DOI: 10.28934/ea.21.54.1.pp41-59

ORIGINAL SCIENTIFIC PAPER

Housing Quality and Human Capital Formation in Developing Countries

Guilherme Shirazawa[1] | Marta Simões[2*]

1 University of Coimbra, Faculty of Economics, Coimbra, Portugal
2 University of Coimbra, CeBER, Coimbra, Portugal

ABSTRACT
The main objective of this work is to analyse the relationship between the quality of housing and human capital formation in the context of developing countries. The analysis attempts to fill a gap in the current literature regarding the lack of empirical studies that address the impact that living conditions can have on human capital. The study was performed using cross-sectional data, mostly taken from the UNESCO database, for 52 low and middle-income countries. The estimated empirical models consider average years of schooling as the dependent variable and as the explanatory variable of interest the proportion of the population living in houses with below minimum quality standards. The OLS results obtained suggest a negative association between housing quality and average years of schooling, but with little or no statistical significance, making the empirical analysis inconclusive. We pose that this result might relate to the comparability of the housing quality data provided by UNESCO, highlighting the need to gather more data and produce new, more reliable indicators on the topic.

Key words: *Human capital, housing, developing countries, Cross-Sectional Analysis, economic growth, economic development*

JEL Classification: I25, O18, O15, O50, R20

INTRODUCTION

The economic impact of housing investment has gone through several debates over the last few decades. Housing has mainly been treated as a factor that influences short-run macroeconomic performance through its (wealth) effect on consumption and investment (Harris & Arku. 2006). In this study we take a different perspective focusing on the potential long-run economic impact of housing investment through education improvements, an important source of human capital accumulation, in turn a key driver of economic growth and development (Mankiw, Romer and Weil (1992); Lucas (1988), Benhabib & Spiegel (1994); Hanushek & Woessmann (2011)). Our main hypothesis is that housing quality can play an important role in the explanation of human capital availability differences.

Having adequate dwellings has been recognized by the literature as a basic requirement for individuals to develop and become more prosperous (e.g., Healy (1971); Bradley and Putnick (2012)). The main argument is that housing provides basic facilities, such as having access to a good shelter that protects from the elements, to electricity, clean water and a proper

* Corresponding author, e-mail: mcsimoes@fe.uc.pt

environment for cooking. Access to these factors would in turn have a positive effect on productivity and overall personal health, as well as on the performance of children at school. However, the analysis of this relationship has seldom been pursued from an empirical perspective and has never considered the specific impact on human capital.

Given the present gap in the literature, this study aims to empirically investigate the relationship between housing quality and human capital formation in the context of developing countries through the estimation of an empirical model where human capital is the dependent variable and an indicator of the lack of housing quality is our explanatory variable of interest, alongside a set of control variables selected from previous empirical studies on the determinants of human capital. The majority of the data used comes from the UNESCO Institute for Statistics.

The remainder of this study consists of four sections. In the first section we provide an overview of the literature concerning the relationship between housing, human capital formation and macroeconomic performance. In the second section, we present the empirical model and data used. The third section contains the discussion of the results obtained and, finally, the fourth section concludes.

LITERATURE OVERVIEW

In this section we start by giving an overview of the literature on human capital and economic growth in order to motivate our analysis of the relationship between housing and human capital formation and better identify the gap in the literature that the present study tries to address. We next review the scarce (at least at the macro level) literature on the relationship between access to housing and human capital formation.

HUMAN CAPITAL AND ECONOMIC GROWTH

After the initial methodological and empirical foundations laid in the 1960s, economic growth models began to incorporate the concept human capital in the 1980s and 1990s. A pioneer work is the model proposed by Mankiw, Romer and Weil (1992) where the authors extend the Solow (1956) exogenous neoclassical growth model to include human capital as another input into final goods production and subject to diminishing returns just like physical capital. The model shows that this augmented version of the Solow model provides a better explanation of the differences in income per-capita across countries although it is not able to explain the growth rate of output in the long run. In the empirical validation of their model, the authors considered annual data (1960-1985) for a sample of 98 countries and found that the inclusion of human capital made possible to explain about 80 per cent of the variation in income across countries. A different approach is that proposed by Lucas (1988), which lies within the category of AK growth models. In this model, human capital accumulation creates positive externalities due to “learning-by-studying”. In the final goods sector the increase in human capital at the individual level raises the average human capital in the economy by making workers that are in contact with the more educated individuals more productive. In this way, the economy is able to continue growing as a whole even if there are diminishing returns to individual human capital accumulation. Another important landmark in the analysis of the relationship between human capital and economic growth is the work of Romer (1990) in which the growth of output in the long run is the result of intentional decisions made by economic agents in terms of the allocation of resources to an R&D sector that produces new knowledge (non-rival) usable in final goods production. Human capital is viewed as the main input in this R&D sector and thus a major driver of growth.

At the empirical level, Benhabib & Spiegel (1994) benchmark study analyses the relative importance of human capital for economic growth through the different channels discussed above, i.e. distinguishing between the role of human capital in final goods production and as an input into innovation and imitation activities. The empirical model uses time series cross-country data for 78 countries with annual observations from 1965 to 1985, with the proxy for human capital corresponding to average years of schooling retrieved from the Barro-Lee and Kyriacou datasets. The results obtained indicate that human capital plays a major role in the adoption and implementation of new technologies. Additionally, human capital seems to be more relevant to absorb technology from the leading countries than it is to internally develop new technologies, supporting the idea that the cost of imitation activities is lower than that of innovation activities for follower countries. The empirical identification of the role human capital plays in economic growth has also revolved around measurement issues, in particular in what concerns quality vs. quantity of human capital. An example is the work by Hanushek & Woessmann (2011). The authors developed an empirical analysis that focus on the role of human capital, as measured by cognitive skills, in explaining the differences in income per capita across OECD countries, from 1980 to 2000. They use microdata from international achievements tests (PISA– Programme for International Student Assessment scores) for measuring, separately, basic and top skills. This is a more sophisticated approach than considering just measures of the quantity of human capital such as average years of schooling, literacy rates or educational attainment rates, which the authors consider to be a potentially incomplete and misleading measures for human capital as they implicitly assume that learning outcomes from additional years of education are the same across countries. The results from the regressions indicate that cognitive skills are a better predictor of economic growth than average years of schooling, confirming that the quality of human capital is more important for growth than its quantity. In any case, the results obtained still point to average years of schooling as able to explain an important part of long-run growth, which is in line with the empirical findings from Mankiw, Romer and Weil (1992). Even though human capital quality plays a more important role than its quantity in the explanation of economic growth, the two previous studies show that human capital in general is crucial if countries are in pursuit of long-term prosperity.

