 2009 w Polsce*; IBE: Warsaw, Poland, 2010.
26. Schleicher, A.; Zimmer, K.; Evans, J.; Clements, N. *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*; OECD Publishing (NJ1): Paris, France, 2009.
27. Dolata, R.; Jakubowski, M.; Pokropek, A. *Polska Oświata w Międzynarodowych Badaniach Umiejętności Uczniów PISA OECD. Wyniki, Trendy, Kontekst i Porównywalność*; Wydawnictwa Uniwersytetu Warszawskiego: Warszawa, Poland, 2013.
28. Anderson, J.O.; Chiu, M.H.; Yore, L.D. First cycle of PISA (2000–2006)—International perspectives on successes and challenges: Research and policy directions. *Int. J. Sci. Math. Educ.* **2010**, *8*, 373–388. [[CrossRef](#)]
29. Haskell, R.E. *Transfer of Learning: Cognition and Instruction*; Academic Press: Cambridge, MA, USA, 2020.
30. Sellar, S.; Lingard, B. The OECD and global governance in education. *J. Educ. Policy* **2013**, *28*, 710–725. [[CrossRef](#)]
31. OECD. *PISA Results: Excellence through Equity-Giving Every Student the Chance to Succeed*; OECD: Paris, France, 2012.
32. West, M. Global lessons for improving US education. *Issues Sci. Technol.* **2012**, *28*, 37–44.
33. Cheung, H.Y.; Chan, A.W. Understanding the relationships among PISA scores, economic growth and employment in different sectors: A cross-country study. *Res. Educ.* **2008**, *80*, 93–106. [[CrossRef](#)]
34. Chen, S.S.; Luoh, M.C. Are mathematics and science test scores good indicators of labor-force quality? *Soc. Indic. Res.* **2010**, *96*, 133–143. [[CrossRef](#)]
35. Hanushek, E.A.; Kimko, D.D. Schooling, labor-force quality, and the growth of nations. *Am. Econ. Rev.* **2000**, *90*, 1184–1208. [[CrossRef](#)]
36. Hanushek, E.A.; Woessmann, L. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *J. Econ. Growth* **2012**, *17*, 267–321. [[CrossRef](#)]
37. Hanushek, E.A.; Woessmann, L. The role of cognitive skills in economic development. *J. Econ. Lit.* **2008**, *46*, 607–668. [[CrossRef](#)]
38. Hanushek, E.A.; Woessmann, L. The economics of international differences in educational achievement. In *Handbook of the Economics of Education*; Elsevier: Amsterdam, The Netherlands, 2011; Volume 3, pp. 89–200.
39. Barro, R.J. Economic growth in a cross section of countries. *Q. J. Econ.* **1991**, *106*, 407–443. [[CrossRef](#)]
40. Próchniak, M.; Witkowski, B. Bayesian model averaging in modelling GDP convergence with the use of panel data. *ROczniki Kolegium Analiz Ekonomicznych.* **2012**, *26*, 45–60.
41. Arellano, M.; Bond, S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Rev. Econ. Stud.* **1991**, *58*, 277–297. [[CrossRef](#)]
42. Lucas, R.E., Jr. On the mechanics of economic development. *J. Monet. Econ.* **1988**, *22*, 3–42. [[CrossRef](#)]
43. Aghion, P.; Ljungqvist, L.; Howitt, P.; Howitt, P.W.; Brant-Collett, M.; García-Peñalosa, C. *Endogenous Growth Theory*; MIT Press: Cambridge, MA, USA, 1998.
44. Arrow, K.J. *The Theory of Risk Aversion, in Aspects of the Theory of Risk Bearing*; Ynjo Jahnsolin Saatio: Helsinki, Finland, 1965.

45. French, J.; French, A.; Li, W. The relationship among cultural dimensions, education expenditure, and PISA performance. *Int. J. Educ. Dev.* **2015**, *42*, 25–34. [[CrossRef](#)]
46. Heyneman, S.P. Economics of education: Disappointments and potential. *Prospects* **1995**, *25*, 557–583. [[CrossRef](#)]
47. Beck, N.; Katz, J.N. What to do (and not to do) with time-series cross-section data. *Am. Political Sci. Rev.* **1995**, *89*, 634–647. [[CrossRef](#)]
48. Lee, J.W.; Lee, H. Human capital in the long run. *J. Dev. Econ.* **2016**, *122*, 147–169. [[CrossRef](#)]
49. Roodman, D. A note on the theme of too many instruments. *Oxf. Bull. Econ. Stat.* **2009**, *71*, 135–158. [[CrossRef](#)]
50. Sala-I-Martin, X.; Doppelhofer, G.; Miller, R. Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *Am. Econ. Rev.* **2004**, *94*, 813–835. [[CrossRef](#)]
51. Próchniak, M.; Witkowski, B. Time stability of the beta convergence among EU countries: Bayesian model averaging perspective. *Econ. Model.* **2013**, *30*, 322–333. [[CrossRef](#)]
52. Leamer, E.E. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*; John Wiley & Sons Incorporated: New York, NY, USA, 1978.
53. Oreopoulos, P.; Salvanes, K.G. Priceless: The nonpecuniary benefits of schooling. *J. Econ. Perspect.* **2011**, *25*, 159–184. [[CrossRef](#)]
54. Witkowska, E.; Witkowski, B. Młodzi profesjonaliści czy doświadczeni eksperci? Kluczowi dla wzrostu gospodarczego wyspecjalizowani uczestnicy rynku pracy. In *Metody Probabilistyczne w Zastosowaniach Ekonomicznych*; Męczarski, M., Ed.; Oficyna Wydawnicza SGH: Warszawa, Poland, 2020.
55. Calkin, S. *Human Capital in Gender and Development*; Routledge: London, UK, 2018.
56. Han, J.S.; Lee, J.W. Demographic change, human capital, and economic growth in Korea and Japan. *World Econ.* **2020**, *3*, 1–13.
57. Piergiuseppe, F. Panizza, U. Democracy, education and the quality of government. *J. Econ. Growth* **2015**, *20*, 333–363.
58. Hahn, F.; Coricelli, F.; Di Matteo, M. *New Theories in Growth and Development*; Springer: Cham, Switzerland, 2016.
59. Glaeser, E.L.; Ponzetto, M.; Shleifer, A. Why does democracy need education? *J. Econ. Growth* **2006**, *12*, 77–99. [[CrossRef](#)]
60. Wensley, A.; Evans, M. *17th International Conference on Intellectual Capital, Knowledge Management & Organizational Learning*; University of Toronto: Toronto, ON, Canada, 2020.
61. Banerjee, R.; Roy, S.S. Human capital, technological progress, and trade: What explains India's long-run growth? *J. Asian Econ.* **2014**, *30*, 15–31. [[CrossRef](#)]
62. Alfada, A. The destructive effect of corruption on economic growth in Indonesia: A threshold model. *Heliyon* **2019**, *5*, 1–8. [[CrossRef](#)] [[PubMed](#)]
63. Xu, H.; Wei-Ling, H.; Teen-Hang, M.; Hua Zhu, J. Can higher education, economic growth and innovation ability improve each other? *Sustainability* **2020**, *12*, 2515. [[CrossRef](#)]
64. Assembly, G. Sustainable development goals. *SDGs Transform. Our World* **2015**, *2030*, 1–41.
65. Guo, L.; Huang, J.; Zhang, Y. Education development in China: Education return, quality, and equity. *Sustainability* **2019**, *11*, 3750. [[CrossRef](#)]