UNIVERSIDADE D
COIMBRA

Guilherme de França Shirazawa

Housing quality and human capital availability in developing countries

Dissertation for Master's Program in Economics, with specialization in Economic Growth and Structural Policies, presented to the Faculty of Economics of the University of Coimbra for obtaining a Master's degree

Advised by: Marta Simões, Ph.D.

January 2020

Agradecimentos

Estes agradecimentos são dedicados a todos aqueles que contribuíram de diversas formas para todo o processo de escrita desta dissertação, sem os quais certamente a mesma não poderia ter sido concluída.

Primeiramente eu gostaria de agradecer à minha orientadora, Dra. Marta Simões, por sua intensa dedicação durante todo o percurso de escrita desta dissertação, por sempre me receber ao gabinete de portas abertas disposta esclarecer todas as minhas dúvidas e por sempre me incentivar mesmo quando obstáculos surgiam. A sua orientação contribuiu inestimavelmente para a qualidade deste trabalho.

Gostaria também de agradecer a todos os professores que fizeram parte do meu percurso acadêmico durante o Mestrado em Economia da Universidade de Coimbra, sem os quais eu não poderia ter adquirido os conhecimentos necessários para a realização desta dissertação.

Aos meus pais e meus avós, que sempre me apoiaram desde o início do meu percurso acadêmico, primeiro no Brasil e agora em Portugal. Sou eternamente grato por tudo o que eles fizeram e ainda fazem por mim.

Finalmente gostaria de agradecer a todos os meus amigos por seu companheirismo nos bons e maus tempos, em especial à Alexandra Adão, Davi Paulino e à Real Tertúlia, composta por João Bento, Derick Almeida, Diogo Mazeron, Francisco Vieira e José Miguel Rebolho. Nós acompanhamos de perto a dissertação de cada um e nunca desistimos uns dos outros.

Housing is more frequently regarded as a target of development than as a tool for development planning – as an end in itself rather than as a means to another end.

Burns & Tjioe (1968, p. 396)

Abstract

The main objective of this work is to analyse the relationship between the quality of housing and the availability of human capital 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, acquired in an individual's education system. The analysis was performed using cross-sectional data, mostly taken from the UNESCO database for 52 low and middle-income countries, following the general income classification from the World Bank Atlas Methodology, in which we estimated a set of regressions by OLS with the variable of interest as the proportion of the population living in houses below the minimum standard of quality and the dependent variable, the average years of schooling of the population. The obtained results from the regressions indicated a positive relationship between an increase in the proportion of the population living in substandard housing and average years of schooling, but with little or no statistical significance, making the empirical analysis inconclusive. We concluded that the reliability of the housing quality data provided by UNESCO may be the source of the previous inconclusive results, pointing to the need to invest more resources in the gathering of data and the production of indicators. This investment may have a relevant role in future research for addressing the gap in the literature and for the definition of public policies.

Keywords: human capital, housing, developing countries, cross-sectional analysis.

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

Resumo

O objetivo principal deste trabalho é analisar a relação entre a qualidade da habitação e a disponibilidade de capital humano no contexto dos países em desenvolvimento. A análise tenta preencher uma lacuna na literatura atual referente à falta de estudos empíricos que abordem o impacto que as condições de moradia podem ter sobre o capital humano, adquirido no sistema de ensino, de um indivíduo. A análise foi realizada utilizando dados *cross section*, maioritariamente adquiridos da base de dados da UNESCO para 52 países de baixa e média renda, seguindo a classificação geral de renda da metodologia Atlas do Banco Mundial, em que estimamos um conjunto de regressões por meio de OLS com a variável de interesse sendo a proporção da população que vive em habitações abaixo do padrão mínimo de qualidade e a variável dependente os anos médios de escolaridade da população. Os resultados obtidos nas regressões indicaram uma relação positiva entre um aumento na proporção da população que vive em habitações de qualidade abaixo do padrão e anos médios de escolaridade, mas com pouca ou nenhuma significância estatística, tornando a análise empírica inconclusiva. Nós concluímos que a fiabilidade dos dados relativos à qualidade da habitação fornecidos pela UNESCO pode estar na origem dos resultados inconclusivos anteriores, apontando para a necessidade de investir mais recursos na recolha de dados e na produção de indicadores. Este investimento pode ter um papel relevante em futuras pesquisas que possam preencher a presente lacuna na literatura e para a definição de políticas públicas.

Palavras-chave: capital humano, habitação, países em desenvolvimento, análise *cross section*

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

Summary

Index of Tables

Index of Figures

1. Introduction

The impact of housing on the economy has gone through several debates over the last few decades. Mainly it has been treated as a factor that has an influence in the short-run macroeconomic scenario, usually through its effect on aggregated demand (Harris & Arku. 2006). In this work project we take a different perspective from the previous literature, focusing on a long-run perspective regarding the impacts that housing has on the lives of individuals, namely through education improvements, an important source of human capital accumulation.

Following the exogenous growth model proposed by Mankiw, Romer and Weil (1992) and the endogenous growth model proposed by Lucas (1988), human capital accumulation became a crucial component to better understand the differences in standards of living and economic growth rates across countries, respectively, as was confirmed by empirical studies of Benhabib & Spiegel (1994) and Hanushek & Woessmann (2011). With the advent of these papers, human capital accumulation has become a relevant and crucial factor to include in the analysis of economic growth and housing quality can play an important role in the explanation of human capital dynamics.

At times, having a house has been recognized by the literature as providing a foundation for individuals to develop and become more prosperous. The usual 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 environment for cooking. The access by individuals to these factors would in turn have a positive effect on productivity and their overall 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 impacts specifically on human capital.

Given the present gap in the literature, this research aims to empirically investigate the relationship between housing quality and human capital availability 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. Most of the data used comes from the UNESCO Institute for Statistics.

The remaining of this work project is organized in four main sections. In section 2 we give the historical context of the debate concerning human capital and housing. Section 3 provides an overview of the literature concerning the relationship of human capital and economic growth along with housing and macroeconomic analysis. In section 4, the empirical model and data are presented followed by a discussion of the results obtained and, lastly, Section 5 concludes.

2. Historical context

The goal of this section is to frame the main concepts used throughout the present study, human capital and housing, and put them in a historical context.

