ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS
ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
SCHOOL OF ECONOMIC SCIENCES
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΕΠΙΣΤΗΜΗΣ
DEPARTMENT OF ECONOMICS

Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 17-2022

Human capital threshold effects in economic development: A panel data approach with endogenous threshold

Dimitris Christopoulos, Dimitris Smyrnakis and Elias Tzavalis

October 2022

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

Human capital threshold effects in economic development: A panel data approach with endogenous threshold

Dimitris Christopoulos[*], Dimitris Smyrnakis[†] and Elias Tzavalis[‡]

October 7, 2022

Abstract

We investigate for human capital threshold effects on economic growth dealing with the issue of endogeneity of the threshold variable. We consider two measures of human capital: (i) the education level and (ii) an extended measure of it accounting for the health level. We provide clearcut evidence of significant human capital threshold effects on economic growth distinguished between low and high levels, which supports the club convergence hypothesis within different levels of human capital. For the higher human capital regime, we find positive and strong relationship, while for the lower the relationship is negative mainly due to poor health conditions.

Keywords: Club Convergence, Human Capital, Dynamic Panel Threshold Model, Copulas, FD-GMM

JEL classification: C23, C24, I00, O47

[*]Department of International and European Economic Studies, Athens University of Economics and Business, email: dchristop@aueb.gr

[†]Department of Economics, Athens University of Economics and Business, email: smyrnakisd@aueb.gr

[‡]Department of Economics, Athens University of Economics and Business, email: e.tzavalis@aueb.gr, tel.: (+30) 210 8203 332

We would like to thank Costas Azariadis, Elias Papaioannou, Thanasis Stengos and Jonathan Temple for their comments and suggestions. Also, we thank participants at the 18th Conference on Research on Economic Theory and Econometrics (2019), for helpful comments on an earlier version of the paper, and Yiannis Karavias for his assistance with coding issues. We acknowledge financial support by the Hellenic Foundation for Research and Innovation under the "First call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant" (project no: HFRI-FM17-3532).

1 Introduction

The vast majority of the empirical growth literature regarding the relationship between human capital and economic growth, focuses on whether it is the level or change of human capital that can be considered as a fundamental source of growth. The results vary substantially in that others report a positive and significant role of human capital on growth (see, e.g., Barro (1991), Mankiw et al. (1992) and Bernanke and Gürkaynak (2001)), while others report insignificant or even negative effects (see, e.g., Benhabib and Spiegel (1994), Islam (1995), Caselli et al. (1996) and Barro (2013a)). Temple (1999) argues that the negative or insignificant effect may be attributed to the heterogeneity of countries used in the empirical studies.[1] What is common in all the aforementioned studies is the linear specification of a convergence regression employed for the testing procedure, accompanied by the common belief that a negative coefficient estimate regarding initial income can be interpreted as evidence of (conditional) convergence towards a common equilibrium.

Since the seminal work of Azariadis and Drazen (1990), introducing a threshold model of human capital externalities and economic growth implying multiple locally stable multiple equilibria, there is growing interest to empirically examine the existence of threshold, or non-linear, effects in the relationship between growth and human capital.[2] Pioneering works in this area include the following: Durlauf and Johnson (1995), using a cross-section regression tree methodology, find significant evidence favouring the multiple equilibria perspective. Hansen (2000), in an illustration of his threshold regression methodology, find similar results. Kalaitzidakis et al. (2001) and Savvides and Stengos (2008), based on earlier work by Liu and Stengos (1999) also conclude favourably towards a non-linear nexus, even after breaking down educational attainment by gender

[1] Other examples of negative effects of human capital on growth include Teulings and van Rens (2008) and Pelinescu (2015).

[2] Note that Azariadis and Drazen (1990) also present preliminary evidence favouring the multiple equilibria perspective to back their theory.

and schooling level.[3,4]

One important issue, however, concerning the estimation of both the panel or cross-section growth threshold model is the endogenous nature of the threshold variable (i.e., its contemporaneous correlation with the error term), which if is ignored can lead to biased estimates of the threshold model parameters including the threshold parameter. As noted in many recent studies (see the comprehensive surveys of Durlauf et al. (2005, 2009), even one can find strong instruments to deal with this issue, these can turn out to be invalid as they can be correlated with the error term, and thus leading to inconsistent estimates. This can be attributed to the fact that there is not one direction causality between growth and its related variables; they are both contemporaneously determined.

In this paper, we investigate for threshold effects in the relationship between economic growth and human capital suggesting a new method to deal with the problem of the endogeneity of human capital based on copula theory, as in Park and Gupta (2012) and Christopoulos et al. (2021). Our method does not require valid instruments for the threshold variable which are difficult to find. The copula theory enables us to control for the endogeneity between the threshold variable and regression error term by adjusting the rhs of the threshold regression by copula-type transformations of the threshold variable, which are orthogonal to the error terms. Furthermore, our method can work satisfactorily in cases where there is a non-linear dependence structure between the threshold variable and the error term, as well as in cases where the threshold variable is not normally distributed.[5]

[3]Further work, similar to Kalaitzidakis et al. (2001), includes Ketteni et al. (2007, 2011) and Kottaridi and Stengos (2010), who confirm the non-linear relationship between growth and human capital after controlling for Information and Communication Technology (ICT) and Foreign Direct Investment (FDI) effects.

[4]Indirect evidence of threshold effects in the relationship between growth and human capital is also provided by the work of Soukiazis and Cravo (2008) in a panel of countries and report variations on the effect both among levels of education and between different samples. A similar approach is followed by Lau et al. (1993), Lau (2010) and Cravo and Soukiazis (2011) who also conclude on a threshold effect of human capital on growth.

[5]A first effort to deal with the problem of endogeneity was made by Pelloni et al. (2019), assuming that institutional quality constitute valid instrument for human capital. However, there is evidence

To eliminate the problem of individual effects type of heterogeneity, we implement our method to the first-difference generalized method of moments (FD-GMM) estimation approach for dynamic panels modified to allow for threshold effects.[6] According to Arellano and Bond (1991), we employ as instruments lagged values of the regressors and the dependent variables which are uncorrelated with the error term. Note that our econometric framework can be also seen as an extension of the dynamic threshold model of Seo and Shin (2016), by using control variables to deal with the endogeneity of the threshold variable, instead of additional instruments for it. This framework can take into account heterogeneity both across individuals and over time. Evidence of such sources of heterogeneity are documented in recent studies (see Johnson and Papageorgiou (2020) for a survey).

To estimate our model, we use annual frequency of the data which can indicate the most likely year of transition across the low and high human capital regimes considered, for each country (if any). The time dimension of the panel covers a long span of data form 1965 to 2017, during which many countries have undertaken various educational reforms. As human capital variable, we use an index which has been recently available within the Penn World Table which could be considered as a more representative and informative measure of education policy. This is due to the fact that this index is constructed upon average years of total schooling (which include all levels of education) and returns to education suggested by Psacharopoulos (1994). Caselli (2005) provides an economic reasoning for using this measure of human capital; this can be attributed to the log-linear relationship between wage and schooling due to perfect competition in factor and good markets.

As education represents only one factor constituting human capital we extend the study to include the health component as well, which has been recognized as an im-

questioning this assumption (see, e.g., Aghion et al. (2004), Acemoglu et al. (2005) and Pargianas (2017).)

[6] Caselli et al. (1996) used this (no-threshold) framework to test for conditional convergence, for the linear growth equation.

portant factor for economic growth complementary to education.[7,8] To measure it, we construct an index in a similar way the educational component is created. We do so to create a linkage with productivity, analogous to that of the educational human capital measure we employed.

A number of interesting conclusions can be drawn from the results of the paper. We provide clear-cut evidence of significant human capital threshold effects, confirmed by a number of tests carried out in our study. We find that controlling for endogeneity plays a significant role in distinguishing the effects of human capital on economic growth. Our results support the convergence club hypothesis and clearly distinguish two human capital regimes: the high and low. The threshold value which identifies these regimes is found to be 1.719 when education is used as threshold variable, which corresponds to 4 years of schooling and a rate of return 13.4%. When a combination of education and health is used as threshold variable, the corresponding value is 2.852, which corresponds to a survival rate 77.5%, if one assumes 4 years of schooling and education return 13.4%.

According to the above regimes, we can categorise the countries as follows: a first group of countries characterised by high levels of human capital and a second one characterised by low levels of human capital. For the first group of countries, we found significant and large in magnitude effects of human capital on growth. For the second group, we find that these effects are insignificant; the fundamental variables positively affecting growth are the rate of investment and population growth. Finally, there is a third group of countries (almost half of the sample) that switch groups from the low to

[7]The importance of health for economic growth is stressed in recent reports by the World Health Organisation and the World Bank (World Health Organization and World Bank (2017) and World Bank (2018)) which highlight the lack of adequate health services in low income countries, as well as a series of studies examining the impact of improvements in health conditions to earnings and costs (see, e.g., Chuma et al. (2006), Alonso et al. (2019), Gallup and Sachs (2001), Sarma et al. (2019), Ahuja et al. (2015) and Dillon et al. (2021).

[8]The complementarity between education and health is summarized by the WHO Regional Office for Europe (2020) report in the following channels: (i) the level of education is a key determinant of health, and in particular the quality of health services offered; (ii) improving the foundations that make people healthier and more resilient, and (iii) enhance people's capacity to access, understand and use information in order to improve their health. See also case studies by Lucas (2010), Belot and James (2011) and Sandjaja et al. (2013) among others.

the high human capital during our sample. We demonstrate that the above groups of countries are closely related to the income group classification reported by the World Bank. Inclusion of health human capital in the regressions does not essentially change the resulting groups. There are, however, some important differences in the year the switch occurs, when combining educational and health human capital into one measure, especially for sub-Saharan African countries, which moved to the higher human capital group when the health conditions improved. This result highlights the importance of improving health conditions in poor countries.

The rest of this paper is organised as follows. In Section 2, we present the theoretical human capital threshold model. In Section 3, we present the empirical methodology and a (limited) Monte Carlo study assessing its performance. In section 4, we present the data and the estimation results. Finally, Section 5 concludes the paper.

2 Model setup

Our starting point is the work of Islam (1995) and Caselli et al. (1996). The former reformulates the convergence equation obtained by Mankiw et al. (1992) (MRW hereafter) based on the human capital augmented Solow model, into a panel data model with individual effects and uses the LSDV and Minimum Distance estimators. To avoid possible estimation biases of these estimators, Caselli et al. (1996) propose the use of the Arellano and Bond (1991) estimator.