In a more recent study, Égert, Botev and Turner (2020) investigate the impact of different educational policies on economic growth. For this purpose, the authors analyse the influence that different educational policies have on human capital formation and, as a consequence, on economic growth. The study found that increasing spending in educational policies such as lowering teacher-to-pupil ratios, providing greater autonomy to schools and universities, reduced barriers for university funding, increasing the age of first education tracking (separating students in different education programmes according to performance) and more primary schooling all had a positive and statistically significant impact on aggregate human capital. The results were next used to estimate the gains in terms of economic growth from the implementation of the best practices in terms of educational policies, and meaningful gains in terms of economic growth in the long run were found.

HOUSING AND HUMAN CAPITAL

Despite some theoretical and empirical analyses on the relationship between housing and economic growth (e.g. Green (1997); Hongyu et al. (2002); Terzi and Bolen (2008)), to the best of our knowledge there are no studies that try to analyse its mediating role through human capital formation[1]. Given the importance of the latter for economic growth, we identified some studies on the relationship between housing and human capital formation at the micro level that provide some arguments on what to expect in terms of the sign of the relationship at the aggregate level.

[1] A related recent study is that by Manzoli, Duarte & Simões (2020). Similarly to our study, the authors investigate the role of housing for macroeconomic performance, proxied by the growth rate of real GDP per capita, although the focus is not on the quality of housing but on the housing deficit in the particular situation of Brazil. They find evidence of a negative association between the housing deficit and economic growth, which supports the promotion of policies that facilitate access to housing as a means to promote social inclusion and economic growth.

An earlier study on the topic is Healy (1971). The author analyses the impact of a rehousing program for a group of workers in a Mexican factory on the respective productivity, starting from the hypothesis that improvements in housing conditions can raise either the capacity to work or the desire to work, resulting in greater output per hour worked and lower absenteeism. The empirical analysis considered two groups of workers, those that were rehoused and those that remained in their original low-quality homes, over a period of four years, two years before the rehousing of the first group and two after. In addition to productivity, the study also investigated the effects of the program on worker's absenteeism and health. The author found that one year after the rehousing programme workers' productivity increased and housing-related health problems decreased. Overall, the improvement in the worker's living conditions had a positive effect on the health component of the worker's human capital and may have impacted positively their ability to concentrate and become more productive. These positive relationships leave room to ask whether these positive outcomes could also have an effect on and individual's educational path. In Bradley and Putnick (2012) the authors analyse the relationship between the home environmental conditions that are associated with child development, e.g., housing quality, material resources, formal and informal learning resources, and the Human Development Index (HDI) for 28 developing countries. The study found that the quality of housing and material resources were positively associated with the HDI. Looking at the issue in the context of low-income households, the study from the Citizens Housing and Planning Council (Housing, C., & Council, P. ,2001) analysed a sample of diverse low-income young adults in New York and found that crowded homes, among other factors such as ethnicity, have a negative impact on the probability of a teenager to finish high school.

One aspect of housing that has been known to affect educational outcomes is tenure. Bramley & Karley (2007), for selected areas in both England and Scotland, found that children living in a household where the parents are the homeowners record higher school attainment and better test scores. The authors attribute these results to better housing conditions that provide higher quality and a more stable environment so that children are able to advance in their educational path. The higher quality of homes in which the owners reside is attributed to the propensity of these owners to take better care of the internal facilities when compared to renting. Mohanty and Raut (2009) find no direct impact of home ownership on educational achievement but conclude that it creates a better home environment, which has a positive effect on children school outcomes. These conclusions are based on data from the Panel Study of Income Dynamics (PSID) Child Development Supplement for the USA. Using data from the National Longitudinal Survey of Youth, USA, Blau et al. (2019) also find a positive association between owner-occupied home during childhood and young adults' educational attainment. Furthermore, the positive effects of homeownership are found to go beyond education, being positively associated with employment and negatively associated with teen pregnancy, criminal convictions and the likelihood of being on welfare. Leviten-Reid & Matthew (2018) also confirm the importance of homeownership for bonding social capital availability in Canada, although other factors such as residential stability exert a bigger effect on all forms of social capital. Closer to the goals of our analysis, Simson & Umblijs (2020) investigate the relationship between home and neighbourhood environment and the educational performance of pupils. Using microdata from Norway, the authors found that factors such as noise pollution, overcrowded homes, lack of homeownership and housing stability (moving frequently) are related to lower test scores.

In summary, housing conditions have been portrayed as exerting a positive influence on health and educational outcomes in single country studies (the exception is Bradley & Putnick (2012)), but the literature lacks a comprehensive empirical analysis covering a wider sample of countries considering human capital measured in a way that may be more useful for economic growth analyses. This approach also adds to the housing literature, allowing it to expand from its usual focus on the impact of housing on short-run economic performance and providing insights on its potential role for economic growth.

EMPIRICAL STRATEGY AND DATA

The empirical analysis considers a sample composed of low- and middle-income countries, based on the World Bank income classification groups. We exclude high-income countries from the analysis due to the small variation in housing quality in this group of countries. The final sample comprises 52 low and middle-income countries for which data on housing quality was available (for the complete list of countries included in the analysis see Table A.1 in the appendix). All the estimations were carried out with the econometric package GRETL (Gnu Regression Econometrics and Time-Series Library) version 2019b.

The baseline empirical model estimated is given by equation *(1)*:

$$lnH_i = \alpha + \beta lnQ_i + \lambda' lnX_i + u_i \qquad (1)$$

where the dependent variable, H, is human capital for country i and the explanatory variable of interest is Q, (lack of) housing quality in country i. The model additionally includes a vector X of control variables with other determinants of human capital formation selected based on previous empirical literature (Baldacci et al., 2008). α is the constant term and u the error term. The variables included in vector X are *GDP* that corresponds to real income per capita, *gov_edu*, that represents state intervention at the educational level, *Mortality* that corresponds to the health status of the population and *Internet*, the proportion of the population with access to the internet. These control variables were selected based on the work of Baldacci et al. (2008) who estimate a regression to predict educational outcomes in 118 developing countries over the period 1971-2000 based on a set of explanatory variables (e.g., population's health, expenditure in education, urbanization and gender equality). Our choice of explanatory variables was also dictated by data availability issues and the need to define a parsimonious empirical model due to the limited number of observations available. We first had to guarantee that we had data for our explanatory variable of interest, housing quality, and the dependent variable, human capital, and the choice of the remaining control variables implied that they had to be available for the sample defined by the previous variables. Table A.2 in the appendix identifies the variables used, describes how they are measured and identifies the sources of the data.