2.1 A Brief History of Human Capital:

Human capital can be defined as "the knowledge, skills, competencies and attributes embodied in individuals that facilitate the creation of personal, social and economic well-being" (Healy et al, 2001, p.18). A first mention of the concept dates back to the 18th century and can be found in Adam Smith's Wealth of Nations with the notion that individuals that have acquired useful abilities from education or skills from their labour will earn higher wages (Savvides & Stengos, 2008). These abilities that lie within the individual increase labour productivity, similar to the effect of physical capital on workers' productivity.

Moving forward to the end of the 19th and beginning of the 20th century, various attempts were made to estimate the value of human capital in industrial nations, for example, Nicholson for the United Kingdom in 1891, Barriol for France in 1908 and Huebner for the United States in 1920 (Savvides & Stengos, 2008). The concept was also utilized in the subsequent period of the first World War (1914-18) to show that the cost of the conflict was considerably higher than that associated to the loss of physical capital.

A major change in the approach to human capital occurred in the University of Chicago in the 1960s with the works of Theodore W. Schultz and Gary Becker, among others. In a 1960 paper entitled "Capital formation by education", Schultz proposes to treat education as an investment and its consequences as a form of capital based on the hypothesis that human capital leads to important increases in national income (Schultz,

1960). Another major contribution was that of Becker (1962) also concerning education and training, where the author estimates the impact of education, considering primary, secondary and tertiary education, on an individual's earnings over his or her lifetime. With the advent of these papers and others, the treatment of the concept of human capital changed and it became possible to measure it and estimate its economic impact, thus becoming a valuable variable for economic analysis.

2.2 Housing in macroeconomic analysis

The discussion on the role of housing in the economy has shifted intensively in terms of perspective during the 20th century. Starting in the 1930s, housing began to be considered as a matter of public concern, with governments developing construction projects aimed at eradicating slums in the interest of fighting diseases and crime within dangerous neighbourhoods. During the post-war period of 1945 until the 1960s, development economics regained a new interest with the main goal of moving economies out of poverty, with capital accumulation and improvements in labour productivity considered as priority to achieve the former objective (Tibaijuca, 2013). This presented a challenge to the developing world as capital was in short supply because savings rates were low due to low wages. In this context, the poverty relieving effort was focused on investments that showed promising productivity returns, namely through basic manufacturing industry and infrastructure investments. Housing was thus not perceived as a priority within the allocation of scarce resources in the pursuit of long-run prosperity due to its low productivity, high capital-output ratio and pressure exerted on the balance of payments due to its high import content. The main role for this sector was addressing occasional social problems and not as a permanent tool to address structural problems in the economy, as housing construction in general was viewed as a non-productive social expenditure, not an investment. Expenditure in housing construction was only admitted to be productive in narrow cases, such as housing for workers that act in key industries, only then housing would have its contribution to economic development (Ibidem).

One example of the perspective on the economic role of housing described above can be found in the work of Howenstine (1957). The author states that for a poor country to develop it needs to allocate its resources to key industrial sectors instead of housing projects. Only after this first investment, is it wise to shift the focus towards

housing construction with the aim of providing for workers in those key industries, "whose contribution to national productivity could be expected to benefit the most from better housing" (Howenstine (1957), p. 25). However, Howenstine also acknowledged that the provision of better quality shelter to the underemployed or the unemployed could boost their productivity through its positive effect on health and creating a greater incentive to work and save, concluding that productivity gains could fall in the 10-30 per cent range.

Associated with concerns about productivity in the decades following the end of the Second World War, the discussion on the role of the housing sector was also present in the work of Keynesian economists. In fact, the housing sector became a largely accepted counter-cyclical public policy tool to be utilized in industrial economies like those of the U.S and Europe. In the context of developing countries, this logic could be applied, with some modifications, given its economic characteristic of widespread unemployment. In the view of Currie (1966), housing construction could be utilized as a semi-permanent tool for economic management due to its ability to absorb unskilled labour that flowed from the countryside to new urban areas, rather than just as a temporary measure as in the case of industrial economies.

By the mid-1960s to the 1970s, there was a shift in paradigm, from viewing housing as a public or social programme to the notion that housing can play a major in economic growth. The relevance of the housing sector changed substantially in the 1970s when the World Bank began to fund slum upgrading programmes in developing countries. This new approach has been credited to Robert McNamara, who, as President of the World Bank, promoted the idea that human capital (in the context of health) is crucial to development. The World Bank produced its landmark Housing Sector Policy Paper in 1975, heralding its formal commitment to housing and lending guidelines in general. The perspective of the World Bank on housing at the time can be summarized by the following excerpt:

> "Housing has substantial social benefits, including the welfare effects of shelter from the elements, sanitation facilities and access to health and education services. Improved health and education and better access to income-earning opportunities can lead to higher productivity and earnings for low-income families. It is thus for sound economic reasons that, after food, housing is typically the largest item of household expenditure for poor families; and that they are willing to go to great lengths to obtain housing at locations with access to employment, even if this means incurring the risks of illegal 'squatting'." (World Bank, 1975, p.3)

It was during this time that housing steadily became seen as an investment that yields a positive socio-economic outcome for developing countries through its effect

on helping lifting individuals out of poverty and, consequently, propelling economic growth. From the 1970s onwards, other returns to the investment on housing became clear such as the continuous stream of services (i.e. shelter from the environment, a place to cook and to rest) that are enjoyed by its owners during an extended period of time. Additionally, having a house opens up the possibility for its inhabitants to rent part or the whole of it, or to set up home-based businesses, which in both cases can make a significant difference in the lives of individuals and households in developing countries.

The last main series of events that helped set a new perspective on role of the housing sector in economic growth was the emergence of the New Industrialized Countries of South-East Asia (Taiwan, Singapore and Hong Kong) and also Japan. In these countries, investment in the housing sector made for a large portion of gross fixed capital formation (GFCF) through the 1970s to 1990s, largely induced by government programs for public housing. The pursue for urban development in these countries has led to the creation of much-needed jobs, which helped incorporate rural unskilled workers into the urban context. Thus, the positive impact that housing construction had in these emerging economies contradicted the past idea of housing as solely a consumer good (Harris & Arku, 2006).