The corresponding convergence equation considered by Islam between two successive periods $t-1$ and t is:[9]

$$\begin{aligned}\ln(y_t) =& e^{-\lambda\tau}\ln(y_{t-1}) + (1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(s_t) - (1-e^{-\lambda\tau})\frac{\alpha}{1-\alpha}\ln(n_t+g+\delta)\\ &+ (1-e^{-\lambda\tau})\frac{\beta}{1-\alpha}\ln(\hat{h}^*) + (1-e^{-\lambda\tau})\ln(A_0) + g(t-e^{-\lambda\tau}(t-1))\end{aligned} \tag{1}$$

[9] Equation (1) is the unrestricted version of equation (18) in Islam (1995) for $t_2 = t$ and $t_1 = t-1$.

where $\hat{h}^*$ denotes the human capital (per efficient unit of labour) in the Balanced Growth Path, n_t, and s_t denote population growth and saving rate respectively between the periods $t-1$ and t, τ is the actual time distance between the two successive periods $t-1$ and t, e.g. $\tau = 1$ for annual data, and we have suppressed the subscript i denoting an individual country.[10] g and δ denote the rate of technological progress and depreciation rate respectively, which are often considered constant both across i and t (see, e.g., Mankiw et al. (1992) and Islam (1995)). Since $\hat{h}_t$ is defined as:

$$\hat{h}_t \equiv \frac{H_t}{A_t L_t}, \tag{2}$$

it holds that:

$$\hat{h}_t = \frac{h_t}{A_t} \tag{3}$$

where $h_t \equiv \frac{H_t}{L_t}$. Substituting equation (3) into (1), and after some necessary rearrangements we obtain:

$$\begin{aligned} \ln(y_t) =& e^{-\lambda\tau} \ln(y_{t-1}) + (1 - e^{-\lambda\tau}) \frac{\alpha}{1-\alpha} \ln(s_t) - (1 - e^{-\lambda\tau}) \frac{\alpha}{1-\alpha} \ln(n_t + g + \delta) \\ &+ (1 - e^{-\lambda\tau}) \frac{\beta}{1-\alpha} \ln(h_t) + (1 - \frac{\beta}{1-\alpha}) \left[(1 - e^{-\lambda\tau}) \ln(A_0) + g(t - e^{-\lambda\tau}(t-1)) \right] \end{aligned} \tag{4}$$

In the above specification we add a threshold property, as in Durlauf and Johnson (1995)

[10] Note that equation (1) is equivalent to

$$\begin{aligned} \ln(y_t) - \ln(y_{t-1}) =& -(1 - e^{-\lambda\tau}) \ln(y_{t-1}) + (1 - e^{-\lambda\tau}) \frac{\alpha}{1-\alpha} \ln(s_t) - (1 - e^{-\lambda\tau}) \frac{\alpha}{1-\alpha} \ln(n_t + g + \delta) \\ &+ (1 - e^{-\lambda\tau}) \frac{\beta}{1-\alpha} \ln(\hat{h}^*) + (1 - e^{-\lambda\tau}) \ln(A_0) + g(t - e^{-\lambda\tau}(t-1)) \end{aligned}$$

which corresponds to the MRW convergence equation variant that includes the stock measure of human capital. The choice is driven by the human capital measures of our dataset that represent stocks rather than flows.

and Hansen (2000), such that

$$\beta = \begin{cases} \beta_1, \text{ if } h_t < \gamma \\ \beta_2, \text{ if } h_t \geq \gamma \end{cases} \tag{5}$$

Obviously, this threshold property will directly affect the slope of the human capital regressor, yet, it will also affect indirectly all others through the convergence rate λ which is a function of β, yielding the following threshold model:

$$\ln(y_t) = \begin{cases} \theta_1^{(1)} \ln(y_{t-1}) + \theta_2^{(1)} \ln(s_t) + \theta_3^{(1)} \ln(n_t + g + \delta) + \theta_4^{(1)} \ln(h_t) + \theta_5^{(1)}, & \text{if } h_t < \gamma \\ \theta_1^{(2)} \ln(y_{t-1}) + \theta_2^{(2)} \ln(s_t) + \theta_3^{(2)} \ln(n_t + g + \delta) + \theta_4^{(2)} \ln(h_t) + \theta_5^{(2)}, & \text{if } h_t \geq \gamma \end{cases} \tag{6}$$

where $\theta_1^{(j)} = e^{-\lambda_j \tau}$, $\theta_2^{(j)} = (1 - e^{-\lambda_j \tau})\dfrac{\alpha}{1-\alpha}$, $\theta_3^{(j)} = -(1 - e^{-\lambda_j \tau})\dfrac{\alpha}{1-\alpha}$ and $\theta_4^{(j)} = (1 - e^{-\lambda_j \tau})\dfrac{\beta_j}{1-\alpha}$, for $j = \{1, 2\}$.

The dynamic panel data econometric specification of equation (6) is given as follows

$$\ln(y_{i,t}) = \begin{cases} \theta_1^{(1)} \ln(y_{i,t-1}) + \theta_2^{(1)} \ln(s_{i,t}) + \theta_3^{(1)} \ln(n_{i,t} + g + \delta) + \theta_4^{(1)} \ln(h_{i,t}) + \eta_i + v_{i,t}, \\ \hfill \text{if } h_{i,t} < \gamma \\ \theta_1^{(2)} \ln(y_{i,t-1}) + \theta_2^{(2)} \ln(s_{i,t}) + \theta_3^{(2)} \ln(n_{i,t} + g + \delta) + \theta_4^{(2)} \ln(h_{i,t}) + \eta_i + v_{i,t}, \\ \hfill \text{if } h_{i,t} \geq \gamma \end{cases} \tag{7}$$

where $i = 1, ..., N$ denotes countries. The error components η_i and v_{it} in model (7) represent individual (country specific) effects associated with A_0 and zero mean transitory error term effects that vary both across countries and time (see Islam (1995) for a discussion).

Estimation of model (7) to examine the existence of a non-linear relationship between human capital and economic growth has the following interesting properties: Firstly, the evidence of threshold effects imply the need for large scale policy interventions in order to close the income gap across countries. Secondly, the panel data nature of the models, by combining cross-sectional with time dimension information of the data, can more efficiently identify threshold effects from the data as it increases the number of observations in each regime of the model. Thirdly, the time-dimension of the data can estimate more efficiently the convergence rates implied by the growth equation and can indicate the most likely date of transition across the two regimes, for each country (if any). Finally, it can account for time heterogeneity in the growth-human capital nexus.

3 Empirical methodology

For notational generality, we denote the vector of explanatory variables with x_{it}. Model (7) can be estimated using the following dynamic panel threshold model, written in vector notation:

$$y_{it} = x_{it}'\beta_1 I\{q_{it} \leq \gamma\} + x_{it}'\beta_2 I\{q_{it} > \gamma\} + \eta_i + v_{it}, \qquad i = 1, ..., N, t = 1, ..., T \tag{8}$$

where y_{it} is the dependent variable, x_{it} represents a $K \times 1$ vector of regressors, which include the lagged dependent variable and a constant, i.e. $x_{it} = (1, y_{it-1}, x_{1,it}, ..., x_{K-2,it})'$, and $I\{\cdot\}$ is an indicator function that captures the regime-switching mechanism based on a transition variable q_{it}, while γ is a location parameter that defines the regime switch. Given a value of γ we can define two partitions (regions) of q_{it}: $Q_{1t} = \{-\infty < q_{it} \leq \gamma\}$ and $Q_{2t} = \{\gamma < q_{it} < +\infty\}$, representing the two distinct regimes of the model, with $\beta_1 \neq \beta_2$, for all i and t.[11] Finally, $\eta_i + v_{it}$ represent the error components, where η_i

[11] A more general specification of the model (8) would be

$$y_{it} = x_{1,it}'\beta_0 + x_{2,it}'\beta_1 I\{q_{it} \leq \gamma\} + x_{2,it}'\beta_2 I\{q_{it} > \gamma\} + \eta_i + v_{it}$$

denotes an unobserved individual effect and v_{it} is a zero mean idiosyncratic error term, such that $E(v_{it}|\mathcal{F}_{t-1}) = 0$, which allows for endogeneity in both the regressors and the threshold variable, that is $E(x_{it}v_{is}) \neq 0$ and/or $E(q_{it}v_{is}) \neq 0$, for $s < t$.[12]

To deal with the problem of endogeneity between the threshold variable q_{it} and the error term v_{it}, we can extend the copula approach suggested recently by Christopoulos et al. (2021) to dynamic panel threshold models. This method adjusts the conditional mean of y_{it} on q_{it} for the conditional expectations of the error term, v_{it}, given Q_{1t} or Q_{2t}, that is

$$E(y_{it}|x_{it}, q_{it}) = \begin{cases} x'_{it}\beta_1 + E(v_{it}|Q_{1t}), & \text{if } q_{it} \leq \gamma \\ x'_{it}\beta_2 + E(v_{it}|Q_{2t}), & \text{if } q_{it} > \gamma \end{cases} \tag{9}$$

for all i. Equation (9) shows that in order to consistently estimate the parameters of the regression model (8), it is necessary to control for the effects of the conditional expectations $E(v_{it}|Q_{1t})$ and $E(v_{it}|Q_{2t})$. Otherwise, both the estimates of the threshold parameter γ and slope coefficients β_1 and β_2 can be seriously biased.

To calculate the conditional means $E(v_{it}|Q_{1t})$ and $E(v_{it}|Q_{2t})$, we can rely on Copula theory. This theory enables us to capture the dependence between v_{it} and the threshold variable, q_{it}, across the two regimes "1" and "2" by decomposing the joint distribution of v_{it} and Q_{jt}, $j = 1, 2$ into a part that captures the dependence structure between them based on a copula C and the marginal distribution of the threshold variable, q_{it}, across

so that it also includes regressors that remain unaffected by the threshold, yet, without affecting the estimation procedure. This specification allows for the dynamic term and the intercept to be included in $x_{1,it}$, so that the threshold does not affect them. Note also that the transition/threshold variable may also be one of the regressors included in x_{it}.

[12]Usually in the panel data literature regressors are considered either *strictly exogenous*, i.e. $E(x_{it}v_{is}) = 0$ $\forall t, s$, or *predetermined*, i.e. $E(x_{it}v_{is}) \neq 0$ for $s \leq t$. The latter case allows regressors to be correlated with past values of the structural error, but not present or future ones. By setting $E(x_{it}v_{is}) \neq 0$ for $s < t$, we allow for contemporaneous correlation as well, and there is good reason to do so, as it may be the case that some variables may have been measured with error, or variables omitted from the regression are potentially correlated with the ones included. Technically this translates as x_t and x_{t-1} are no longer valid instruments for Δx_t.

the two regimes. This can be done without assuming any distribution for the threshold variable. Furthermore, the use of copula C can approximate for a non-linear structure of dependence between q_{it} and v_{it}, which can also change across the two regimes $j = 1, 2$.

Next, we derive analytic forms of the conditional means $E(v_{it}|Q_{jt})$, $j = 1, 2$. Consider two copula functions C_j where $C_j : [0, 1]^2 \to [0, 1]$, $j = 1, 2$ defined as follows

$$C_j(F_v(v_{it}), F_{q|Q_j}(q_{it}|Q_{jt})), \qquad j = 1, 2 \tag{10}$$

where $F_v(\cdot)$ denotes the cumulative distribution function of v_{it}, and $F_{q|Q_j}(\cdot)$ is the truncated cumulative distribution function of q_{it}, for $j = 1, 2$, such that

$$F_{q|Q_j}(q_{it}|Q_{jt}) = \begin{cases} \dfrac{F_q(q_{it})}{F_q(\gamma)}, & \text{if } q_{it} \leq \gamma \\ \dfrac{F_q(q_{it}) - F_q(\gamma)}{1 - F_q(\gamma)}, & \text{if } q_{it} > \gamma \end{cases} \tag{11}$$

where $F_q(\cdot)$ denotes the cumulative distribution function of q_{it} and $F_q(\gamma) = P(q_{it} \leq \gamma)$. The truncated distributions $F_{q|Q_j}(q_{it}|Q_{jt})$ are scaled so that they constitute proper distributions integrating to 1 as is required by the definition of the C_j copulas, defined above.