We measure human capital, H; as average years of schooling of the population aged between 25 and 74 years old taken from the UNESCO Institute for Statistics. This proxy is widely used as a measure of human capital in empirical growth studies[2] and one of the main purposes of our analysis is to investigate the relationship between housing quality and human capital availability to reflect on the role of the former as a potential determinant of economic growth, with human capital as the mechanism of transmission. Other often used proxies for educational human capital include enrolment rates, for measuring quantity of schooling, and internationally comparable test scores, for measuring the quality of schooling as in Hanushek & Woessmann (2011). The choice of average years of schooling was based on its wider availability for developing countries and is in line with applied economic growth studies such as Benhabib & Spiegel (1994). One problem we encountered was the matching of the cross-sectional data for the human capital stock and housing quality for some countries and years. To address this problem, we used the Barro-Lee dataset to fill the gaps for countries for which there was no data in the UNESCO database. Although the Barro-Lee dataset computes average years of schooling based on the highest education level attained by individuals aged 15-64 years old (not 25-74), we believe that this approach does not meaningfully influence the results. In fact, when we

[2] For some very recent examples of the vast number of studies that consider the relationship between human capital, measured as average years of schooling, and economic growth see e.g. Cornell, Knutsen, & Teorell (2020), Gutiérrez-Romero (2021), Köppe Malanski & Póvoa (2021), Matousek & Tzeremes (2021), Osei & Kim (2020), Ouedraogo, Sourouema & Sawadogo (2021), Sturn & Epstein (2021), Woo (2020) and Zergawu, Walle & Giménez-Gómez (2020).

compared the two datasets for the countries for which we have data in both datasets, for the same year, we concluded that the values were quite similar. Of the 52 countries considered, we used the Barro-Lee data to fill the gaps for 8 countries, or 15,38% of our sample (see Table A.3 in the Appendix).

Our explanatory variable of interest, Q, is the proportion of the population that lives in sub-standard housing. The use of this variable dictated the structure of the data used in the empirical analysis. In fact, the cross-section approach was chosen due to data limitations associated with the housing quality indicators, where for each country only one data point was available, corresponding to a single year. The year to which each observation refers to was also usually different across a large number of countries. To measure the lack of housing quality (sub-standard housing) we consider the number of occupants of housing units, according to different housing types, retrieved from the UNESCO Institute for Statistics. We chose this indicator due to its comparability across countries, covering ten standardized types of housing. This homogenous international classification allows us to compare different countries despite the large variability in housing standards between different countries, usually dictated by the availability of building materials among other factors. The housing data is divided into several categories corresponding to different housing quality types. Table A.4 and Figure A.1 in the Appendix summarize the different housing quality classifications used by the United Nations and present the respective definitions. Good quality housing according to the description in the database corresponds to a common dwelling with all the basic facilities. According to the UN's Principles and Recommendations for Population and Housing Censuses, a common dwelling has four essential features: it is composed by a room or suite of rooms, it is located in a permanent building, it has a separate access to a street or common space and was intended for occupation by a single household (UN, 2017, p.249). Furthermore, the UN also defines basic facilities for decent living: piped water, a toilet, fixed bath or shower, a kitchen or other space for cooking, with all four located within the same dwelling. All other categories of housing fail to meet the former criteria and so we dub them sub-standard housing.

Due to some inconsistency in the observations for different housing categories we cannot include each separately in the regressions. To overcome this problem, we computed a new variable that considers the population living in any of the housing categories considered to be sub-standard divided by the total population to take into consideration different population sizes, as can be seen in equation *(2)*.

$$Q = \left(\frac{Population\ in\ sub-standard\ housing}{Total\ population}\right) * 100 \qquad (2)$$

As far as the expected sign of the different estimated coefficients is concerned, we expect a negative relationship between lack of housing quality and human capital, with higher shares of the population living in sub-standard housing (higher Q) associated with lower human capital formation because lower housing quality may act as a disincentive for individuals to pursuit more education due to the lack of a study enhancing environment at home or negative health effects caused by low housing quality. Income per capita is expected to have a positive influence on human capital since higher income raises the ability of individuals to afford more education since its relative cost becomes lower as income increases (Baldacci et al. 2008). The same positive influence applies to state intervention at the educational level that gives broader access to the education system and probably allows for poorer, but talented, individuals to acquire skills and competences that would otherwise be unattainable. A less healthy population, proxied by the infant mortality rate, is expected to have a negative influence on human capital because it may act as a barrier for individuals to be able to afford investing in education since the individual's poor health status can act as a disincentivise for school attendance, can lower learning ability or even induce dropping out of school altogether. Finally, the percentage of the population with access to the internet is expected to have a positive influence on human capital formation as it is a tool that helps individuals in the education process through online materials, useful when doing homework and studying, see e.g. Lei & Zhao (2007) and Sanchis-Guarner, Montalbán & Weinhardt (2021).

Table 1 contains some descriptive statistics for the variables of interest, lack of housing quality and human capital (see Table A.5 in the appendix for the descriptive statistics for the control variables). At first glance, it seems the data for both variables shows enough variation across countries in order to allow for the identification of a relationship between the two variables. Indeed, the minimum and maximum values are located apart from each other, indicating a high variation in the dataset. This characteristic is also supported by the high standard deviation, in particular for the lack of housing quality variable. For human capital, the standard deviation is not very high, but this is to be expected since the sample is comprised of low and middle-income countries only, which tend to be associated with similar low levels of education. The high variability of the lack of housing quality variable holds true also when comparing directly to human capital, with the standard deviation of the former being higher than that of the latter.

Looking at the median and the mean, neither of the variables follows a normal distribution, with housing quality having a positive skew and human capital with a negative skew. The positive skew in the lack of housing quality variable is especially worrying as its skewness value is very high (5.2486). Furthermore, the coefficient of correlation between the two variables is negative (-0.1179) but not statistically significant. This negative correlation indicates that the relationship we expect to find is confirmed by the data. However, the correlation coefficient for the same variables in logs changes in sign (+0.1052), although it remains insignificant.

Table 1. Descriptive Statistics for the lack of housing quality and human capital variables

Variable	Mean	Median	Min	Max	Std Dev	C.V
Q	4.8489	2.1928	0.1912	73.07	10.950	2.2583
H	8.4023	8.7687	1.9193	12.632	2.6627	0.3169
Ln_Q	0.8615	0.7851	-1.6540	4.2915	1.0226	1.1869
Ln_H	2.0587	2.1706	0.6519	2.5362	0.4188	0.2034

Notes: Q is the share of the population that lives in substandard quality housing. H is average years of schooling.