At the empirical level, Green (1997) finds that residential investment in the United States Granger causes GDP, which does not mean that fluctuations in residential investment causes fluctuations in GDP (i.e. the business cycle), but that it is a good predictor of its future behaviour. In line with Green's research, and with updated data, Hongyu et al. (2002) finds that housing investment in China predicts a growth in GDP in the short-run and is also a more important driving force to the economy than non-housing investment. The same relationship can be seen in Terzi and Bolen (2008), where the authors concluded that there is a positive correlation between GNP and housing construction, concluding that housing construction is one of the important driving forces of the national economy of Turkey. Considering Green's reference study in the field and two examples for relevant developing economies we can see that, contrary to what was previously believed by the housing literature, housing investment is indeed a factor that influences economic growth. What remains to be seen is if the housing sector has an influence over the economy through channels other than investment. This will be explored in section 3.2.

3. Literature Review

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 and better identify the gap in the literature that the present study tries to address.

3.1 Human Capital and Economic Growth

After the initial methodological and empirical foundations laid in the 1960s, crucial works made in the field of economic growth gave new importance to the concept of human capital. During the 1980s and 1990s, a series of works were published with the intent of incorporating human capital into the field of economic growth through different approaches.

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 just another input into final goods production and also 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 consider annual data (from 1960 to 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 of income in this sample of countries.

A different approach to the relationship between human capital and economic growth is that proposed by Lucas (1988) which lies within the category of AK growth models. In this model, human capital accumulation creates positive externalities to the economy which was designated as “learning-by-studying”. In this model, there are two distinct sectors in the economy, the education sector where human capital is produced through education and the final goods sector where the increase in human capital at the individual levels 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 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) are some of the first authors that try to identify the relative importance of human capital 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 cross-country panel data for 78 countries with annual observations from 1965 to 1985, with the proxy for human capital 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, meaning that the cost to follow the technological leader is lower than to innovate.

The empirical identification of the role of human capital for 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 200 using microdata from international achievements tests (PISA scores – Programme for International Student Assessment) 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 measure for education as it, implicitly, assumes that learning outcomes from additional years of education are the same across countries. The results from the regressions estimated indicate that cognitive skills are a better predictor of economic growth than average years of schooling, confirming the indeed that the quality of human capital (in its education component) is more important for the growth than its quantity. In any case, the estimations still indicated that average

years of schooling can still explain an important part of long-run growth, which is in line with the empirical findings from Mankiw, Romer and Weil (1992).

The largest contribution made by the article that concerns the importance of human capital to economic growth models is where the authors try, extensively, to apply both endogenous and exogenous growth models to different future scenarios of general education for the OECD countries. However, independently of the model or education reform that was calculated, the results clearly indicate that the improvement in education (i.e. human capital, in this study) has a positive and substantial impact on every economy considered.

Even though the human capital quality has a larger role than its quantity in economic growth of countries, what the two previous studies show is that human capital in general is a crucial factor to develop when countries are in pursuit of long-term prosperity.

3.2 Housing and human capital

Despite some theoretical and empirical analyses on the role of housing in economic growth, to the best of our knowledge there are no studies that try to analyse its mediating role through human capital. Due to the importance of the latter for economic growth, as reviewed in the previous section, this section analyses the relationship between housing and human capital whist providing some arguments on what to expect in terms of the sign of the former relationship based on a few previous studies found.

Robert Healy (1971) provides some clues on what to expect as far as the relationship between housing and human capital accumulation is concerned. The author analyses the impact of a rehousing program for a group of workers in a Mexican factory on their 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 man-hour and lower absenteeism. The analysis covered two groups, the workers that were rehoused and the workers that remained in their original low-quality homes, for a period of four years, which covers data for two years before the rehousing of the first group and two after, and, in addition to productivity, considered the effects of the program on worker's absenteeism and health. The author found that after one year of the rehousing programme, productivity

increased and housing-related health issues 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 if these positive outcomes could also have an effect on individual's educational path.

In Bradley and Putnik (2012) the authors analyse the relationship between home environment conditions that are associated with child development, such as 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 status, in that they were of higher quality in high-HDI countries and of lower quality in all low-HDI countries.

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 have a negative impact on the probability of a teenager to finish high school, among other factors such as ethnicity.

In summary, housing conditions have been portrayed as exerting a positive influence on health and educational outcomes in studies for single country studies, with the exception of Bradley & Putnik (2012) that covers a broader sample of countries, 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 analysis. An approach like this would also expand the housing literature, deviating from its focus that mainly lies in the housing's impact on variables that affect only the short-run economic scenario and giving new clues for its potential role in terms of affecting the economy in the long-run.

4. Empirical analysis

The main goal of the empirical analysis that will be carried out in this section is to investigate the relationship between housing quality and human capital for developing countries. We will start by presenting the empirical model, followed by the description of the data and the estimation methodology used. Next we will present and discuss the results obtained.

4.1 Empirical model, data and methodology

The empirical analysis considers low- and middle-income countries. We utilize the definition from the World Bank of US$12.375 of GNI *per capita* as a threshold for a country to be considered rich, in 2018 values, so that countries below this threshold are classified as middle-high income (US$3,996<GNI *per capita*< US$12,365), middle-low income (US$1,026<GNI *per capita*< US$3,995) or low income (GNI *per capita*< US$1,025) countries, the ones included in our analysis. These values change for past years, so we utilize the corresponding thresholds for each country in its respective year of observation utilizing a conversion table from the World Bank's Atlas Methodology. The choice for the exclusion of high-income countries was due to the small variation in housing quality in this group of countries. The final sample thus 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 that will be 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*, housing quality in country i. The model additionally includes a vector *X* of control variables with other determinants of human capital 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 identified expressed in equation (2), where GDP is real income *per capita* , gov_edu is state intervention at the educational level, Mortality represents 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 from 1971 to 2000 based on a set of explanatory variables (e.g. population's health, expenditure in education, urbanization and gender equality). Table A.2 in the appendix identifies the variables used, describes how they are measured and identifies the sources of the data.

$$X = \begin{bmatrix} GDP \\ Gov\_edu \\ Mortality \\ Internet \end{bmatrix} \qquad \textbf{(2)}$$

We measure human capital as average years of schooling of the population for each country taken from the UNESCO Institute for Statistics. This variable corresponds to the average years of the highest level of education attained by the individuals aged between 25 and 74 years old. We acknowledge that there are other ways to approach the education component of human capital such as enrolment rates, for measuring quantity of schooling, and results from test scores, for measuring the quality of schooling as in Hanushek & Woessmann. Despite these options, the choice for average years of schooling was due to its wider availability for developing countries and is in line with applied economic growth studies such as Benhabib & Spiegel (1994).