Given C_j, we can define the conditional distribution v_{it} conditional on Q_j by differentiating it with respect to $F_{q|Q_j}(\cdot)$ as follows:[13]

$$F_{v|Q_j}(v_{it}|Q_{jt}) = \frac{\partial}{\partial F_{q|Q_j}} C_j(F_v(v_{it}), F_{q|Q_j}(q_{it}|Q_{jt})) \tag{12}$$

From this distribution, we can obtain the conditional probability density function by

[13] See, e.g., Erdely (2017). The copula functions $C_j, j = 1, 2$, can be thought of as scaled functions of gluing copulas along the support interval of Q_{1t} and Q_{2t}.

differentiating it with respect to v_{it}, that is

$$
\begin{aligned}
f_{v|Q_j}(v_{it}|Q_{jt}) &= \frac{\partial^2}{\partial v_{it}\partial F_{q|Q_j}} C_j(F_v(v_{it}), F_{q|Q_j}(q_{it}|Q_{jt})) \\
&= c_j(F_v(v_{it}), F_{q|Q_j}(q_{it}|Q_{jt}))f_v(v_{it}) \qquad (13)
\end{aligned}
$$

where $c_j(\cdot)$ denotes the copula density function corresponding to $C_j(\cdot)$ and $f_v(v_{it}) = \frac{\partial}{\partial v_{it}}F_v(v_{it})$ denotes the marginal probability density of v_{it}.

If we consider that C_j are Gaussian Copulas and $v_{it} \sim N(0, \sigma^2)$, it can be shown that the error term v_{it} has the following single factor representation:

$$
v_{it} = \lambda_j q_{it}^{(j)*} + Var(v_{it}|q_{it}^{(j)*})^{1/2}\varepsilon_{it} \qquad (14)
$$

where $q^{(j)*}$ represents the Copula transformation $q_{it}^{(j)*} = \Phi^{-1}(F_{q|Q_j}(q_{it}|Q_{jt}))$, where Φ^{-1} is the quantile function of the Standard Normal distribution, $\lambda_j = \sigma\rho_{v,q^{(j)*}}$, $\rho_{v,q^{(j)*}}$ denotes the Pearson correlation coefficient between v and $q^{(j)*}$, $Var(u_{it}|q_{it}^{(j)*}) = \sigma^2(1 - \rho^2_{v,q^{(j)*}})$ is the conditional variance of v_{it} on $q^{(j)*}$ and, finally, $\varepsilon_{it} \sim iidN(0,1)$ is independent of $q_{it}^{(j)*}$ and q_{it}.[14]

From relationship (14), we can easily see that the conditional mean of v_{it} on Q_{jt} is given as $E(v_{it}|Q_{jt}) = \lambda_j q_{it}^{(j)*}$ and equation (9) implies the following augmented threshold regression model:

$$
y_{it} = (x'_{it}\beta_1 + \lambda_1 q_{it}^{(1)*})I\{q_{it} \leq \gamma\} + (x'_{it}\beta_2 + \lambda_2 q_{it}^{(2)*})I\{q_{it} > \gamma\} + \eta_i + e_{it} \qquad (15)
$$

where $e_{it} = \left[Var(v_{it}|q_{it}^{(1)*})^{1/2}I(q_{it} \leq \gamma) + Var(v_{it}|q_{it}^{(2)*})^{1/2}I(q_{it} > \gamma)\right]\varepsilon_{it}$. The regression model (15) extends model (8) to control for the threshold variable endogeneity effect.[15]

[14] For a proof, see Joe (2014), or Christopoulos et al. (2021).

[15] Note that our approach shares some similarities to the binary choice selectivity models (known as endogenous switching models) - see, e.g., Vella (1998), for a comprehensive survey. See also Lee (1983), who considers similar transformations to ours of the error term of the selection equation of the censored model (referred to as strictly increasing transformations) to adjust the structural equation of the model

The variables $q_{it}^{(j)*}$, $j = 1, 2$, can be thought of as control functions added to the model (15) to render the threshold variable q_{it} (which play the role of the conditional state variable) appropriately exogenous with respect to the new error term e_{it}. As noted before, in constructing the variables $q_{it}^{(j)*}$, $j = 1, 2$ we make no assumption about the distribution of the threshold variable q_{it}. We can estimate the marginal distribution of q_{it} non-parametrically, in a first step.[16]

We can employ this augmented model to obtain more accurate estimates of the parameters of the model if there is a problem of weak and/or invalid instruments of the threshold variable q_{it}. Furthermore, this problem becomes more serious when a large set of instruments is used (see, e.g., Zhang and Zhou (2020) for a recent survey). Such problems may arise especially in long panels when the number of moment conditions becomes very large (see, e.g., Roodman (2009)).

The first difference GMM estimator (Arellano-Bond) for model (15) relies on the first difference transformation of (15) given as follows:

$$\begin{aligned} \Delta y_{it} =& b' \Delta x_{it} + \delta' X_{it}' 1_{it}(\gamma) \\ &+ \lambda_1 (q_{it}^{(1)*} I\{q_{it} \leq \gamma\} - q_{it-1}^{(1)*} I\{q_{it-1} \leq \gamma\}) \\ &+ \lambda_2 (q_{it}^{(2)*} I\{q_{it} > \gamma\} - q_{it-1}^{(2)*} I\{q_{it-1} > \gamma\}) + \Delta e_{it} \end{aligned} \tag{16}$$

where $\delta = \beta_2 - \beta_1$, $b = (\beta_{12}, ..., \beta_{1K})'$, $X_{it} = \begin{pmatrix} x_{it}' \\ x_{it-1}' \end{pmatrix}$ and $1_{it} = \begin{pmatrix} I\{q_{it} > \gamma\} \\ -I\{q_{it-1} > \gamma\} \end{pmatrix}$. If $\lambda_1 = \lambda_2 = 0$, this representation reduces to that of Seo and Shin (2016), that is

$$\Delta y_{it} = \beta' \Delta x_{it} + \delta' X_{it}' 1_{it}(\gamma) + \Delta v_{it} \tag{17}$$

for the selectivity bias problem without relying on a specific distribution of the selection equation error.

[16] The above approach of augmenting the threshold model (5) is in the spirit of Kourtellos et al. (2016), who assume joint normality of the error term and the threshold variable. To relax the assumption of normality for q_{it}, recently, Kourtellos et al. (2021) proposed a semi-parametric approach.

Let the vector of the parameters of equation (16) be $\theta = (b', \delta', \lambda_1, \lambda_2, \gamma)'$. This vector can be estimated by the FD-GMM estimator based on equation (16) relying on the following moment conditions:

$$E(\Delta e_{it}|z_{it}) = 0, \quad \text{for} \quad t = t_0, ..., T \tag{18}$$

where the vector of instruments z_{it} can include lagged values of y_{it} and x_{it}. The use of the first differenced equation (16), instead of (17), allows us to not use lagged values of q_{it} as instruments to account for the endogeneity between the threshold variable and the error term, given that e_{it} is orthogonal to both q_{it} and its copula transformed variables $q_{it}^{(j)*}$. Thus, the FD-GMM estimator based on (16) can also mitigate any weak instruments problem related to the threshold variable, especially in the case of too many instruments.

3.1 Estimation and Monte Carlo results

If the threshold parameter, γ, is known, estimation of regression model (16) can be straightforwardly estimated by the FD-GMM procedure. If it is unknown, the estimation can proceed in two steps by concentrating the GMM function. In the first step, we can retrieve estimates of γ by solving the following problem through a grid search procedure:

$$\hat{\gamma} = arg \min_{\gamma \in \Gamma} J(\gamma) \tag{19}$$

where $J(\gamma)$ is the GMM function corresponding to the sample moment conditions of (18) for a given value of $\gamma \in \Gamma$, where $\Gamma = \left[\underline{\gamma}, \bar{\gamma}\right]$ is the sample set of all possible values of q_{it} trimmed at a lower and upper bound to ensure a sufficient number of data points for the estimation of the equation (16) across the two regimes. Given $\hat{\gamma}$, we can obtain estimates of vector of the remaining parameters of θ, $(b', \delta', \lambda_1, \lambda_2)'$. Seo and Shin (2016) show that the above estimation procedure leads to consistent estimates of γ and $(b', \delta', \lambda_1, \lambda_2)'$, which is asymptotically normally distributed.

As a final, note that, in the above estimation procedure in order to derive values of $q_{it}^{(j)*} = \Phi^{-1}(F_{q|Q_j}(q_{it}|Q_{jt}))$, for $j = \{1, 2\}$, we estimate the distribution $F_{q|Q_j}(q_{it}|Q_{jt})$ based on a non-parametric procedure. We can obtain estimates of the cumulative distribution function $F_{q|Q_j}(q_{it}|Q_{jt})$ by employing a Silverman's Kernel estimator (Silverman (1986)), or by using the empirical cumulative distribution function, ECDF (Rice (2007), Amengual and Sentana (2020)).[17] The Glivenko - Cantelli theorem establishes that these estimates almost surely converge to the true cumulative distribution function (see Glivenko (1933) and Cantelli (1933)).

To evaluate the performance of the proposed estimation method, next, we carry out a Monte Carlo study. The details of the study are given in Appendix A.2. We consider various scenarios and combinations of parameters and distributions. Yet, due to reasons of space we present a limited set of results. In particular, we consider (i) the case of weak instruments for both the explanatory and the threshold variables, (ii) the case of weak instruments only for the explanatory variables, and (iii) the case that the instruments for the threshold variable, q_{it}, are invalid. For all the above simulation scenarios, we compare the performance of our method to Seo and Shin's, henceforth denoted S-S. Our method can be thought of as a useful extension of the S-S estimation method in cases when it's hard to find appropriate instruments. For both the S-S and our approach (denoted as CST) of the FD-GMM estimation we restrict the number of instruments to the second lag of the regressors and the threshold variable, such that the estimator to be consistent regardless of the size of T (see, e.g., Hsiao and Zhang (2015)).

The results of the Monte Carlo exercise, reported in Appendix A.2, indicate that our method constitutes a refined extension of the S-S method. It can efficiently deal with the problem of endogeneity of the threshold variable either for short or long panels.

[17] The empirical CDF is defined as

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x)$$

though some authors suggest dividing the sum by $n + 1$ instead of n.

Our approach reduces substantially the bias and root mean square error (RMSE) of the parameters of the threshold model, including the threshold parameter γ. Most of the gains are found for the case of invalid instruments (simulation scenario (iii)), as was expected. Yet, for the simulation scenarios (i) and (ii), the gains are found to mainly concern the estimates of the slope coefficients of the explanatory variables.

4 Estimation

4.1 Data

Our dataset covers the period 1965-2017 using annual frequency, implying $T = 53$ observations, and include real GDP and gross capital formation as a share of GDP taken from the *Penn World Table v.9.1* (Feenstra et al. (2015)). As a measure of human capital, we employ a Human Capital Index that has become available since version 8 of the Penn World Table. The index is constructed upon average years of total schooling and returns to education, establishing a link between educational human capital and productivity (see, e.g., Caselli (2016)).[18] Regarding labour, we collected data from the World Bank's *Health, Nutrition and Population Statistics* database and *Taiwan's National Development Council* on population aged 15-64 to compute the respective growth rate n_{it}, as well as real GDP per working age person. Technological progress and depreciation rates, g and δ, are kept fixed at 0.02 and 0.03 respectively, as is usually assumed in the growth literature.