Source: authors' own calculations using the econometric package Gretl

RESULTS

The results from the OLS estimation of the baseline equation (1) are presented in Table 2. We present the results for four distinct regressions corresponding to different model specifications depending on the set of control variables considered in order to check the robustness of the results to different combinations of the control variables. We eliminated the control variables according to its importance to the explanation of differences in educational attainment based on the findings of previous empirical literature (e.g. Baldacci et al. (2008)) or due to its lack of statistical significance. We first leave out the variables that are less consensual as determinants of human capital formation, such as access to the internet, up to the most parsimonious model that considers only GDP per capita as a control variable, according to the relevance attributed to these variables by Baldacci et al. (2008). Column (1), Table 2, contains the results considering all control variables; column (2) leaves out internet access; column (3) additionally leaves out mortality; and, finally, column (4) also does not consider government spending on education.

Table 2. Results with OLS

	(1)	(2)	(3)	(4)
const	1.4773* (0.8470)	0.9789 (0.7937)	−1.0653** (0.4323)	−0.9320** (0.3918)
Ln_Q	0.0806** (0.0379)	0.0885** (0.0381)	0.0821* (0.0411)	0.0743* (0.0396)
Ln_GDP	0.1211 (0.0749)	0.1848*** (0.0633)	0.3306*** (0.0435)	0.3288*** (0.0432)
Ln_gov_edu	−0.0219 (0.0982)	0.0036 (0.0982)	0.0765 (0.1026)	
Ln_mortality	−0.2039*** (0.0664)	−0.2010*** (0.0673)		
Ln_Internet	0.0523 (0.0340)			
Countries	52	52	52	52
Adjusted R-squared	0.6025	0.5910	0.5237	0.5280
p-value(F-stats)	2.71e-09	1.72e-09	1.84e-08	3.85e-09
Akaike criterion	14.7156	15.3191	22.3365	20.9351
p-Value (Breusch-Pagan)	0.0010	0.0003	0.0010	0.0002

*Notes: standard error in parenthesis. ***; **; * indicate statistical significance at the 1%, 5% and 10% level, respectively. Heteroskedasticity-robust standard errors, variant HC1.*

Source: authors' own calculation using the econometric package Gretl

According to the results presented in Table 2, in all of the estimated models we obtain a positive and statistically significant relationship between lack of housing quality and human capital, at either 5 or 10% significance levels (but never at the 1% level). From model (1) to model (4) the coefficient for lack of housing quality remains basically unchanged, ranging from 0.07 and 0.08, indicating that if this variable increases by 10% the human capital stock will increase by 0.7-0.8 percentage points, depending on the model. This positive relationship implies that a country that has a larger share of its population living in sub-standard housing also has available higher levels of human capital, a result that goes against our initial expectations.

As for the control variables, the estimated coefficient for GDP per capita has the expected positive sign in all models, with statistical significance at the 1% level for models (2) to (4), confirming the prediction that countries with higher levels of income per capita are also the ones with higher average years of schooling. The coefficient for public spending in education appears with a negative sign in model (1), contrary to our expectations when considering the results from the work of Baldacci et al. (2008), implying that the more governments spend on education, the less human capital stock is available. This could indicate that higher public spending on education results in less efficiency in terms of resource allocation. However, the former coefficient is not statistically significant and turns positive in models (2) and (3), when the variables for the health status and access to the internet are removed from the regression. Again, none of the coefficients is statistically significant and the estimated coefficient is relatively low. It thus seems that state intervention in the education system has not had a significant impact on human capital formation in developing countries. The result for the health status of the population is in line with initial predictions, presenting a large negative estimated coefficient, corresponding to a negative elasticity of 2 percentage points for models (1) and (2), and statistically significant at the 1% level. This is in line with Baldacci et al. (2008), which concluded that countries that have a population with better health have a higher amount of human capital available. The estimated impact of internet access on human capital is positive in sign but not statistically significant, which might indicate that having access to information does not provide enough aid in the personal educational process, as oppose to income for example.

Overall, when looking at the adjusted R-squared for model (1), we can see that the model explains 60.2% of the change in the dependent variable. Considering the relatively small number of explanatory variables and observations, we can say that the model provides a satisfactory prediction ability. For the F-test's p-value, it is possible to reject the null hypothesis of the test for all the models which means that the coefficients obtained have more explanatory power than if the model had no explanatory variables, i.e., an intercept-only model. Comparing the performance of the models by the Akaike-information-criteria, in which lower values indicates a higher predictive ability of the model, we observe that the best model, with the lowest value for the Akaike criteria, is model (1). However, when we apply the Breusch-Pagan test, for which the null hypothesis is that of homoscedastic errors, the p-value is always lower than 0.01, indicating that all the models suffer from heteroskedasticity. This indicates that the regression results can be biased, which is caused by the omission of an unknown variable, and so the results we obtained are not robust.

Table 3. Results with OLS and robust standard errors

	(1)	**(2)**	**(3)**	**(4)**
const	1.4773* (0.7909)	0.9789 (0.7681)	−1.0653* (0.5358)	−0.9320* (0.5118)
Ln_Q	0.0806 (0.0501)	0.0885* (0.0483)	0.0821 (0.0543)	0.0743 (0.0493)
Ln_GDP	0.1211 (0.07220)	0.1848*** (0.0657)	0.3306*** (0.0522)	0.3288*** (0.0524)
Ln_gov_edu	−0.0219 (0.1032)	0.0036 (0.1053)	0.0765 (0.1120)	
Ln_mortality	−0.2039*** (0.0533)	−0.2010*** (0.0540)		
Ln_Internet	0.0523* (0.0299)			
Countries	52	52	52	52
R-squared	0.6415	0.6230	0.5517	0.5465
Adjusted R-squared	0.6025	0.5910	0.5237	0.5280
P-value(F)	1.15e-08	4.26e-09	7.44e-07	3.96e-07
Akaike criterion	14.7156	15.3191	22.3365	20.9351
P-value (Breusch-Pagan)	0.0010	0.0003	0.0010	0.0002

*Notes: standard error in parenthesis. ***; **; * indicate statistical significance at the 1%, 5% and 10% level, respectively. Heteroskedasticity-robust standard errors, variant HC1.*

Source: authors' own calculations using the econometric package Gretl

To address the problem of heteroskedasticity we estimated equation (1) correcting for this problem. The results can be found in Table 3 where we ran the same models but now considering robust standard errors, in which heteroskedasticity is eliminated from the calculation of the matrix of variances-covariances. It is important to notice that this procedure does not eliminate the problem of heteroskedasticity from the regressions, but considers robust standard errors, making statistical inference possible while maintaining the same value and sign of the coefficients as in the former estimations but potentially changing its statistical significance, i.e., standard errors and t-statistics. As can be seen in Table 3 from the p-value of the Breusch-Pagan test, the heteroskedasticity problem remains after the inclusion of robust standard errors specification and, thus, we continue to have the problem of omitted variable bias. As for the estimated coefficients, although the estimated coefficient for lack of housing quality is still positive, the respective statistical significance changed considerably, since it is only significant in model (2) and only at the 10% level. Overall, these results suggest that housing quality is not an important determinant of human capital availability in developing countries, contrary to our initial expectations. The results for the control variables indicate now that the relevant determinants of human capital availability are the mortality rate and GDP per capita. According to the adjusted R-squared and the Akaike criterion, model (1) in Table 3 outperforms the other models. Also, all the models still managed to reject the null hypothesis of the F-test. Notice one interesting change in the results for the control variables relative to the ones in Table 2: now access of the population to the internet presents the positive expected sign and is also statistically significant, in line with our expectations that the ability to access the large pool of useful information online, such as educational materials, can affect positively the accumulation of human capital.