One problem we encountered before the start of the regressions was the matching of the cross-sectional data for the human capital stock and housing quality for some countries. 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 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 the sample (see Table A.3 in the Appendix).

Our explanatory variable of interest is the proportion of the population that lives in sub-standard housing, Q, as a proxy variable for the lack of housing quality. The use of this variable dictated the structure of the data for 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 for which each observation was recorded was also usually different across a large number of countries.

To measure the lack of housing quality 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 homogenous criteria for

classification of housing for all countries, including 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 we retrieved is divided into several categories of housing qualities. Table A.4 and Figure A.1 in the Appendix summarize the different housing quality classifications by the United Nations and present the respective definitions. The most relevant definition to our analysis is that of housing good quality habitation, or as it is defined in the database, i.e. 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 to be occupied by one household (UN., 2017, p.249). Furthermore, the UN also defines basic facilities for decent living: piped water, toilet, fixed bath or shower, kitchen or other space for cooking, with all four located within the same dwelling. All other categories of housing quality fail to meet the former criteria and so we decided to call 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 (3).

$$Q = \left(\frac{Population\ in\ sub-standard\ housing}{Total\ population}\right) * 100 \qquad (3)$$

As far as the sign of the different estimated coefficients is concerned, we expect a negative relationship between lack of housing quality and human capital, this sign would indicate that higher shares of the population living in sub-standard housing (higher Q) are associated with lower stocks of human capital because lower standards of living in terms of housing may act as a disincentive for individuals to pursuit more education due to poor environment of study 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, for its relative cost lowers 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 be a relevant barrier for individuals to be able to afford investing in education since the individual's poor health status can disincentivise attendance in school, lower learning ability or even induce dropping out of school altogether. Finally, the estimated percentage of the population with access to the internet is expected to have a positive sign, as it is a tool that helps individuals in the education process through online materials that are of aid when doing homework and for studying a new subject as the study by Lei & Zhao (2007) indicates. Given this aid in the day-to-day learning process, we believe that having access to the internet can be a relevant factor that enables individuals to endure longer in their educational course than they could otherwise endure.

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 far 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 only of low and middle-income countries, 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 first being larger than the second.

Concerning the behaviour of the variables' distribution we can see, by looking at the median and the mean, that 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). The high skew for this variable poses a challenge for generating a scatterplot graph with the human capital data that can expose a clear relationship. A visual representation of this skew can be seen in Figure A.2 in the

appendix, where most of the observations are concentrated in low values to the left, with the exception of some outliers. However, when we repeat the same scatterplot for the variables in logarithms the extreme values/outliers disappear, but the relationship still remains weak (see Figure A.3 in the appendix).

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 wish to find might be true. However, when we ran the correlation test for the same variables in logarithm the correlation coefficient changes the sign (0.1052), although it remains statistically insignificant. This change in sign is probably due to the reduction in the distance between the outliers and the mean caused by application of the logarithm.

Table 1 – Descriptive Statistics for housing 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 housing quality. H is average years of schooling.
Source: elaborated by the author using GRETL.

4.2 Results

The results from the OLS estimation of the baseline equation (1) are presented in Table 2. We present four distinct regressions corresponding to different model specifications, i.e. different versions of equation (1) depending on the set of control variables considered in order to check the robustness of the result of interest, the estimated coefficient for the lack of housing quality variable, to different combinations of the control variables. We eliminated the control variables according to its importance to the explanation of differences in schooling levels following the previous empirical literature or by its lack of statistical significance. We start by leaving out the variables that are less consensual as determinants of human capital, like access to the internet, up to the most parsimonious model that considers only GDP *per capita* as a control variable, following the relevance given for these variables in Baldacci et al. (2008). Thus, column (1) in Table 2 presents the results considering all control

variables; column (2) takes 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*	0.9789	−1.0653**	−0.9320**
	(0.8470)	(0.7937)	(0.4323)	(0.3918)
Ln_Q	0.0806**	0.0885**	0.0821*	0.0743*
	(0.0379)	(0.0381)	(0.0411)	(0.0396)
Ln_GDP	0.1211	0.1848***	0.3306***	0.3288***
	(0.0749)	(0.0633)	(0.0435)	(0.0432)
Ln_gov_edu	−0.0219	0.0036	0.0765	
	(0.0982)	(0.0982)	(0.1026)	
Ln_mortality	−0.2039***	−0.2010***		
	(0.0664)	(0.0673)		
Ln_Internet	0.0523			
	(0.0340)			
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-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. Source: elaborated by the author utilizing the software Gretl.

According to the results presented in Table 2, in all of the models estimated we obtained 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 housing quality remains basically unchanged, ranging from 0.07 and 0.08, indicating that if this variable increases by 10% then the human capital stock would increase by 0.7-0.8% 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. This result is in line with the existing literature on economic growth (Benhabib & Spiegel, 1994). 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 is the result of poor efficiency in resource allocation. However, the former coefficient is not statistically significant and turns positive in models (2) and (3), when the variables for population health and access to the internet are removed from the estimation. Again, none of the coefficients is statistically significant and the estimated elasticity is relatively low. It thus seems that state intervention in the education system has not had a significant impact on human capital availability 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 educational process that individuals go through, 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 observation, we can say that the model provides a satisfactory prediction ability. For the F-test's p-value, all the models managed to reject the null hypothesis of the test, meaning that the coefficients estimated can better explain the dependent variable than a model that explains the dependent variable with its own mean and all the coefficients are equal to zero. 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, in 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, i.e. there is correlation in the error terms. 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.

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.