We also employ data on adult survival rates which we use to construct a measure of health human capital in a similar way the educational human capital from the Penn World Table is constructed. We do so, in order to obtain a measure that is also connected with productivity. These data are also taken from the World Bank's *Health, Nutrition*

[18] To maximize the set of countries included in the study we employ data on years of schooling from the educational attainment database of *Wittgenstein Centre for Demography and Global Human Capital* (Lutz et al. (2018)) and follow the procedure described in the Penn World Tables' manuals and Caselli (2005) to construct the human capital index for Cape Verde, Chad and Comoros.

and Population Statistics database and *Taiwan's National Development Council.*

We end up with a balanced panel of $N = 106$ countries for which all data are available and $T = 53$ time observations. From this sample, we have excluded China as a massively centrally planned economy, Lesotho because the sum of private and government consumption still exceeds GDP, indicating that labour income from abroad remains a large fraction of national income (see Mankiw et al. (1992)) and Iran as its economy is heavily based on oil exports. We also excluded Nicaragua, as gross capital formation to GDP in 1979 appears to be negative as a result of the Nicaraguan Revolution during 1978-1979, and any transformation we tried significantly altered the coefficient estimates for $\log(s_k)$ as well as its statistical significance.

Before proceeding to the estimation results, one final remark is in order. The choice of annual frequency is driven by the fact that a sufficiently large number of time series observations is important to identify the threshold effects and their dynamic influence on y_t, over the time dimension of the panel. According to our MC results, increasing the time dimension of the panel relative to N improves the accuracy of the estimates of the threshold parameter and the remaining parameters of the model. The high number of time series observations employed can also overcome possible common business cycle effects across countries (see also Cerra and Saxena (2008) and Christopoulos and León-Ledesma (2014)).

4.2 Estimation of the linear model

We start by estimating a restricted linear form of equation (7) such that there are no threshold effects, i.e. $\theta_l^{(1)} = \theta_l^{(2)}$, $l = 1, 2, 3, 4$. Table 1 presents the estimation results. The findings are rather interesting. Firstly, the autoregressive coefficient nearly equals one, which means that there is no tendency for poor countries to grow faster than rich countries.[19,20]

Table 1: Estimation results for equation (1)

$\ln(y_{it-1})$	0.995^{***}
	(0.005)
$\ln(s_{it})$	0.062^{***}
	(0.005)
$\ln(n_{it} + g + \delta)$	0.009
	(0.009)
$\ln(ehc_{it})$	-0.025^{***}
	(0.009)
J_{stat}	105.66

Notes: Dependent variable is real GDP per working age person (y_{it}). y_{it-1} represents the one period lagged value of output per working age person, and *ehc* stands for educational human capital. Notation on significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Secondly, the slope coefficient estimate of saving is highly significant at the 1% significance level and its sign is consistent with the prediction of the Mankiw et al. (1992) theory (see also Knight et al. (1993), Islam (1995) and Caselli et al. (1996)). On the other hand, the slope coefficient estimate of population growth is insignificant. Similar results with ours have been reported by Benhabib and Spiegel (1994), Knowles and Owen (1995) and Ram (2007).

Finally, and more interestingly, the slope of human capital possesses a negative sign and is highly significant at the 1% significance level. The latter is not uncommon in the growth and convergence literature and is often interpreted as evidence of a weak relationship between human capital and growth (see, e.g., Sala-i-Martin (2002)). This weakness could be attributed to ignoring the non-linear specification of the model and, more specifically, the existence of multiple regimes.

[19] see Mankiw et al. (1992), Islam (1995) and Caselli et al. (1996) for similar evidence.

[20] Actually, we have found that panel data unit root tests, namely the Harris and Tzavalis (1999) and the Levin et al. (2002), cannot reject the hull hypothesis of a unit root for the unconditional convergence hypothesis. The resulting p-values are found to be 0.938 and 0.209 respectively.

4.3 Estimation of the model using educational human capital

In Table 2, we present estimation results of the threshold regression model (7). We present estimates based on S-S FD-GMM dynamic panel threshold estimator as well as our proposed method, including the copula transformations of the threshold variable, $q_{it}^{(1)*}$ and $q_{it}^{(2)*}$, as control functions for its endogeneity. In the estimation of both methods, we used second order lagged values of the regressors, according to the Arellano-Bond procedure (as in our Monte Carlo exercise). For the S-S approach, we have chosen second lagged values of the threshold variable as instrument for it to mitigate possible endogeneity effects by using its lagged one value.

The p-values of the bootstrap LR test, reported in the table, support the existence of human capital threshold effects in model (4).[21] This test soundly rejects the null hypothesis of no threshold effects based on our approach. It is worth noting that, both $q_{it}^{(1)*}$ and $q_{it}^{(2)*}$ are significant at least at the 5% significance level, thus justifying the existence of threshold variable endogeneity effects.

The threshold estimate $\hat{\gamma}$ is identical for both methods, at a value of 1.719, lying at the 37.9% of the human capital distribution. This value corresponds to about 4 years of schooling, a value close to that reported by Lau et al. (1993), as well as Savvides and Stengos (2008). It also implies a rate of return on human capital investment of about 13.4%. It is worth noting, at this point, that the bootstrap confidence interval of γ, reported in the table, is tighter under our method, meaning that the above estimate of γ is more likely to reflect its true value. The above implied rate of return, 13.4%, based on the estimate of γ means that if a worker with no education earns \$1 and a worker with four years of schooling earns \$1.134 we could think of \$0.134 as being part of the wage attributed to human capital and the remaining \$1.00 due to raw labour.

For both human capital regimes identified by our method, we found that the estimates of the autoregressive coefficients $\theta_1^{(1)}$ and $\theta_1^{(2)}$ have significantly dropped from one,

[21] The test procedure is described in Appendix A.1.

Table 2: Estimation results for equation (7)

	S-S	CST
	Lower Regime	
$\ln(y_{it-1})$	0.909^{***}	0.925^{***}
	(0.012)	(0.013)
$\ln(s_{it})$	0.048^{***}	0.041^{***}
	(0.006)	(0.006)
$\ln(n_{it} + g + \delta)$	0.253^{***}	0.255^{***}
	(0.023)	(0.026)
$\ln(ehc_{it})$	-0.132	0.124
	(0.085)	(0.223)
	Upper Regime	
$\ln(y_{it-1})$	0.769^{***}	0.774^{***}
	(0.010)	(0.011)
$\ln(s_{it})$	0.172^{***}	0.176^{***}
	(0.011)	(0.011)
$\ln(n_{it} + g + \delta)$	-0.106^{***}	-0.123^{***}
	(0.018)	(0.018)
$\ln(ehc_{it})$	0.683^{***}	0.887^{**}
	(0.057)	(0.079)
$q_{it}^{(1)*}$		-0.050^{**}
		(0.024)
$q_{it}^{(2)*}$		-0.069^{***}
		(0.016)
$\hat{\gamma}$	1.719	1.719
95% C.I.	[1.213,2.722]	[1.586,1.857]
J_{stat}	104.95	102.52
$LR_{p-value}$	0.734	0.020

Notes: Dependent variable is real GDP per working age person (y_{it}). y_{it-1} represents the one period lagged value of output per working age person. In these regressions educational human capital (ehc) is considered as a regressor and the threshold variable. Notation on significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

and both of them are highly significant. These results imply that there exists conditional convergence within the groups of countries classified by the human capital level (which is our threshold variable). Regarding saving and population growth, we find that, within the lower regime, their effects are positive and highly significant at the 1% significance level, with the effect of population growth having a stronger effect than saving. These results indicate that labour quantity is more important for the low education group of countries than the rate of investment. Similar results have been reported for low development countries (LDCs) by Simon (1976, 1990).[22] On the contrary, within the upper regime, both saving and population growth are found significant at the 1% significance level, and possess the expected signs. Contrary to the lower human capital regime, saving appears more important than population growth, yet both are overshadowed by human capital, in terms of their slope coefficient estimates. The stronger effect of saving combined with faster conditional convergence within the upper regime, are consistent with evidence provided by Sirimaneetham and Temple (2009) based on a cross-section threshold model using as threshold variable the quality of macroeconomic policy. This can be related with the possible association between macroeconomic policy and the level of human capital.

Regarding human capital, the existence of threshold effects significantly alters the results compared with the linear model. Within the lower regime, we find an insignificant effect, implying that low levels in human capital investment have no impact on growth. The opposite happens within the upper human capital regime, where its slope coefficient is positive and significant at the 5% significance level. In fact, the standard errors reported in the table imply that the estimate of this coefficient based on our method is not significantly different than one, highlighting the importance of human capital on economic growth.

Based on the estimate of the threshold parameter γ, next we divide our sample in three categories. Within the lower human capital category lie countries that fall in the lower regime for the entire period studied. Similarly, the higher human capital category

[22] A positive effect of population on economic growth can be also justified within the endogenous growth framework, see e.g. Jones (1999), Mierau and Turnovsky (2014) and Bucci (2015).

consists of countries that lie within the upper regime for the entire period, while the intermediate group consists of countries which switch between the human capital regimes. Note that countries in the intermediate group switch only once from the lower to the higher human capital regime. Furthermore, this group includes a significant number of countries (24) which moved to the higher human capital regime after 1990 (near the middle of the time period considered), leaving a sufficient number of observations across both N, T dimensions of the panel to better identify the two human capital regimes from the data. These groups are reported in table 3.

Table 3: Country groups based on the h.c. threshold value $\hat{\gamma} = 1.719$

low h.c.	intermediate			high h.c.	
Burkina Faso	Algeria (1996)	Benin (2014)	Fiji (1971)	USA	Jamaica
Burundi	Bangladesh (2004)	India (1999)	Portugal (1983)	Australia	Japan
Central Africa	Honduras (1987)	Indonesia (1985)	Rwanda (2014)	Austria	Luxembourg
Chad	Guatemala (2009)	Jordan (1985)	Singapore (1982)	Barbados	Belgium
Ivory Coast	Botswana (1986)	Kenya (1991)	Sri Lanka (1967)	Canada	Hong Kong
D.R. Congo	Venezuela (1979)	Liberia (2010)	Syria (1989)	Chile	Netherlands
Ethiopia	Cape Verde (2002)	Malawi (2011)	Taiwan (1975)	Malta	New Zealand
Gambia	Cameroon (1999)	Malaysia (1978)	Thailand (1984)	Cyprus	Norway
Haiti	Colombia (1974)	Ghana (1986)	Togo (2002)	Denmark	Panama
Madagascar	Comoros (2002)	Egypt (1992)	Tunisia (1995)	Ecuador	South Korea
Mali	Congo R. (1989)	Mexico (1969)	Turkey (1986)	Finland	Romania
Niger	Costa Rica (1973)	Morocco (2009)	Uganda (2002)	France	Seychelles
Senegal	Dominican R. (1980)	Gabon (1989)	Brazil (1991)	Germany	South Africa
Sierra Leone	Mauritius (1979)	Namibia (1978)	Zambia (1986)	Greece	Argentina
Tanzania	El Salvador (1994)	Nepal (2016)	Bolivia (1974)	Iceland	Sweden
	Philippines (1971)	Nigeria (2009)		Ireland	Switzerland
	Myanmar (2010)	Pakistan (2004)		Israel	Trinidad & Tobago
	Mauritania (2014)	Paraguay (1975)		Italy	United Kingdom
	Zimbabwe (1986)	Peru (1973)		Spain	Uruguay

Notes: The year in parentheses for the intermediate group countries represents the year at which a regime switch occurs. Educational human capital is considered as the regime switching variable.