Given the poor performance of the coefficient for the housing quality variable, we moved on to test for different hypothesis always considering robust standard errors. In the previous estimations we considered the whole sample of countries corresponding to low and middle-income countries according to the World Bank classification. This implies considering a set of countries still with quite distinct realities, since the sample includes low income, lower-middle income and upper-middle income countries. Therefore, it is important to consider the possibility of a difference in the behaviour of human capital in relation to housing quality for these distinct levels of income. To address this possibility, we estimated the model with interactions terms between dummies for each of the three levels of income interacted with the lack of housing quality variable, where the dummy variables are defined as:

$$dummy_{low} = 1, for\ i = low\ income\ country, and\ 0\ otherwise;$$

$$dummy_{middle_L} = 1, for\ i = lower - middle\ income\ country, and\ 0\ otherwise;$$

$$dummy_{middle_H} = 1, for\ i = upper - middle\ income\ country, and\ 0\ otherwise,$$

This specification allows us to investigate if the relationship between lack of housing quality and human capital availability differs according to income levels. The former difference would correspond to different estimated coefficients for each of the interaction terms where we could have also different signs and statistically significance. The results of the regressions with interaction terms can be seen in Table 4. The number of countries included in each country group is discriminated in the notes to the table.

Table 4. OLS regressions with interaction terms for income groups

	(1)	**(2)**	**(3)**	**(4)**
const	1.1363 (0.8975)	0.5835 (0.8307)	−1.3747** (0.6198)	−1.2629** (0.6058)
lnQ*dummy_low	0.1551 (0.1026)	0.1525 (0.1008)	0.1742 (0.1097)	0.1719 (0.1097)
lnQ*dummy_middle_L	0.0552 (0.0595)	0.0751 (0.0584)	0.0600 (0.0680)	0.0504 (0.0561)
lnQ*dummy_middle_H	0.0364 (0.0400)	0.0371 (0.0393)	0.0169 (0.0428)	0.0135 (0.0408)
Ln_GDP	0.1549* (0.0851)	0.2259*** (0.0739)	0.3701*** (0.0642)	0.3681*** (0.0649)
Ln_gov_edu	−0.0337 (0.1088)	0.0002 (0.1098)	0.0615 (0.1190)	
Ln_mortality	−0.1857*** (0.0545)	−0.1855*** (0.0533)		
Ln_Internet	0.0585* (0.0305)			
Total countries	52	52	52	52
R-squared	0.6573	0.6353	0.5768	0.5735

	(1)	(2)	(3)	(4)
Adjusted R-squared	0.6028	0.5867	0.5308	0.5373
P-value(F)	1.08e-07	1.88e-08	1.12e-06	2.47e-06
Akaike criterion	16.3651	17.5961	23.3398	21.7360

*Notes: standard error in parenthesis. ***; **; * indicate statistical significance at the 1%, 5% and 10% level, respectively. Heteroskedasticity-robust standard errors, variant HC1. For the dummy variables, we have in dummy_low 12 countries, in dummy_middle_L 15 countries and in dummy_middle_H 27 countries.*

Source: authors' own calculations using the econometric package Gretl

As can be seen in Table 4, all four models reveal again a positive association between lack of housing quality and human capital in the three country groups under analysis based on the estimated coefficients of the three interaction terms, although none is statistically significant. For the control variables, we left them out of the regression in the same order as used in Tables 2 and 3. Public expenditure on education still presents no statistical significance, indicating that this variable is not a relevant determinant of in human capital formation. Infant mortality maintains its statistical significance at the 1% level as in the previous tables. The sign of the estimated coefficients for GDP per capita and access to the internet remain basically the same. Concerning the overall performance of the regressions, all models reject the null hypothesis of the F-test. Model (1) records the best performance based on the results of the Akaike criterion and the adjusted R-squared.

Finally, we also tested in a different way for the possibility of non-linearities in the relationship between housing quality and human capital formation by considering that the response of human capital to housing quality might correspond to an inverted U: for small levels of Q an increase in the former variable leads to an increase in human capital but, beyond a certain threshold, the relationship becomes negative. In line with the regressions in Table 4 that assume a different response of human capital to housing quality depending on the level of income of countries, we also believe that the influence of housing quality on human capital can have different responses depending on the intensity of the former, according to a quadratic function. In practical terms, an inverted U would mean that, beyond the maximum point of the function, the ratio of the population living in sub-standard housing would become too detrimental to the well-being of the population, impacting negatively the ability to attend school. However, this inverted U also implies that for lower values in the housing quality ratio, the relationship is positive, meaning that, until the maximum is reached, having a portion of the population living in sub-standard housing is actually positively correlated with human capital. This could apply if lower housing quality functioned as an incentive for individuals to search for more education to access better paying jobs and later improve their housing quality. To test this hypothesis, we ran the regressions including additionally the square of the housing quality variable.