Table 3 – Results with OLS and robust standard errors

	(1)	(2)	(3)	(4)
const	1.4773*	0.9789	−1.0653*	−0.9320*
	(0.7909)	(0.7681)	(0.5358)	(0.5118)
Ln_Q	0.0806	0.0885*	0.0821	0.0743
	(0.0501)	(0.0483)	(0.0543)	(0.0493)
Ln_GDP	0.1211	0.1848***	0.3306***	0.3288***
	(0.07220)	(0.0657)	(0.0522)	(0.0524)
Ln_gov_edu	−0.0219	0.0036	0.0765	
	(0.1032)	(0.1053)	(0.1120)	
Ln_mortality	−0.2039***	−0.2010***		
	(0.0533)	(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: elaborated by the author using Gretl.

First, as can be seen 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. These results indicate the statistical significance obtained in the models present in Table 2 was indeed biased due to the presence of heteroskedasticity. Overall, these results indicate 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.* Judging by the adjusted R-squared and by the Akaike criterion, model (1) still 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 continued 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 the classification from the World Bank. This implies still considering a set of countries with quite different levels of income and, as a consequence, comprising populations living under 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 for each of the three levels of income interacted with the lack of housing quality variable, where the terms correspond to:

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

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

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

This specification allows us to investigate if the relationship between lack of housing quality and human capital availability differs according to different 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 table's notes.

Table 4 – OLS regressions with interaction terms for income groups

	(1)	(2)	(3)	(4)
const	1.1363	0.5835	−1.3747**	−1.2629**
	(0.8975)	(0.8307)	(0.6198)	(0.6058)
lnQ*dummy_low	0.1551	0.1525	0.1742	0.1719
	(0.1026)	(0.1008)	(0.1097)	(0.1097)
lnQ*dummy_middle_L	0.0552	0.0751	0.0600	0.0504
	(0.0595)	(0.0584)	(0.0680)	(0.0561)
lnQ*dummy_middle_H	0.0364	0.0371	0.0169	0.0135
	(0.0400)	(0.0393)	(0.0428)	(0.0408)
Ln_GDP	0.1549*	0.2259***	0.3701***	0.3681***
	(0.0851)	(0.0739)	(0.0642)	(0.0649)
Ln_gov_edu	−0.0337	0.0002	0.0615	
	(0.1088)	(0.1098)	(0.1190)	
Ln_mortality	−0.1857***	−0.1855***		
	(0.0545)	(0.0533)		
Ln_Internet	0.0585*			
	(0.0305)			
Total countries	52	52	52	52
R-squared	0.6573	0.6353	0.5768	0.5735
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: elaborated by the author using 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. In all the models, we see that the coefficient for the interaction term for low-income countries presents the highest value, indicating that, for this group of countries an increase in the housing quality ratio has a higher positive association with human capital availability. This result is surprising, it not only goes against our initial expectations but it additionally indicates that having a worse household living condition has a larger positive effect for individuals that live in an economic environment considerably more restricted than high or even medium-income countries. However, as was mentioned before, none of the coefficients of the three interaction terms is statistically significant. For the control variables, we retrieved them from the model in the same order as in Tables 2 and 3. The public expenditure in education still appears with no statistical significance, indicating that

this variable may not be of relevant aid in explaining variations in human capital stocks. Now, the child mortality rate variable maintains its statistical significance at the 1% level as in the previous tables, indicating the health status of a country is a relevant factor that impacts human capital. Now, the coefficients for GDP *per capita* and access to the internet remain relatively unchanged from the previous models.

Concerning the overall performance of the regressions, all models rejected the null hypothesis of the F-test. Model (1) had the best performance judging by the Akaike criterion and has managed to better explain the variations in human capital when compared to its peers by their adjusted R-squared.

Finally, we test in a different way for the possibility of non-linearities in the relationship between housing quality and human capital availability 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 over 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 threshold of the function, having a portion of the population living in sub-standard housing is actually positively correlated to human capital. In this scenario, this could be true 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 with the housing quality variable squared, representing the quadratic term of the function.

Table 5 – Results with quadratic regressions

	(1)	(2)	(3)	(4)
const	1.7311*	1.1181	−0.7736	−0.7375
	(0.8684)	(0.8168)	(0.5283)	(0.5292)
Ln_Q	0.1618*	0.1563*	0.1661*	0.1677
	(0.0894)	(0.0919)	(0.0987)	(0.0985)
$(\mathbf{Ln\_Q})^2$	−0.0362	−0.0297	−0.0366	−0.0379
	(0.0272)	(0.0280)	(0.0987)	(0.0293)
Ln_GDP	0.0951	0.1723**	0.3067***	0.3055
	(0.0811)	(0.0715)	(0.0561)	(0.0570)
Ln_gov_edu	−0.0811	−0.0413	0.0167	
	(0.1072)	(0.1066)	(0.1084)	
Ln_mortality	−0.1927***	−0.1914***		
	(0.0496)	(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: elaborated by the author using 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 the 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 (4)) we can compute the maximum of the function, i.e. the value of Q beyond which the relationship becomes negative:

$$\left(\frac{dlnH}{dlnQ} = 0 <=> \beta_1 + 2\beta_2 Ln\_Q = 0\right), \qquad \textbf{(4)}$$

The maximum is located at $Ln\_Q = 2.667$ (as visual representation of this function can be seen at Figure A.4 in the appendix). Now, this turning point in the function would indicate that countries with values higher than this maximum value in Ln_Q would have the negative relationship we first proposed, which, with the values in levels, would be beyond a ratio of 14.39% in Q. 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 the considered countries, 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.

4.3 Discussion

In this section we will try to assess the main issues that hinder the robustness of the results obtained, in particular the lack of statistical significance and heteroskedasticity, as well as the potential implications as far as structural policies that promote economic growth in developing countries are concerned.

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

$$Human\ capital = \beta_1 family\ inputs + \beta_2 schooling\ inputs + \beta_3 individual\ ability + \beta_4 other\ factors + \epsilon \qquad \textbf{(5)}^{1}$$

As it is possible to see, this model differs from ours in particular because it considers factors that impact individuals human capital at the micro level where family inputs and individual abilities define the context in which an individual develops its cognitive abilities to absorb knowledge provided by the educational system. However,

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

in this work the authors do not pursuit an empirical analysis to calculate the degree in which these inputs help to predict the quality of human capital.