In Table 4, we present descriptive statistics for our income, saving and human capital variables used in estimation of the model for the countries as classified in Table 3, according to our model estimates. These statistics include the mean, standard deviation and the correlation coefficients for the above variables. The results of the table clearly indicate that the low human capital group consists of countries with the lowest income, saving and human capital. According to Azariadis and Drazen (1990), these countries can be considered as trapped in poverty, characterised by low saving and human capital investment. As can be seen by the table, these countries have a low correlation between income and human capital. On the other hand, the high human capital group consists of countries with the highest income, saving and human capital. These countries exhibit the strongest correlation between income and human capital.

Turning to the intermediate group of countries, transitioning to the high human capital regime, the results of the table indicate that income, saving and human capital lie in between the values of the low and high human capital groups. The same is true for the correlation coefficient between income and human capital. The high correlation for saving and human capital with income found for this group implies that these countries have invested both in physical and human capital, which allowed them to escape from the poverty trap.[23]

The income levels of the countries associated with the above human capital country groups are largely consistent with the World Bank's Income Group classification (2017), which divides income groups based on GNI per capita. All high income countries belong to the higher human capital group identified by our method, except for Portugal, Singapore and Taiwan, which, however, switch regimes rather early during the time period considered. Moreover, most lower income countries belong explicitly to the lower human capital group. Last but not least, almost all middle income countries belong to

[23]Illustrative examples of countries switching from the lower to the upper human capital regime and which have undertaken educational reforms are the following, with the switching date indicated in parenthesis: First, Zimbabwe (1987), following the 1980 educational reform recognized education as a basic right and set a goal of achieving universal education for all students by 2000. Second, Taiwan (1975), following the 1968 reform established compulsory primary and lower secondary schooling. Third, Portugal (1983), following a sequence of reforms starting with 1973 José Veiga Simão's reform which strengthened educational institutions and set the foundations for the democratization of the country, and followed by a continuous restructuring process of the educational system to include compulsory, secondary and higher education (see Da Cunha (1993) for a discussion). Finally, Singapore (1982) following the *Survival phase* (1959-1978) of its educational system, aimed to build a skilled workforce to contribute to the country's economic development. This phase was followed by the *Efficiency* (1979-1996), *Ability* (1997-2011) and *Values* (2012-2018) phases strengthening the educational institutions into establishing its position among the leading countries of the high human capital group (see Tan (2008) and Tien (2019) for a discussion).

Table 4: Descriptive statistics for income, saving and human capital

	low h.c. group			high h.c. group			intermediate group		
	inc.	sav.	h.c.	inc.	sav.	h.c.	inc.	sav.	h.c.
mean	2741.66	0.139	1.224	33293.56	0.266	2.789	9333.71	0.205	1.800
st. dev.	582.15	0.064	0.141	12614.67	0.054	0.303	3671.20	0.065	0.354
Correlations (logs)									
inc.	1			1			1		
sav.	0.331	1		−0.131	1		0.315	1	
h.c.	0.084	0.317	1	0.836	−0.188	1	0.581	0.167	1

Notes: The mean and standard deviation are calculated for the level of variable, while the correlation coefficients for the logarithms of the variables, used in the estimation of the model. "h.c." here represents educational human capital.

the intermediate group. More specifically, we find that most lower middle income countries switch regimes relatively late, that is post 1985, while most upper middle income countries make the switch prior to 1985.

4.4 Extension of the model to allow for health human capital effects

In the previous section we considered human capital in the form of educational attainment, as is the case in most studies on the human capital - growth nexus. However, human capital is a multidimensional concept, including the health conditions in a country among others. Its importance for economic development has been highlighted in recent reports by the World Health Organisation and the World Bank (World Health Organization and World Bank (2017) and World Bank (2018)), which state nearly half of the world population does not receive essential health services, while four out of five poor citizens in low income countries lack any social security net.[24]

In this section we modify our model to include health human capital in two ways: (i) as a separate form of capital in the production function, as in Knowles and Owen (1995) and McDonald and Roberts (2002), and (ii) a formation of human capital that

[24]See also Weil (2014) for a comprehensive survey on the growth - health relationship.

combines education and health, as in Weil (2007). To measure the health component of human capital we construct an index in a similar fashion to that included in the Penn World Table for educational attainment. We base this index on Adult Survival Rates (ASR) and the "return to health" computed by Weil (2007), that is

$$x_{it} = \exp(\phi_{ASR} ASR_{it}), \tag{20}$$

where $\phi_{ASR} = 0.653$ denotes return to health. This way we can connect health with productivity, contrary to other measures often used in the literature.[25]

In columns 1 and 2 of Table 5 we present results of the threshold model augmented with the health capital proxy. As threshold variable, we still consider educational human capital. The LR test rejects the null hypothesis of no threshold at the 5% significance level. The threshold value estimate is slightly lower than the 1.719 reported in Table 2, meaning that the country categories in Table 3 remain largely the same, except the transition occurs a couple of years earlier for some countries.

Inclusion of the health conditions of a country does not appear to have any significant impact on the rest of the regressors in the lower regime. Its own effect is negative and only marginally significant (p-value equals 0.095). The countries in the lower regime are also characterised by poor health conditions, which can justify the negative sign and significance of the health coefficient. On the contrary, within the upper regime the health coefficient is positive and highly significant, and even larger than that of educational human capital, which is also positive and highly significant.[26] These results

[25]Life expectancy at birth is often used to proxy for health (see, e.g., Barro and Lee (1994), Barro and Sala-i-Martin (2003) and Desbordes (2011) and Barro (2013a,b) more recently) or some transformation of it, e.g. $-\ln(80 - LE)$ (see, e.g., Knowles and Owen (1995) and McDonald and Roberts (2002)). Such measures, however, are not tied with productivity. Scatter-plots, though, reveal that there is a linear relationship between adult survival rates and life expectancy at birth. Consequently, we can follow a similar procedure to Caselli (2016) and extract the return to life expectancy, yet due to the linearity of the transformation the results are very similar and we do not report them.

[26]Cooray (2013) also finds a positive impact of health on growth for high and upper middle income countries.

highlight that education and health complement well each other in the mechanics of growth. The other Solow determinants are weakened, with population growth being insignificant as well.

In columns 3 and 4 of Table 5, we present results of the threshold model when we combine educational and health human capital.[27] Now the combined variable is used as threshold explanatory variable. Once again the LR test rejects the null hypothesis of no threshold at the 5% significance level. The threshold value estimate is 2.852 lying at the 39.9% of the combined human capital distribution, also implying that the resulting categorisation of countries is largely the same as that presented in Table 3. There are, however, some noteworthy differences. Some countries switch regimes a lot earlier than the year reported in Table 3, e.g. Singapore in 1973 (1982) and Costa Rica which now lies in the higher human capital group (1973). There are also countries which switch regimes a lot later than the year reported in Table 3, mostly sub-Saharan African: Cameroon in 2012 (1999), Rep. of the Congo in 1998 (1989), Kenya in 2004 (1991), Togo in 2014 (2002), Zambia in 2004 (1986), Zimbabwe in 2006 (1986) and South Africa in 1979 and which was found in the higher human capital group. These changes in the groups of countries clearly demonstrate the importance of achieving better health conditions in poorer countries.

Regarding the coefficient of the combined human capital measure, within the lower regime it is negative and highly significant indicating the devastating combined effects that bad health and low schooling levels have on growth. On the contrary, within the upper regime the effect remains strong, positive and highly significant. Finally, for the rest of the slope coefficients in the lower regime, we find negligible differences compared to those reported in Table 2. Within the upper regime, the deviations are a bit larger without, however, affecting either the significance or the signs of the coefficients.

[27] In so doing we multiply the educational human capital Index from the Penn World Table with the health measure we constructed, as implied in Caselli (2005, 2016) and Weil (2014).

Table 5: Estimation results for equation (7) including health capital

	S-S	CST	S-S	CST
		Lower Regime		
$\ln(y_{it-1})$	0.993***	0.987***	0.918***	0.918***
	(0.027)	(0.023)	(0.012)	(0.013)
$\ln(s_{it})$	0.046***	0.047***	0.035***	0.033***
	(0.008)	(0.007)	(0.009)	(0.007)
$\ln(n_{it}+g+\delta)$	0.251***	0.259***	0.283***	0.266***
	(0.028)	(0.027)	(0.027)	(0.021)
$\ln(ehc_{it})$	−0.025	0.118		
	(0.173)	(0.359)		
$\ln(hhc_{it})$	−0.722*	−0.639*		
	(0.432)	(0.380)		
$\ln(chc_{it})$			−0.134**	−0.633***
			(0.062)	(0.207)
		Upper Regime		
$\ln(y_{it-1})$	0.802***	0.810***	0.735***	0.736***
	(0.015)	(0.015)	(0.015)	(0.014)
$\ln(s_{it})$	0.092***	0.087***	0.208***	0.231***
	(0.010)	(0.010)	(0.012)	(0.013)
$\ln(n_{it}+g+\delta)$	−0.022	−0.013	−0.083***	−0.057***
	(0.018)	(0.017)	(0.014)	(0.021)
$\ln(ehc_{it})$	0.513***	0.480***		
	(0.064)	(0.097)		
$\ln(hhc_{it})$	0.649***	0.622***		
	(0.240)	(0.235)		
$\ln(chc_{it})$			0.699***	0.570***
			(0.053)	(0.078)
$q_{it}^{(1)*}$		−0.032		0.079***
		(0.036)		(0.029)
$q_{it}^{(2)*}$		−0.003		0.034*
		(0.018)		(0.018)
$\hat{\gamma}$	1.586	1.586	2.700	2.852
95% C.I.	[1.213,2.723]	[1.536,1.719]	[1.873,3.826]	[2.238,3.858]
J_{stat}	103.70	103.35	104.61	103.12
$LR_{p-value}$	0.754	0.014	0.580	0.012

Notes: Dependent variable is real GDP per working age person (y_{it}). y_{it-1} represents the one period lagged value of output per working age person. In columns 1 and 2 we have separated the effect of educational human capital (*ehc*) and health human capital (*hhc*). As a threshold variable we consider the educational human capital. In columns 3 and 4 we present estimates of model (7) where we consider the combined human capital measure (*chc*) as a regressor and threshold variable. Notation on significance: * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

5 Conclusions

In the present study, we empirically examine the existence of human capital threshold effects on economic growth. In so doing, we use as an econometric framework a variant of the Mankiw, Romer and Weil model that possesses a threshold property which can create regime switches in the output elasticity. As a threshold variable, we initially consider a Human Capital index recently available in the Penn World Table constructed upon years of schooling and returns to education. Such a variable can be used for educational policy recommendations, as those by the World Bank and UNESCO. Then, we extend the human capital variable to allow for health effects, which can be important especially for poorer countries.