Table 5. Results with a quadratic term for lack of housing quality

	(1)	(2)	(3)	(4)
$const$	1.7311* (0.8684)	1.1181 (0.8168)	−0.7736 (0.5283)	−0.7375 (0.5292)
Ln_Q	0.1618* (0.0894)	0.1563* (0.0919)	0.1661* (0.0987)	0.1677 (0.0985)
$(Ln\_Q)^2$	−0.0362 (0.0272)	−0.0297 (0.0280)	−0.0366 (0.0987)	−0.0379 (0.0293)
Ln_GDP	0.0951 (0.0811)	0.1723** (0.0715)	0.3067*** (0.0561)	0.3055 (0.0570)
Ln_gov_edu	−0.0811 (0.1072)	−0.0413 (0.1066)	0.0167 (0.1084)	

	(1)	(2)	(3)	(4)
Ln_mortality	−0.1927*** (0.0496)	−0.1914*** (0.0504)		
Ln_Internet	0.0612* (0.0321)			
Countries	52	52	52	52
R-squared	0.6613	0.6368	0.5728	0.5726
Adjusted R-squared	0.6162	0.5973	0.5365	0.5459
P-value(F)	3.65e-08	1.10e-08	1.43e-06	4.87e-07
Akaike criterion	13.7499	15.3844	21.8217	19.8480

*Notes: standard error in parenthesis. ***; **; * indicate statistical significance at the 1%, 5% and 10% level, respectively. Heteroskedasticity-robust standard errors, variant HC1.*

Source: authors' own calculations using the econometric package Gretl

The results considering a quadratic relationship are presented in Table 5. Each column considers different sets of control variables selected according to the strategy described for Table 2. Overall, the results do not support the existence of a non-linear relationship in any of the four models estimated since the estimated coefficient for the square of housing quality, although negative, is never statistically significant. In any case, the estimated coefficient for the linear term of housing quality is positive and statistically significant at the 10% level in all models except model 4 and, based on the p-value for the F statistic, we confirm the joint significance of the variables in the models. Model (1) presents the highest adjusted R-squared and the lowest value for the Akaike information criterion and so is our preferred model. By deriving human capital relative to housing quality in model (1) and equalizing it to zero (see equation (3)) we can compute the maximum of the function, i.e., the value of Q beyond which the relationship becomes negative:

$$\left(\frac{dlnH}{dlnQ} = 0 \leq=> \ \beta_1 + 2\beta_2 Ln\_Q = 0\right) \qquad (3)$$

The maximum is located at $Ln\_Q = 2.667$. This turning point implies that countries with values of lnQ higher than this maximum record a negative relationship between Q and H, as expected. The maximum for lnQ implies a value for Q of 14.39%. However, there is a small number of countries in our dataset that record housing quality ratios higher than the threshold and so are located in the part of the curve where the relationship is positive. This means that, for the majority of countries under analysis, the positive relationship that the previous linear models in Tables 2 to 4 indicated still holds true. As for the control variables, the results remain basically unchanged when compared to the results in Table 2.

DISCUSSION

In this section we will discuss some issues that might explain the inability of our results to confirm our initial hypothesis, that lack of housing quality hinders human capital formation, and the potential implications of the results found for structural policies that promote economic growth in developing countries.

Regarding the presence of heteroskedasticity in all the estimated models, we believe that this problem might be due omitted variables in the model specification, a consequence of limited data availability for our sample of developing countries. It is possible to find in the previous literature some potential candidates for these missing explanatory variables. For instance, Hanushek & Woessmann (2011) propose a production function approach for the estimation of the quality of human capital that includes other variables not considered in the present study. The empirical model proposed by the authors is given by equation *(5)*:

$$Human\ capital = \beta_1 family\ inputs + \beta_2 schooling\ inputs + \beta_3 individual\ ability + \beta_4 other\ factors + \epsilon \quad (5)$$ [3]

As it is possible to see in equation (5), this model differs from ours mainly because it considers factors that impact the human capital of individuals at the micro level where family inputs and individual abilities define the context in which an individual develops its cognitive abilities to absorb the knowledge provided by the educational system. However, in the present study we are not able to access data on the quality of human capital for our sample of countries.

Lee & Barro (2001) carry out an empirical analysis of schooling quality in a cross-section of countries considering a similar production function, summarized in equation *(6)*:

$$Q_{ijt} = \alpha_{ijt} + \beta_1 F_t + \beta_1 R_t + \varepsilon_{ijt} \quad (6)$$

where Q stands for individual tests scores, F includes family factors, such as parents' income and educational attainment, which affects the probability that children enrol in, attend and complete school, but also the ability of a child to learn. R stands for school resources, such as pupil-teacher ratios, average teacher salary, educational expenditure per pupil and school duration, with all these factors influencing the ability of the schooling system to provide a good quality environment for learning. The authors found that family background and school resources have a strong positive association with student performance. This positive association is also supported by the results found in Égert, Botev and Turner (2020).

The inputs of the human capital production function in Lee & Barro (2001) gives us a clue for one of the reasons why our analysis probably does not allow us to reach robust results. Even though we included inputs such as public expenditure in education and GDP per capita, which Lee & Barro (2001) considered as a proxy for parental income, we left out other potentially important determinants of human capital, which explains our need to perform the Breusch-Pagan heteroskedasticity test in our original models (Table 2). The test results indicate that the variance of the error term is not white-noise, indicating that the models suffer from misspecification, thus limiting the ability of the control variables to isolate the effect of our explanatory variable and limiting its predicting ability. As much as we would like to fill this gap in the model's specification, the unavailability of data for the countries under analysis, given that, being developing countries, most suffer from limited data collection at the national level and data processing by the national statistical agencies, did not enable us to define comprehensive model specifications. Moreover, in Lee & Barro (2001) the authors analyse the quality of human capital and not the quantity, as we do in our study. This difference in measurement might also give a clue for the lack of meaningful results in our regressions, as housing quality probably impacts more human capital quality than quantity. However, we were not able incorporate human capital quality in our analysis as internationally comparable data on student's performance is not available for most of the countries in our sample.

Other limitations that might influence the performance of the models can be found in the housing quality data. Housing quality data is scarce in the UN's database, resulting in a maximum of 52 developing countries with which we could work with since we also had to guarantee that we had human capital data for those same countries. Additionally, the data collected in the UN's database lacks a periodic time frame and so most of the countries in upper middle income to low middle income groups have only one-time observation. This lack of observations over time limits the estimations methodologies that can be applied to correct certain issues, which would become possible with a panel data structure. The lack of observations is not only due to lack of time coverage. The number of countries for which data is available is higher than 52, but we had to reduce this number due to lack of data for average years of schooling for some countries.

---

[3] Hanushek & Woessmann (2011, p.433)

Besides the small number of observations, we also encountered other problems with the housing quality variable which had to be computed as an aggregate of different classifications/categories of housing quality. This was done to overcome the problem of having different coverage of housing quality categories across countries and enabled us to reach a comprehensive measure that represents the overall problems in access to decent housing for each country. However, as we saw during the construction of the variable, it appears that the lack of information in various categories, which was represented as a zero in the spreadsheet provided in the UNESCO dataset, is improbable, leading us to believe that the dataset suffered from a problem of poor data collection. If this applies then the indicator that we constructed does not accurately represent the proportion of individuals living in sub-standard housing, making it harder to find a robust relationship with human capital. Although this problem might seem to impede an econometric analysis, we went ahead with the study due to the possibility that the inclusion of control variables that are known to be measured with a good degree of precision might help to isolate the housing indicator effects even if its accuracy is not as high as that of the remaining variables.