In another study, Lee & Barro (2001) develop an empirical analysis of schooling quality in a cross-section of countries considering a similar production function, summarized in equation (6):

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

Where Q stands for tests scores for individual *i, F* is for family factors, such as the parents' income and educational attainment for student *i,* which affects the probability that children enrol in, attend and complete school, but also the ability of the child to learn. *R* stands for school resources, such as pupil-teacher ratios, average teacher salary, educational expenditure per pupil and school length, with all these factors encompassing the ability of the schooling system to provide a good quality environment of learning. The empirical study found that family background and school resources have a strong positive association with student performance.

The inputs of the human capital production function in Lee & Barro (2001) gives us a clue for one of the reasons why our analysis could not provide 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 potentially important variables, which explains our need to perform the Breusch-Pagan heteroskedasticity test in our original models (Table 2). The test results indicate that the error term has a variance that does not have a white-noise behaviour, indicating that the models suffers from a 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 encompassing 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 can have its effect over human capital quality, not quantity. Even if the relationship occurs through this channel, we were not able

incorporate human capital quality in this analysis as internationally comparable data on student's performance is not available for the group of countries under analysis.

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 with which we could work since we also had to guarantee that we had human capital data for the 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 the 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 observations is not only due to the problem with the temporal component. The sample of countries available is actually larger than 52, but we were obligated to reduce this number due to lack of data for mean years of schooling for some countries.

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 having inconsistent observations for each category of housing quality, in which countries did not have information on all categories, and enabled an encompassing variable that represents the overall problems in 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. This being true, 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 forward 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 precision was not as high as for the remaining variables.

Despite all that was discussed above, there is also the question about the presence of endogeneity within the general model that we constructed, which considers the housing quality ratio as independent from human capital. In reality, it could also

be the case that human capital influences housing quality through the positive effect that the first has on income, which in turn may induce an improvement in housing quality. This problem could in theory be solved by using instrumental variables estimations methodologies, but we encountered no literature that might indicate which variables could serve as good instruments and even if we had it would be unlikely the respective availability would match our limited dataset.

Given the problems we have described so far, we are led to reject the positive relationship between housing quality and human capital indicated by all the estimated models. This rejection is further supported by the overall lack of statistical significance of the estimated coefficient in the models considering robust standard errors and the fact that even when the coefficient was statistically significant this never happened at the 1 or 5% levels, only at the 10% level and in a very small subset of the regressions using robust standard errors.

5. Conclusion

The main objective of this work was to empirically investigate the relationship between housing quality and the stock of human capital in the context of developing countries. This empirical research was motivated by a gap in the literature concerning quantitative studies that explore this relationship, since the studies on the impacts of housing quality tend to focus on its impact on health and productivity of individuals, so at the micro level, and leaving the education context out of the analysis.

For the estimation of the relationship, we carried out a cross-sectional analysis for which we estimated different regression specifications considering different control variables, robust standard errors and the possibility of non-linearities according to countries income levels or the size of the lack housing quality variable represented by a quadratic relationship.

The findings indicate that overall there is no statistically significant relationship between the lack of housing quality and human capital although the respective estimated coefficient is positive, contrary to what we were expecting. However, we are led to believe that there is a problem in the quality of the housing data. What we found was that the way in which the housing quality are compiled presents some problems in terms of its structure, e.g. inconsistent observations for different categories of housing quality and in time. In the end, this problem, along with the small number of observations in the UNESCO database for the group of countries

under analysis (developing countries but for which we could only select a subset of 52), produced inconclusive results.

Regarding public policy implications, our results do not endorse investing in access to better quality housing as a means to improve human capital availability 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 when considering child mortality rates.

Although, the inconclusive nature of the results did not aid in bridging the gap identified in the literature, we believe that a major contribution of this work lies in raising awareness to the need for better collected and structured data for housing quality and human capital to enable for future empirical analysis that can consider also human capital quality instead of just human capital quantity. Considering the rapid rate of urbanization that has been happening through the end of the 20th century and the beginning of the 21st in developing countries, such future researches can become helpful guidelines for broad housing policies.

List of 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. https://doi.org/10.1016/j.worlddev.2007.08.003

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, *70*(5, Part 2), 9-49. https://doi.org/10.1086/258724

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. https://doi.org/10.1016/0304-3932(94)90047-7

Bradley, R. H., & Putnick, D. L. (2012). Housing quality and access to material and learning resources within the home environment in developing countries. *Child development*, *83*(1), 76-91. https://doi.org/10.1111/j.1467-8624.2011.01674.x

Burns, L. S., & Tjioe, B. K. (1968). Housing and Human Resource Development. Journal of the American Institute of Planners, 34(6), 396–401. doi:10.1080/01944366808977568

Currie, L. B. (1966). *Accelerating development: the necessity and the means* (No. HC59. 7 C85). https://doi.org/10.2307/2510987

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. https://doi.org/10.1111/1540-6229.00714

Hanushek, E. A., & Woessmann, L. (2011). How much do educational outcomes matter in OECD countries?. *Economic policy*, *26*(67), 427-491. https://doi.org/10.1111/j.1468-0327.2011.00265.x

Harris, R., & Arku, G. (2006). Housing and economic development: The evolution of an idea since 1945. *Habitat International*, *30*(4), 1007-1017. https://doi.org/10.1016/j.habitatint.2005.10.003

Healy, T., Cote, S., et al. *The Well-Being of Nations: The role of human and social capital*, Centre for Educational Research and Innovation, OECD (2001).https://doi.org/10.1787/9789264189515-sum-es

Healy, R. G. (1971). Effects of improved housing on worker performance. *Journal of Human Resources*, 297-308. https://doi.org/10.2307/144952

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.

Howenstine, E. Jay. (1957). Appraising the Role of Housing in Economic Development. International Labour Review, No.75, pp. 21-33.

Lee, J. W., & Barro, R. J. (2001). Schooling quality in a cross–section of countries. *Economica*, *68*(272), 465-488. https://doi.org/10.1111/1468-0335.00257

Lei, J., & Zhao, Y. (2007). Technology uses and student achievement: A longitudinal study. *Computers & Education*, *49*(2), 284-296. https://doi.org/10.1016/j.compedu.2005.06.013

Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, *22*(1), 3-42. https://doi.org/10.1016/0304-3932(88)90168-7

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. https://doi.org/10.2307/2118477

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, 2008, Liverpool, UK.