To deal with the problem of endogeneity of the threshold variable, we exploit recent developments in the dynamic panel threshold literature. We develop a modified first difference GMM estimator based on the Arellano-Bond principle and we rely on copula theory to account for endogenous threshold variables, without the need of instrumental variables for it. Through a simulation study, we provide evidence that our suggested method performs well in cases of weak and invalid instruments.

The estimation of the model from our data leads to a number of conclusions, with interesting policy implications. We find significant human capital threshold effects on economic growth, which support the club convergence hypothesis within different levels of human capital and suggest large scale policy interventions to close the gap between the poor and rich countries. These include significant investments in education and health services.

Regarding education, we found a critical value of the threshold parameter $\hat{\gamma} = 1.719$, which corresponds to 4 years of schooling and a rate of return on education at 13.4%. When combining education and health human capital the threshold parameter value is $\hat{\gamma} = 2.852$. If one assumes 4 years of schooling and education return 13.4% this value of γ

implies a survival rate 77.5%. Based on the above estimates of the threshold parameter values, we can distinguish three groups of countries. There is a group of countries for which the effects of human capital on economic growth is significant and always strong, and another group for which the effects of human capital mainly come from health. The latter group consists of very poor/low income countries. We also demonstrate that, during our sample, there is large number of countries that transitioned from the low to the higher human capital level regime, and thus achieved significant economic gains. We found that poor health conditions do not help the transition to the higher human capital regime.

A Appendices

A.1 An algorithm for the parametric bootstrap LR test for thresholds

For the model (8), we want to test the null hypothesis of no threshold effects, i.e. $H_0 : \beta_1 = \beta_2$, against the alternative of threshold. The testing procedure is set up as follows:

1. We estimate the models under the null and under alternative hypotheses, and compute the LR statistic defined as

$$LR = -N(J_1 - J_0), \tag{21}$$

   where J_0 is the value of the GMM objective function under the null, and J_1 under the alternative.

2. We save the residuals produced under the null hypothesis, $\widehat{\Delta u_{it}}$. For each bootstrap iteration $b = 1, 2, ..., B$, we draw a random sample with replacement from the distribution of $\widehat{\Delta u_{it}}$ and compute $\widehat{\Delta u}_{it}^{*(b)} = f(\widehat{\Delta u_{it}})v_{it}^{*}$, where v_{it}^{*} is random variable

following the Rademacher distribution, that is

$$v_{it}^{*} = \begin{cases} -1, & \text{with probability} \quad p = 0.5 \\ 1, & \text{with probability} \quad 1 - p = 0.5 \end{cases} \tag{22}$$

and $f(\widehat{\Delta u}_{it}) = \left(\frac{N}{N-K}\right)^{1/2} \widehat{\Delta u}_{it}$ (see Davidson and MacKinnon (2010) for a discussion).

3. Based on sample values of the vector of regressors x_{it} (which includes the lagged dependent variable), the estimates $\hat{\beta}$ under H_0, the bootstrap samples $\widehat{\Delta u}_{it}^{*(b)}$, and fixed initial values (y_{i0}, y_{i1}) for all $i = 1, 2, ..., N$, we calculate the associated bootstrap samples for the dependent variable, denoted by $y_{it}^{*(b)}$.

4. Given $y_{it}^{*(b)}$, we estimate the regression under both the null and alternative hypotheses, and calculate the LR statistic defined above, denoted by $LR^{(b)}$.

5. Finally, we calculate the bootstrap $p-value$ of the $sup-LR$ test as the percentage of $LR^{(b)}$ statistics that exceed the LR computed in step 1.

A.2 Monte carlo study

In this appendix, we present our simulation study and its results in more detail. For the first two simulation scenarios, (i) the case of weak instruments for both the explanatory and the threshold variables and (ii) the case of weak instruments only for the explanatory variables, we consider the following data generating process:

$$y_{it} = (\alpha_1 y_{it-1} + \beta_1 x_{it}) I(q_{it} \leq \gamma) + (\alpha_2 y_{it-1} + \beta_2 x_{it}) I(q_{it} > \gamma) + \eta_i + e_{it} \tag{23}$$

where $\alpha_1 = 0.75$, $\alpha_2 = 0.95$, $\beta_1 = -1$ and $\beta_2 = 1$. x_{it} and q_{it} are defined as follows:

$$q_{it} = cx_{it} + u_{it} \tag{24}$$

$$x_{it} = \phi_0 + \phi_1 x_{it-1} + \xi_{it} \tag{25}$$

where ϕ_0 is fixed at 0.2, while we consider the cases of $\phi_1 = 0.05$ and $\phi_1 = 0.99$, and c is set to 0.276. The above values of ϕ_1 and c are chosen so that they imply a high correlation between the explanatory variable x_{it} and the threshold variable q_{it}, that is $corr(x_{it}, q_{it}) = 0.9$, for all i and t. This reflects the case often encountered in practice, when the threshold variable constitutes a transformation of an explanatory variable. Note that the value of $\phi_1 = 0.99$ corresponds to the case of weak instruments for Δx_{it}, while $\phi_1 = 0.05$ to that of weak instruments for both q_{it} and Δx_{it}.

We set the individual component $\eta_i \sim iidN(0, 1)$. Regarding the disturbance terms e_{it}, ξ_{it} and u_{it}, we consider the following process:

$$\begin{bmatrix} e^*_{it} \\ \xi^*_{it} \\ u^*_{it} \end{bmatrix} \sim GC \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.7 & 0.7 \\ 0.7 & 1 & 0.7 \\ 0.7 & 0.7 & 1 \end{bmatrix} \right)$$

where GC denotes a Gaussian Copula, and then transform e_{it}, ξ_{it} and u_{it}, as

$$e_{it} = \Phi^{-1}(e^*_{it})$$
$$\xi_{it} = -\sqrt{3} + 2\sqrt{3}\xi^*_{it}$$
$$u_{it} = -\sqrt{3} + 2\sqrt{3}u^*_{it}$$

such that $e_{it} \sim N(0, 1)$, while ξ_{it} and u_{it} both follow a continuous uniform distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$, i.e. $U[-\sqrt{3}, \sqrt{3}]$.

For our third simulation scenario, (iii) the case that the instruments for the threshold variable, q_{it}, are invalid, we assume that $c = 0$, $\phi_1 = 0.8$ which excludes the case of weak instruments for x_{it}. For this case the error term e_{it}, the threshold variable q_{it} and the invalid instrument z_{it} are generated by the following process:

$$\begin{bmatrix} e_{it}^* \\ q_{it}^* \\ z_{it}^* \end{bmatrix} \sim GC\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0.7 & 0.7 \\ 0.7 & 1 & 0.7 \\ 0.7 & 0.7 & 1 \end{bmatrix}\right)$$

where GC again denotes a Gaussian Copula, and then transform e_{it}, q_{it} and z_{it}, as

$$\begin{aligned} e_{it} &= \Phi^{-1}(e_{it}^*) \\ q_{it} &= -\sqrt{3} + 2\sqrt{3}q_{it}^* \\ z_{it} &= -\sqrt{3} + 2\sqrt{3}z_{it}^* \end{aligned}$$

such that $e_{it} \sim N(0, 1)$, while q_{it} and z_{it} both follow a continuous uniform distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$, i.e. $U[-\sqrt{3}, \sqrt{3}]$.

For all the above simulation scenarios, we calculate γ to represent the 75th percentile of the distribution of the threshold variable q_{it}. Initial values of x_{i0} and y_{i0} are both fixed at 0, for all $i = 1, 2, ...N$, and, as in Arellano and Bond (1991), we drop the first 10 observations in every iteration in order to limit the initial value effect on our results. We consider samples, with $N = \{50, 100\}$ and $T = \{10, 25, 50\}$. For each simulation exercise, we carry out 1000 iterations, and calculate the bias of each estimator, and its squared error. We report the mean of this bias (denoted as *BIAS*) and the square root of the mean of the squared errors (denoted as *RMSE*), over all iterations.

In Table 6, we present estimates for the first scenario (weak instruments both for Δx_{it} and q_{it}). The results of the table demonstrate that both approaches can efficiently retrieve the true threshold parameter, γ, as well as the autoregressive slope coefficients in either regime. This happens regardless of N and T. This can be attributed to the fact that y_{it-2} constitutes a strong instrument for y_{it-1} and the fact that the estimator $\hat{\gamma}$ is strongly consistent, as noted before.[28] Regarding the estimates of the slope coefficients β_1 and β_2, the results of the table indicate that our method improves upon S-S, for all N and T considered. We find that the S-S estimator produces biases at around 10%, while those by our approach within the range 1-2.5%. Obviously, this can be attributed to the fact that our approach controls for the endogeneity problem of the threshold variable by augmenting regression (23) with the copula transformed terms $q_{it}^{(j)*}$, $j = 1, 2$. In so doing, it mitigates the effect of the endogeneity of x_{it} and/or q_{it} on the FD-GMM estimator in cases where weak instruments problem exists. As a final, note that our approach improves in terms of efficiency as T increases. This can be attributed to the fact that more observations become available, and thus the distribution of the threshold variable q_{it} is more precisely estimated.

In Table 7, we present results for the second simulation scenario (weak instruments only for Δx_{it}). As was expected, the results of the table indicate that our approach performs similarly to the case of simulation scenario (i), for all N and T. It provides efficient estimates for all the parameters, for all N and T. S-S estimator performs better than in case (i) in estimating β_1 and β_2 and this can be attributed to the fact that the moment conditions for q_{it}, are more valid than before and this helps to identify the parameters of x_{it}, given that x_{it} and q_{it} are strongly correlated.

In Table 8, we present results for the simulation scenario (iii). Recall that this scenario represents the case where x_{it} is strictly exogenous, q_{it} is still endogenous and the instruments for q_{it} are invalid (endogenous). The correlation between the instruments and the error term may be contemporaneous or at lags. The results of the table indicate that our approach performs very well for N and T, and improves its efficiency as T increases. Our method improves considerably upon the performance of the S-S method, including the threshold parameter in small T samples. For such samples (e.g. $T =$

[28] See also results for single equation threshold regression models given by Kourtellos et al. (2016) and Samia and Chan (2011)

Table 6: Simulation results for case (i): Weak instruments for both x_{it} and q_{it}

	N=50				N=100			
	Seo & Shin		Copulas		Seo & Shin		Copulas	
$T=10$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	−0.014	0.037	−0.011	0.038	−0.013	0.028	−0.010	0.027
β_1	0.094	0.109	0.026	0.101	0.095	0.108	0.017	0.092
α_2	−0.004	0.041	−0.002	0.045	−0.004	0.032	−0.001	0.033
β_2	0.102	0.133	0.012	0.173	0.102	0.132	0.015	0.164
γ	0.002	0.126	0.007	0.168	0.001	0.084	0.009	0.113
$T=25$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	−0.007	0.013	−0.005	0.011	−0.006	0.010	−0.004	0.009
β_1	0.094	0.098	0.022	0.051	0.091	0.095	0.016	0.051
α_2	0.001	0.014	0.002	0.014	0.003	0.012	0.004	0.013
β_2	0.109	0.119	0.022	0.088	0.108	0.118	0.016	0.091
γ	−0.001	0.055	0.000	0.059	−0.000	0.041	−0.000	0.045
$T=50$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	−0.014	0.022	−0.006	0.013	−0.005	0.007	−0.004	0.006
β_1	0.097	0.101	0.024	0.045	0.091	0.093	0.015	0.034
α_2	−0.007	0.022	−0.001	0.014	0.003	0.008	0.003	0.008
β_2	0.104	0.115	0.022	0.074	0.106	0.111	0.014	0.059
γ	0.002	0.041	0.000	0.042	0.001	0.032	0.002	0.034

Notes: The value 0.000 signifies values that are less than 5×10^{-4}.