Additionally, we can also be in the presence of endogeneity with human capital availability influencing access to quality housing. This problem could in principle be solved using instrumental variables estimators, but we could not find good instruments and even if we had it would be unlikely the respective availability would match our limited dataset.

The previously discussed shortcomings thus represent hurdles to the interpretation of our findings leading us to doubt the positive relationship between the lack of housing quality and human capital obtained. This conclusion is further supported by the lack of statistical significance of the estimated coefficient for the variable of interest in the models considering robust standard errors and the fact that even when the coefficient was statistically significant this happened only at a 10% level and in a very small subset of the regressions using robust standard errors.

Overall, if new datasets emerge in the future, both for housing quality and human capital quality in the context of developing economies, a new version of this analysis can be done which could not only answer our initial question, but also enable a research on the impact that housing quality has on economic growth. This could be done much in the same ways as in the study by Égert, Botev and Turner (2020), in which the authors estimate the impact of educational policies on human capital and, consequently, on economic growth.

CONCLUSION

The main objective of this work was to empirically investigate the relationship between housing quality and human capital formation in the context of developing countries. For this purpose, we performed a cross-sectional analysis estimating different model specifications, considering different control variables, robust standard errors and the possibility of non-linearities in the relationship, according to countries' income levels or described by a quadratic relationship.

Our findings suggest that, overall, there is no statistically significant association between the lack of housing quality and human capital although the respective estimated coefficient is positive, contrary to theoretical prediction. However, we believe that this is likely due to the quality of the housing data that results in measurement error. Additionally, the small number of observations for developing countries under analysis leads to inconclusive results.

From a policy design perspective, our results do not endorse investing in access to better quality housing as a means to increase human capital formation in developing countries. As we were not able to obtain conclusive results for the relationship under analysis, the study also could not support access to better housing quality as a factor that impacts long-run economic growth through its interaction with higher levels of human capital, along with other factors, such as investment in physical capital and technology, or even the health status of a country, that has consistently been shown to impact human capital negatively.

Although the inconclusive nature of the results did not allow us to bridge the gap identified in the literature concerning empirical analysis of the relationship between housing and human capital, we believe that a major contribution of this work lies in raising awareness to the need for better data for housing quality and human capital to enable for future empirical analyses that e.g. consider also human capital quality not just quantity. Considering the rapid rate of urbanization that has been taking place since the end of the 20th century in developing countries, future research on this topic can help produce helpful guidelines for the design of effective and efficient policies directed at housing that can additionally promote growth and development.

ACKNOWLEDGMENTS

We are grateful to the anonymous reviewers for their time and effort reading our manuscript and for the constructive input on an earlier version of the text.

REFERENCES

Baldacci, E., Clements, B., Gupta, S., & Cui, Q. 2008. "Social spending, human capital, and growth in developing countries." *World development*, 36(8): 1317-1341.

Barro, Robert and Jong-Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104:184-198.

Benhabib, J., & Spiegel, M. M. 1994. "The role of human capital in economic development evidence from aggregate cross-country data." *Journal of Monetary economics*, 34(2): 143-173.

Blau, David M., Haskell, Nancy L., Haurin, Donald R. 2019. "Are Housing Characteristics Experienced by Children Associated with their Outcomes as Young Adults?" *Journal of Housing Economics,* 46.

Bradley, R. H., & Putnick, D. L. 2012. "Housing quality and access to material and learning resources within the home environment in developing countries." *Journal of Child development*, 83(1): 76-91.

Bramley, Glen & Karley, Noah K. 2007. "Homeownership, Poverty and Educational Achievement: School Effects as Neighbourhood Effects". *Housing Studies*, 22(5): 693-721.

Cornell, A., Knutsen, C.H. & Teorell, J. 2020. "Bureaucracy and Growth." Comparative Political Studies, 53(14): 2246-2282.

Égert, B., Botev, J., & Turner, D. 2020. "The contribution of human capital and its policies to per capita income in Europe and the OECD". *European Economic Review*, 129: 1-16

Green, R. K. 1997. "Follow the leader: how changes in residential and non-residential investment predict changes in GDP". *Real Estate Economics*, 25(2): 253-270.

Gutiérrez-Romero, R. 2021. "How does inequality affect long-run growth? Cross-industry, cross-country evidence." Economic Modelling, 95: 274-297.

Hanushek, E. A., & Woessmann, L. 2011. "How much do educational outcomes matter in OECD countries?" *Economic policy*, 26(67): 427-491.

Harris, R., & Arku, G. 2006. "Housing and economic development: The evolution of an idea since 1945." *Habitat International*, 30(4): 1007-1017.

Healy, R. G. 1971. "Effects of improved housing on worker performance." *Journal of Human Resources,* 6(3): 297-308.

Hongyu, L., Park, Y. W., & Siqi, Z. 2002. "The interaction between housing investment and economic growth in China". *International real estate review*, 5(1): 40-60.

Housing, C., & Council, P. 2001. "Housing and schooling." *The Urban Prospect*, *7*(2): 1-4.

Köppe Malanski, L. & Póvoa, A. C. S. 2021. "Economic growth and corruption in emerging markets: Does economic freedom matter?" International Economics, 166: 58-70.

Lee, J. W., & Barro, R. J. 2001. "Schooling quality in a cross–section of countries." *Economica*, 68(272): 465-488.

Lei, J., & Zhao, Y. 2007. "Technology uses and student achievement: A longitudinal study." *Computers & Education*, 49(2): 284-296.

Leviten-Reid, C. & Matthew, R. A. 2018. "Housing Tenure and Neighbourhood Social Capital." *Housing, Theory and Society*, 35(3): 300-328.

Matousek, R. & Tzeremes, N.G. 2021. "The asymmetric impact of human capital on economic growth." Empirical Economics, 60(3): 1309-1334.

Lucas Jr, R. E. 1988. "On the mechanics of economic development." *Journal of monetary economics*, 22(1): 3-42.

Mankiw, N. G., Romer, D., & Weil, D. N. 1992. "A contribution to the empirics of economic growth." *The Quarterly Journal of Economics*, 107(2): 407-437.

Manzoli, V.; Duarte, A.P.; Simões, M. 2020. "Housing, Inequality and Economic Growth: Evidence from a Sample of Brazilian States". *Economic Analysis* 53(1), pp. 42-58.

Mohanty, L.L. & Raut, L.K. 2009. "Home Ownership and School Outcomes of Children: Evidence from the PSID Child Development Supplement." American Journal of Economics and Sociology, 68: 465-489.

Osei, M.J. & Kim, J. 2020. "Foreign direct investment and economic growth: Is more financial development better?" Economic Modelling, 93: 154-161.