UN. (2017). *Principles and Recommendations for Population and Housing Censuses, Revision 3*. UN. https://doi.org/10.18356/bb3ea73e-en

Savvides, A., & Stengos, T. (2008). *Human capital and economic growth*. Stanford University Press. https://doi.org/10.11126/stanford/9780804755405.001.0001

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, *70*(1), 65-94. https://doi.org/10.2307/1884513

Schultz, Theodore. (1960). Capital Formation by Education. Journal of Political Economy, Vol. 68, No. 6, pp. 571-583. https://doi.org/10.1086/258393

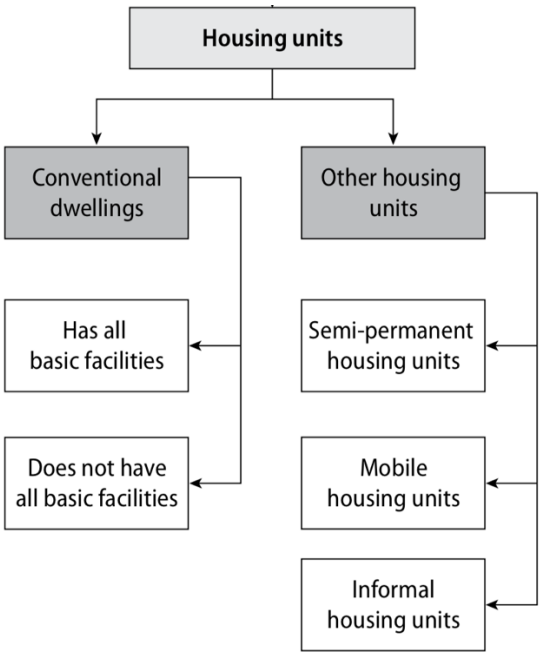
Schultz, Theodore. (1961). Investment in Human Capital. The American Economic Review, Vol. 51, No. 1, pp. 1-17.

Tibaijuka, A. (2013). *Building prosperity: Housing and economic development*. Routledge. https://doi.org/10.4324/9781849770040

World Bank. (1975). Housing Sector Policy Paper. World Bank, Washington, DC.

Appendix

Figure A.1 – Classification of housing units



Source: Principles and Recommendations for Population and Housing Censuses (2017, p. 250)

Figure A.2 – Actual and fitted values (Human capital versus Housing quality)

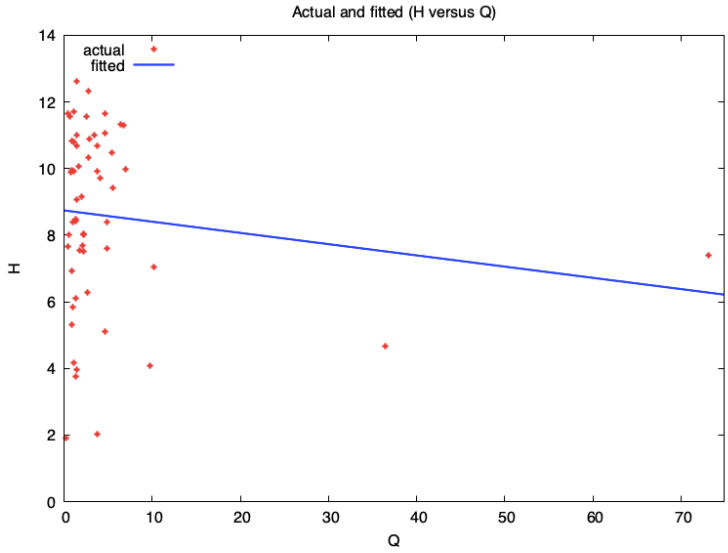


Figure A.3 - Actual and fitted values (ln_Human capital versus ln_Housing quality)

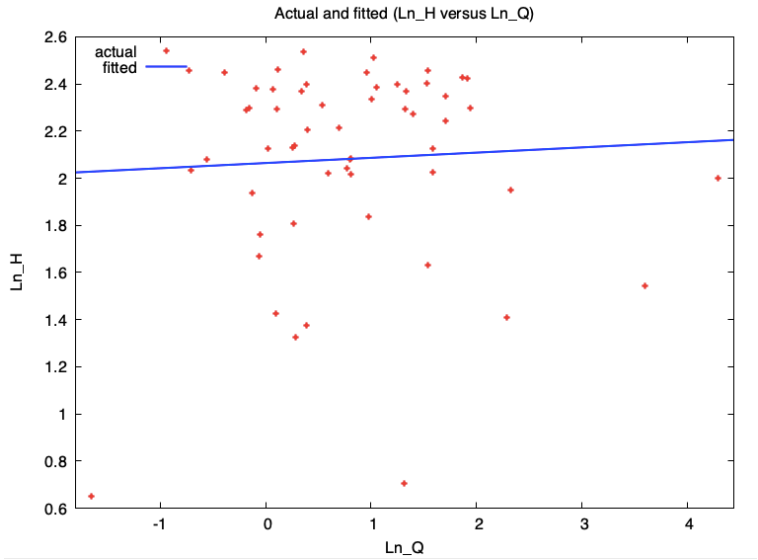
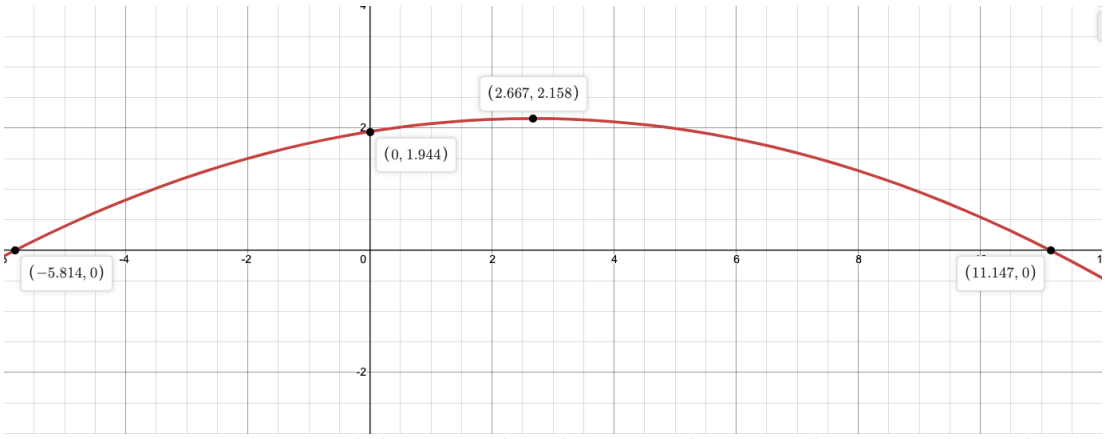


Figure A.4 – Visual representation of the quadratic regression function



Notes: Y axis represents human capital (Ln_H) and X axis represents housing quality (Ln_Q). The function also considers the control variables, in which each variable takes the value of its mean and then is multiplied by the respective coefficient. This process thus transforms the four variables into constants. Source: author with the help of Desmos graphing calculator.