10), the S-S approach produces higher values of bias and RMSE of γ, even for the case that N is large. Such biases can explain those of all the remaining coefficients of the model (namely $(\alpha_1, \alpha_2, \beta_1, \beta_2)$), despite the fact that the instruments for Δx_{it} and Δy_{it-1} are valid and strong. Note that, as T increases relative to N, the bias of the threshold parameter γ based on the S-S method tends to reduce considerably, but still our method performs more satisfactorily. This result shows that the S-S estimator of γ may be strongly consistent as T grows large. However, this does not apply to the slope coefficients of the explanatory variable x_{it} (see β_1 and β_2) which still remain biased, suffering from the problem of the invalid instruments.

Table 7: Simulation results with $\phi_1 = 0.99$.

	N=50				N=100			
	Seo & Shin		Copulas		Seo & Shin		Copulas	
$T = 10$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	-0.006	0.012	-0.003	0.012	-0.005	0.011	-0.002	0.011
β_1	0.061	0.080	0.011	0.063	0.060	0.080	0.010	0.062
α_2	-0.004	0.011	-0.003	0.012	-0.003	0.008	-0.002	0.009
β_2	0.078	0.094	0.035	0.075	0.076	0.090	0.029	0.071
γ	0.000	0.095	0.001	0.118	0.005	0.061	0.005	0.076
$T = 25$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	-0.002	0.005	0.000	0.005	-0.001	0.005	-0.000	0.004
β_1	0.030	0.040	-0.016	0.031	0.025	0.036	-0.017	0.029
α_2	-0.001	0.004	-0.001	0.004	-0.000	0.003	-0.000	0.003
β_2	0.051	0.058	0.011	0.037	0.045	0.052	0.004	0.031
γ	-0.002	0.047	0.000	0.053	0.002	0.038	0.002	0.041
$T = 50$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	-0.001	0.003	-0.001	0.003	-0.000	0.004	0.000	0.004
β_1	0.023	0.032	-0.018	0.028	0.013	0.022	-0.024	0.031
α_2	-0.000	0.002	-0.000	0.002	0.000	0.002	-0.000	0.002
β_2	0.039	0.045	-0.003	0.027	0.031	0.039	-0.009	0.037
γ	-0.002	0.039	-0.003	0.040	-0.002	0.033	-0.001	0.033

Notes: The value 0.000 signifies values that are less than 5×10^{-4}.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). Institutions as a Fundamental Cause of Long-Run Growth. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, pages 385–472, London. Elsevier.

Aghion, P., Alesina, A., and Trebbi, F. (2004). Endogenous Political Institutions. *The Quarterly Journal of Economics*, 119(2):565–611.

Ahuja, A., Baird, S., Hicks, J. H., Kremer, M., Miguel, E., and Powers, S. (2015). When Should Governments Subsidize Health? The Case of Mass Deworming. *The World Bank Economic Review*, 29(suupl 1):S9–S24.

Table 8: Simulation results with an invalid instrument z_{it} for q_{it}.

	N=50				N=100			
	Seo & Shin		Copulas		Seo & Shin		Copulas	
$T = 10$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	0.034	0.129	−0.007	0.056	0.051	0.142	0.000	0.044
β_1	−0.351	0.503	−0.010	0.149	−0.479	0.603	−0.017	0.120
α_2	−0.094	0.294	−0.001	0.145	−0.146	0.355	−0.003	0.137
β_2	0.794	0.984	0.022	0.306	1.098	1.286	0.032	0.280
γ	−0.707	1.934	−0.039	0.703	−0.917	2.180	−0.011	0.643
$T = 25$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	0.036	0.063	−0.003	0.024	0.048	0.070	−0.003	0.023
β_1	−0.261	0.292	−0.010	0.067	−0.375	0.391	−0.004	0.056
α_2	−0.075	0.150	−0.004	0.057	−0.127	0.196	0.004	0.060
β_2	0.785	0.816	0.017	0.121	1.061	1.089	0.020	0.121
γ	−0.127	0.396	0.008	0.123	−0.237	0.612	0.010	0.103
$T = 50$	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
α_1	0.058	0.082	−0.005	0.023	0.048	0.060	−0.001	0.013
β_1	−0.292	0.341	−0.016	0.097	−0.355	0.365	−0.006	0.033
α_2	−0.022	0.131	−0.005	0.053	−0.113	0.150	−0.002	0.033
β_2	0.897	0.928	0.001	0.140	1.095	1.107	0.015	0.072
γ	−0.058	0.181	0.009	0.066	−0.114	0.281	0.004	0.046

Notes: The value 0.000 signifies values that are less than 5×10^{-4}.

Alonso, S., Chaccour, C. J., Elobolobo, E., Nacima, A., Candrinho, B., Saifodine, A., Saute, F., Robertson, M., and Zulliger, R. (2019). The economic burden of malaria on households and the health system in a high transmission district of Mozambique. *Malaria Journal*, 18(360).

Amengual, D. and Sentana, E. (2020). Is a Normal Copula the Right Copula? *Journal of Business and Economic Statistics*, 38(2):350–366.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277–297.

Azariadis, C. and Drazen, A. (1990). Threshold Externalities in Economic Development.

The Quarterly Journal of Economics, 105(2):501–526.

Barro, R. J. (1991). Economic Growth in a Cross Section of countries. *The Quarterly Journal of Economics*, 106(2):407–443.

Barro, R. J. (2013a). Education and Economic Growth. *Annals of Economics and Finance*, 14(2):277–304.

Barro, R. J. (2013b). Health and Economic Growth. *Annals of Economics and Finance*, 14(2):305–342.

Barro, R. J. and Lee, J.-W. (1994). Sources of economic growth. *Carnegie-Rochester Conference Series on Public Policy*, 40:1–46.

Barro, R. J. and Sala-i-Martin, X. (2003). *Economic Growth, 2nd Ed.* MIT Press, Cambridge, MA.

Belot, M. and James, J. (2011). Healthy school meals and educational outcomes. *Journal of Health Economics*, 30(3):489–504.

Benhabib, J. and 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.

Bernanke, B. S. and Gürkaynak, R. S. (2001). Is Growth Exogenous? Taking Mankiw, Romer and Weil Seriously. *NBER Macroeconomics Annual*, 16(1):11–57.

Bucci, A. (2015). Product Proliferation, Population, and Economic Growth. *Journal of Human Capital*, 9(2):170–197.

Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giornale dell'Istituto Italiano degli Attuari*, 4:421–424.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, pages 679–741. Elsevier.

Caselli, F. (2016). *Accounting for Cross-Country Income Differences: Ten Years Later.* World Bank, Washington DC.

Caselli, F., Esquivel, G., and Lefort, F. (1996). Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics. *Journal of Economic Growth*, 1(3):363–389.

Cerra, V. and Saxena, S. C. (2008). Growth Dynamics: The Myth of Economic Recovery. *American Economic Review*, 98(1):439–457.

Christopoulos, D. and León-Ledesma, M. A. (2014). Efficiency and production frontiers in the aftermath of recessions: International evidence. *Macroeconomic Dynamics*, 18(6):1326–1350.

Christopoulos, D., McAdam, P., and Tzavalis, E. (2021). Dealing With Endogeneity in Threshold Models Using Copulas. *Journal of Business and Economic Statistics*, 39(1):166–178.

Chuma, J. M., Thiede, M., and Molyneux, C. S. (2006). Rethinking the economic costs of malaria at the household level: Evidence from applying a new analytical framework in rural Kenya. *Malaria Journal*, 5(76).

Cooray, A. (2013). Does health capital have differential effects on economic growth? *Applied Economics Letters*, 20(3):244–249.

Cravo, T. A. and Soukiazis, E. (2011). Human Capital Thresholds and Economic Growth in Brazil. *RDE Revista De Desenvolvimento Econômico*, 13(23):106–120.

Da Cunha, S. V. L. (1993). Educational Reforms in Portugal: Helping to Raise the Profile of VET. *European Journal of Education*, 28(2):215–227.

Davidson, R. and MacKinnon, J. G. (2010). Wild Bootstrap Tests for IV Regression. *Journal of Business and Economic Statistics*, 28(1):128–144.

Desbordes, R. (2011). The non-linear effects of life expectancy on economic growth. *Economics Letters*, 112(1):116–118.

Dillon, A., Friedman, J., and Serneels, P. (2021). Health Information, Treatment, and Worker Productivity. *Journal of the European Economic Association*, 19(2):1077–1115.

Durlauf, S. N. and Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, 10(4):365–384.

Durlauf, S. N., Johnson, P. A., and Temple, J. R. (2005). Growth Econometrics. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, pages 555–677, London. Elsevier.

Durlauf, S. N., Johnson, P. A., and Temple, J. R. (2009). The Methods of Growth Econometrics. In Mills, T. C. and Patterson, K., editors, *Palgrave Handbook of Econometrics*, volume 2, pages 1119–1179, London. Palgrave Macmillan.

Erdely, A. (2017). Copula-based piecewise regression. In Úbeda Flores, M., de Amo Artero, E., Durante, F., and Sánchez, J. F., editors, *Copulas and Dependence Models with Applications*, pages 69–81, Cham. Springer International Publishing.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, 105(10):3150–3182. available at http://www.ggdc.net/pwt.

Gallup, J. L. and Sachs, J. D. (2001). The economic burden of malaria. *The American Journal of Tropical Medicine and Hygiene*, 64(1 suppl):85–96.

Glivenko, V. I. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giornale dell'Istituto Italiano degli Attuari*, 4:92–99.

Hansen, B. E. (2000). Sample Splitting and Threshold Estimation. *Econometrica*, 68(3):575–603.

Harris, R. D. and Tzavalis, E. (1999). Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics*, 91(2):201–226.

Hsiao, C. and Zhang, J. (2015). IV, GMM or likelihood approach to estimate dynamic panel models when either N or T or both are large. *Journal of Econometrics*, 187(1):312–322.

Islam, N. (1995). Growth Empirics: A Panel Data Approach. *The Quarterly Journal of Economics*, 110(4):1127–1170.

Joe, H. (2014). *Dependence Modelling with Copulas.* Chapman & Hall, London.

Johnson, P. and Papageorgiou, C. (2020). What Remains of Cross-Country Convergence? *Journal of Economic Literature*, 58(1):129–175.

Jones, C. I. (1999). Growth: With or Without Scale Effects? *American Economic Review*, 89(2):139–144.

Kalaitzidakis, P., Mamuneas, T. P., Savvides, A., and Stengos, T. (2001). Measures of Human Capital and Nonlinearities in Economic Growth. *Journal of Economic Growth*, 6(3):229–254.

Ketteni, E., Mamuneas, T. P., and Stengos, T. (2007). Nonlinearities in economic growth: A semiparametric approach applied to information technology data. *Journal of Macroeconomics*, 29(3):555–568.