Ouedraogo, R., Sourouema, W.S. & Sawadogo, H. 2021. "Aid, growth and institutions in Sub-Saharan Africa: New insights using a multiple growth regime approach." World Economy, 44(1): 107-142.

Sanchis-Guarner, R., Montalbán, J., & Weinhardt, F. 2021. "Home Broadband and Human Capital Formation." *CESifo Working Paper* 8846

Solow, R. M. 1956. "A contribution to the theory of economic growth." *The quarterly journal of economics*, *70*(1): 65-94.

Sturn, S. & Epstein, G. 2021. "How much should we trust five-year averaging to purge business cycle effects? A reassessment of the finance-growth and capital accumulation-unemployment nexus." Economic Modelling, 96: 242-256.

Terzi, Fatih & Bölen, Fulin. 2008. "An Analysis of the Relationship between Housing and Economic Development", presented at 48th European Congress of the Regional Science Association International, Liverpool, UK.

UN. 2017. *Principles and Recommendations for Population and Housing Censuses, Revision 3*. UN.

UNESCO Institute for Statistics. 2019. Available from: http://data.un.org/Data.aspx?d=POP&f=tableCode%3a309

Von Simson, Kristine von & Umblijs, Janis. 2020. "Housing conditions and children's school results: evidence from Norwegian register data" *International Journal of Housing Policy,* 1-26.

Woo, J. 2020. "Inequality, redistribution, and growth: new evidence on the trade-off between equality and efficiency." Empirical Economics, 58(6): 2667-2707.

Zergawu, Y.Z., Walle, Y.M. & Giménez-Gómez, J.-M. 2020. "The joint impact of infrastructure and institutions on economic growth." Journal of Institutional Economics, 16(4): 481-502.

APPENDIX

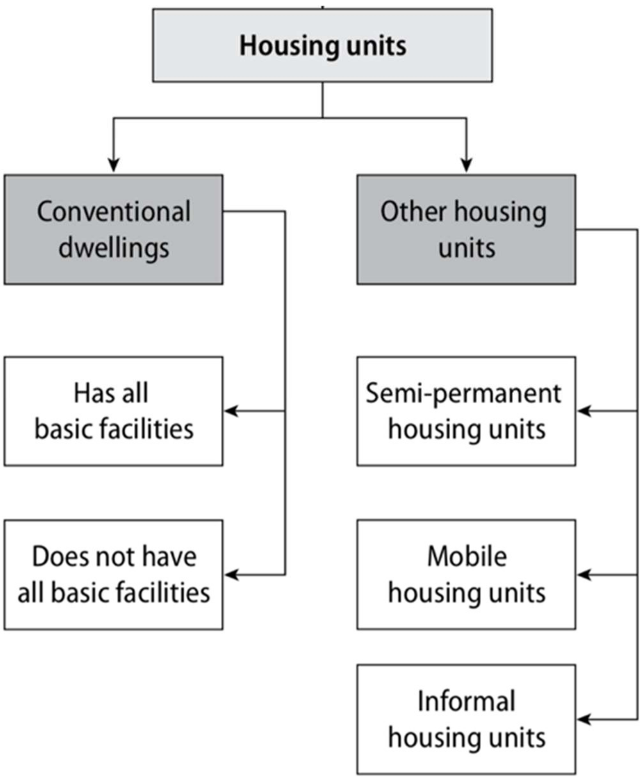

Figure A.1. Classification of housing units

Source: UN (2017), p. 250

Table A.1. List of the 52 countries included in the econometric analysis

Country	Year of the data used	Income Classification	Country	Year of the data used	Income Classification
Albania	2011	M_L	Latvia	2011	M_H
Argentina	2010	M_H	Lesotho	2006	M_L
Armenia	2011	M_L	Liberia	2008	L
Azerbaijan	2009	M_H	Malawi	2008	L
Belarus	2009	M_H	Malta	1995	M_H
Bolivia	2012	M_L	Malaysia	2010	M_H
Brazil	2010	M_H	Mexico	2010	M_H
Bulgaria	2011	M_H	Morocco	2004	M_L
Chile	2002	M_H	Myanmar	2014	M_L
Costa Rica	2011	M_H	Nicaragua	2005	M_L
Croatia	2001	M_H	Peru	2007	M_L
Cuba	2002	M_L	Philippines	2000	M_H
Dominican Republic	2002	M_L	Poland	2002	M_H
Ecuador	2010	M_H	Republic of Moldova	2004	L
Ethiopia	2007	L	Romania	2011	M_H
Egypt	2006	M_L	Russian Federation	2010	M_H
Georgia	2002	L	Rwanda	2012	L
Ghana	2010	M_L	Saint Lucia	2010	M_H
Guinea	2014	L	Serbia	2011	M_H
Hungary	2001	M_H	Slovakia	2001	M_H
India	2001	L	South Africa	2011	M_H
Iran (Islamic Republic of)	2011	M_H	Thailand	2000	M_L
Jamaica	2011	M_H	Tonga	2006	M_L
Kazakhstan	2009	M_H	Turkey	2011	M_H
Kyrgyzstan	2009	L	Uganda	2002	L
			Uruguay	1996	M_H
			Zambia	2010	M_L

Notes: According to the income classification group of the World Bank "L" refers to low-income countries, "M_L" refers to lower-middle income countries and "M_H" refers to upper-middle income countries.

Source: Authors.

Table A.2. Variables and sources

Variable	Definition	Source
H	Average years of schooling of the population aged between 25 and 74 years old	UNESCO Institute for Statistics (2019) and Barro-Leee educational attainment dataset (Barro & Lee, 2013) available at http://barrolee.com/
Q	Proportion of the total population living in sub-standard housing (%)	Own computations based on data from the UNESCO Institute for Statistics (2019). Available at http://data.un.org/Data.aspx?d=POP&f=tableCode%3a309
GDP	Gross domestic product *per capita* (in 2011 USD purchasing power parities)	UNESCO Institute for Statistics (2019)
gov_edu	Government expenditure in education as a percentage of GDP	UNESCO Institute for Statistics (2019)
Mortality	Mortality rate of children under 5 (per 1000 live births)	UNESCO Institute for Statistics (2019)
Internet	Percentage of the population with access to the internet	UNESCO Institute for Statistics (2019)

Source: Authors.

Table A.3. List of countries with human capital data taken from the Barro-Lee dataset

Country	Year of reference in the Barro-Lee dataset
Morocco	2000
Myanmar	2010
Nicaragua	2005
Thailand	2000
Zambia	2010
Malawi	2005
Liberia	2005
India	2000

Source: Authors.

Article history:	Received: February 10, 2021
	Accepted: May 5, 2021