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

Country	Year	Income Classification
Albania	2011	M_L
Argentina	2010	M_H
Armenia	2011	M_L
Azerbaijan	2009	M_H
Belarus	2009	M_H
Bolivia (Plurinational State of)	2012	M_L
Brazil	2010	M_H
Bulgaria	2011	M_H
Chile	2002	M_H
Costa Rica	2011	M_H
Croatia	2001	M_H
Cuba	2002	M_L
Dominican Republic	2002	M_L
Ecuador	2010	M_H
Ethiopia	2007	L
Egypt	2006	M_L
Georgia	2002	L
Ghana	2010	M_L
Guinea	2014	L
Hungary	2001	M_H
India	2001	L
Iran (Islamic Republic of)	2011	M_H
Jamaica	2011	M_H
Kazakhstan	2009	M_H
Kyrgyzstan	2009	L
Latvia	2011	M_H
Lesotho	2006	M_L
Liberia	2008	L
Malawi	2008	L
Malta	1995	M_H
Malaysia	2010	M_H
Mexico	2010	M_H
Morocco	2004	M_L
Myanmar	2014	M_L
Nicaragua	2005	M_L
Peru	2007	M_L

Country	Year	Income Classification
Philippines	2000	M_H
Poland	2002	M_H
Republic of Moldova	2004	L
Romania	2011	M_H
Russian Federation	2010	M_H
Rwanda	2012	L
Saint Lucia	2010	M_H
Serbia	2011	M_H
Slovakia	2001	M_H
South Africa	2011	M_H
Thailand	2000	M_L
Tonga	2006	M_L
Turkey	2011	M_H
Uganda	2002	L
Uruguay	1996	M_H
Zambia	2010	M_L

Notes: This table lists the countries used in the econometric analysis, the year to which the observation for each specific country refers to and the respective income classification group according to the World Bank where "L" refers to low-income countries (shown in red), "M_L" refers to lower-middle income countries (shown in yellow) and "M_H" refers to upper-middle income countries (shown in green).
Source: author.

Table A.2 – Variables and sources

Variable	Definition	Source
H	Average years of schooling of the population between 25 and 74 years old	UNESCO Institute for Statistics (2019) and Barro-Lee educational attainment dataset available in http://barrolee.com/
Q	Proportion of the total population living in sub-standard housing (%)	Own computations based on data from UNESCO Institute for Statistics (2019). Available in http://data.un.org/Data.aspx?d=POP&f=tableCode%3a309
GDP	Gross domestic product *per capita* (in 2011 USD purchasing power parity)	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: Made by the author

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

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

Source: Made by the author

Table A.4 - United Nations definitions for housing quality classification

Classifications	Definitions
Housing unit	Refers to a separate and independent place of abode intended for habitation by a single household, or one not intended for habitation but occupied as living quarters by a household at the time of the census. (p. 249)
CONVENTIONAL DWELLINGS	
Has all the basic facilities	A conventional dwelling that has all basic facilities refers to a unit that meets all the needs of the household within its confines, such as having piped water, a toilet, fixed bath or shower and a kitchen or other space for cooking. (p. 251)
Does not have all basic facilities	The conventional dwellings that fall in this category are dwellings that may have some, but not all, basic facilities (p. 251)
OTHER HOUSING UNITS	
Informal housing unit	Refers to a unit that does not have many of the features of a conventional dwelling and is generally characterized as unfit for human habitation, but that is used for that purpose at the time of the census. (p. 253)
Semi-permanent housing unit	Refers to a structure that is not expected to maintain its durability for as long as a conventional dwelling relative to each country's standards and practices. (p. 251)
Mobile housing unit	Refers to any type of living accommodation that has been produced to be transported (p. 252)
Improvised housing unit	Refers to an independent, makeshift shelter or structure, built of waste materials and without a predetermined plan for the purpose of habitation by one household. (p. 253)
Housing units in permanent buildings not intended for human habitation:	Included in this category are housing units (in permanent buildings) that have not been built, constructed, converted or arranged for human habitation but that are actually in use as living quarters at the time of the census. This category may also cover units and their occupants in buildings initially built for human habitation but later abandoned with all services cut because of deterioration. (p. 253)
Other informal housing units	This category refers to living quarters that are not intended for human habitation or located in permanent buildings but that are nevertheless being used as living quarters at the time of the census. Caves and other natural shelters fall within this category. (p. 253)

Source: Principles and Recommendations for Population and Housing Censuses (2017, p. 249-253)

Table A.5 – Descriptive statistics for the control variables

	Mean	Median	Minimum	Maximum	Std. Dev.	C.V.	Skewness
GDP	10185	9247.3	860.46	23326	6729	0.6606	0.1544
gov_edu	4.6003	4.2339	1.0997	13.858	2.0137	0.43774	2.0853
Mortality	37.08	23.23	6.968	123.48	34.766	0.9376	1.2473
Internet	20.962	17.1	0.225	69.75	18.015	0.85944	0.575
Ln_GDP	8.8985	9.1321	6.7575	10.057	0.9369	0.1053	-0.7370
Ln_gov_edu	1.4454	1.4431	0.0950	2.6288	0.4076	0.2820	-0.2621
Ln_mortality	3.2159	3.1454	1.9413	4.816	0.8922	0.2774	0.3363
Ln_Internet	2.3243	2.837	-1.4917	4.2449	1.5468	0.6654	-0.9022

Notes: The first four lines has the descriptive statistics for the variables in levels and the last four for the variables in natural logarithm. Source: author computations.