Ketteni, E., Mamuneas, T. P., and Stengos, T. (2011). The Effect of Information Technology and Human Capital on Economic Growth. *Macroeconomic Dynamics*, 15(5):595–615.

Knight, M., Loayza, N., and Villanueva, D. (1993). Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach. *IMF Staff Papers*, 40(3):512–541.

Knowles, S. and Owen, D. P. (1995). Health capital and cross-country variation in income per capita in the Mankiw-Romer-Weil model. *Economics Letters*, 48(1):99–106.

Kottaridi, C. and Stengos, T. (2010). Foreign direct investment, human capital and non-linearities in economic growth. *Journal of Macroeconomics*, 32(3):858–871.

Kourtellos, A., Stengos, T., and Sun, Y. (2021). Endogeneity in Semiparametric Threshold Regression. *Econometric Theory forthcoming.*

Kourtellos, A., Stengos, T., and Tan, C. M. (2016). Structural Threshold Regression. *Econometric Theory*, 32(4):827–860.

Lau, C. K. M. (2010). New evidence about regional income divergence in China. *China Economic Review*, 21(2):293–309.

Lau, L. J., Jamison, D. T., Liu, S.-C., and Rivkin, S. (1993). Education and economic growth Some cross-sectional evidence from Brazil. *Journal of Development Economics*, 41(1):45–70.

Lee, L.-F. (1983). Generalized Econometric Models with Selectivity. *Econometrica*, 51(2):507–512.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108(1):1–24.

Liu, Z. and Stengos, T. (1999). Non-linearities in Cross-Country Growth Regressions: A Semiparametric Approach. *Journal of Applied Econometrics*, 14(5):527–538.

Lucas, A. M. (2010). Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka. *American Economic Journal: Applied Economics*, 2(2):46–71.

Lutz, W., Goujon, A., KC, S., Stonawski, M., and Stilianakis, N. (2018). *Demographic and Human Capital Scenarios for the 21st Century: 2018 assessment for 201 countries*. Publications Office of the European Union.

Mankiw, G. N., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2):407–437.

McDonald, S. and Roberts, J. (2002). Growth and multiple forms of human capital in an augmented solow model: a panel data investigation. *Economics Letters*, 74(2):271–276.

Mierau, J. O. and Turnovsky, S. J. (2014). Demography, growth, and inequality. *Economic Theory*, 55(1):29–68.

Pargianas, C. (2017). Endogenous Political Institutions, Wage Inequality, and Economic Growth. *Macroeconomic Dynamics*, 21(1):183–213.

Park, S. and Gupta, S. (2012). Handling Endogenous Regressors by Joint Estimation Using Copulas. *Marketing Science*, 31(4):567–586.

Pelinescu, E. (2015). The Impact of Human Capital on Economic Growth. *Procedia Economics and Finance*, 22:184–190.

Pelloni, A., Stengos, T., and Valenti, F. (2019). The Non-linearity in the Relationship Between Human Capital and Growth. In Bucci, A., Prettner, K., and Prskawetz, A., editors, *Human Capital and Economic Growth: The Impact of Health, Education and Demographic Change*, pages 3–26. Springer International Publishing.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9):1325–1343.

Ram, R. (2007). IQ and economic growth: Further augmentation of Mankiw-Romer-Weil model. *Economics Letters*, 94(1):7–11.

Rice, J. A. (2007). *Mathematical Statistics and Data Analysis, 3rd Ed.* Cengage Learning.

Roodman, D. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71(1):135–158.

Sala-i-Martin, X. (2002). 15 Years of New Growth Economics : What Have we Learnt? *Journal Economía Chilena (The Chilean Economy)*, 5(2):5–15.

Samia, N. I. and Chan, K.-S. (2011). Maximum likelihood estimation of a generalized threshold stochastic regression model. *Biometrika*, 98(2):433–448.

Sandjaja, S., Poh, B. K., Rojroonwasinkul, N., Nyugen, B. K. L., Budiman, B., Ng, L. O., Soonthorndhada, K., Xuyen, H. T., Deurenberg, P., and Parikh, P. (2013). Relationship between anthropometric indicators and cognitive performance in southeast asian school-aged children. *British Journal of Nutrition*, 110(S3):S57–S64.

Sarma, N., Patouillard, E., Cibulskis, R. E., and Arcand, J.-L. (2019). The Economic Burden of Malaria: Revisiting the Evidence. *The American Journal of Tropical Medicine and Hygiene*, 101(6):1405–1415.

Savvides, A. and Stengos, T. (2008). *Human Capital and Economic Growth.* Stanford University Press.

Seo, M. H. and Shin, Y. (2016). Dynamic panels with threshold effect and endogeneity. *Journal of Econometrics*, 195(2):169–186.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.

Simon, J. L. (1976). Population Growth May Be Good for LDCs in the Long Run: A Richer Simulation Model. *Economic Development and Cultural Change*, 24(2):309–337.

Simon, J. L. (1990). *Population Matters: People, Resources, Environment and Immigration*. Transaction Publishers, New Brunswick, N.J.

Sirimaneetham, V. and Temple, J. R. W. (2009). Macroeconomic stability and the distribution of growth rates. *The World Bank Economic Review*, 23(3):443–479.

Soukiazis, E. and Cravo, T. A. (2008). Human Capital and the Convergence Process Among Countries. *Review of Development Economics*, 12(1):124–142.

Tan, C. (2008). Globalisation, the Singapore state and educational reforms: towards performativity. *Education, Knowledge & Economy*, 2(2):111–120.

Temple, J. (1999). A positive effect of human capital on growth. *Economics Letters*, 65(1):131–134.

Teulings, C. and van Rens, T. (2008). Education, Growth, and Income Inequality. *The Review of Economics and Statistics*, 90(1):89–104.

Tien, M. (2019). Singapore's Education Reforms: Learning Is not a Competition. *CommonWealth Magazine*, 685.

Vella, F. (1998). Estimating models with sample selection bias: A survey. *The Journal of Human Resources*, 33(1):127–169.

Weil, D. N. (2007). Accounting for the Effect Of Health on Economic Growth. *The Quarterly Journal of Economics*, 122(3):1265–1306.

Weil, D. N. (2014). Health and Economic Growth. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 2, pages 623–682, London. Elsevier.

WHO Regional Office for Europe (2020). *Health 2020: Education and health through early development*. World Health Organization.

World Bank (2018). *The State of Social Safety Nets 2018*. World Bank, Washington DC.

World Health Organization and World Bank (2017). *Tracking Universal Health Coverage: 2017 Global Monitoring Report*. World Bank, Washington DC.

Zhang, Y. and Zhou, Q. (2020). Correction for the Asymptotical Bias of the Arellano-Bond type GMM Estimation of Dynamic Panel Models. In Li, T., Pesaran, M. H., and Terrell, D., editors, *Advances in Econometrics: Essays in Honor of Cheng Hsiao*, volume 41, pages 1–24. Emerald Publishing Limited.

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2020

01-20 Incubated Early Stage Startuppers in Athens and their Initiatives during the Crisis (2010-2016), Ioannis Besis and Ioanna Sapfo Pepelasis
02-20 Alternative Monetary Policy Rules in an Imperfectly Competitive Dynamic Stochastic General Equilibrium Model, George Alogoskoufis and Stelios Giannoulakis
03-20 Quantifying Qualitative Survey Data: New Insights on the (Ir)Rationality of Firms' Forecasts, Alexandros Botsis, Christoph Görtz and Plutarchos Sakellaris
04-20 Can Country-Specific Interest Rate Factors Explain the Forward Premium Anomaly?, Efthymios Argyropoulos, Nikolaos Elias, Dimitris Smyrnakis and Elias Tzavalis
05-20 Signaling product quality through prices in an oligopoly model with costly consumer search, Minghua Cheny, Konstantinos Serfes and Eleftherios Zacharias
06-20 Thinking ahead of the next big crash: Clues from Athens in classical times, George C. Bitros
07-20 Financial crises, firm-level shocks, and large downturns: Evidence from Greece, Stelios Giannoulakis and Plutarchos Sakellaris. Published in: *International Journal of Finance and Economics*, 2021
08-20 Notes on the Demand Side Secular Stagnation, George D. Demopoulos and Nicholas A. Yannacopoulos
09-20 Institutions and Macroeconomic Performance: Core vs Periphery Countries in the Eurozone, Tryfonas Christou, Apostolis Philippopoulos and Vanghelis Vassilatos
10-20 Macroeconomic Policy Lessons for Greece, George Economides, Dimitris Papageorgiou and Apostolis Philippopoulos
11-20 Energy Transitions and the role of the EU ETS: The case of Greece, Andriana Vlachou and Georgios Pantelias
12-20 Measuring the Default Risk of Small Business Loans: Improved Credit Risk Prediction using Deep Learning, Yiannis Dendramis, Elias Tzavalis and Aikaterini Cheimarioti
14-20 Saving Democracy by Switching to Corporate-like Governance, George C. Bitros
15-20 The impact of the lockdown on the Greek economy and the role of the Recovery Fund, George Economides, Apostolis Philippopoulos and Vanghelis Vassilatos

2021

01-21 Historical Cycles of the Economy of Modern Greece From 1821 to the Present, George Alogoskoufis
02-21 Greece Before and After the Euro: Macroeconomics, Politics and the Quest for Reforms, George Alogoskoufis
03-21 Commodity money and the price level, George C. Bitros. Published in: *Quarterly Journal of Austrian Economics*, 2022
04-21 Destabilizing asymmetries in central banking: With some enlightenment from money in classical Athens, George C. Bitros. Published in: *Journal of Economic Asymmetries*, 2021

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

05-21 Exploring the Long-Term Impact of Maximum Markup Deregulation, Athanasios Dimas and Christos Genakos
06-21 A regularization approach for estimation and variable selection in high dimensional regression models, Y. Dendramis, L. Giraitis, G. Kapetanios
07-21 Tax Competition in the Presence of Environmental Spillovers, Fabio Antoniou, Panos Hatzipanayotou, Michael S. Michael, Nikos Tsakiris
08-21 Firm Dynamics by Age and Size Classes and the Choice of Size Measure, Stelios Giannoulakis and Plutarchos Sakellaris
09-21 Measuring the Systemic Importance of Banks, Georgios Moratis, Plutarchos Sakellaris
10-21 Firms' Financing Dynamics Around Lumpy Capacity Adjustments, Christoph Görtz, Plutarchos Sakellaris, John D. Tsoukalas
11-21 On the provision of excludable public goods General taxes or user prices? George Economides and Apostolis Philippopoulos
12-21 Asymmetries of Financial Openness in an Optimal Growth Model, George Alogoskoufis
13-21 Evaluating the impact of labour market reforms in Greece during 2010-2018 , Georgios Gatopoulos, Alexandros Louka, Ioannis Polycarpou, Nikolaos Vettas
14-21 From the Athenian silver to the bitcoin standard: Private money in a state-enforced free banking model, George C. Bitros
15-21 Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities. Stelios Arvanitis and Thierry Post
16-21 Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka
17-21 Competition and Pass-Through: Evidence from Isolated Markets, Christos Genakos and Mario Pagliero
18-21 Exploring Okun's Law Asymmetry: An Endogenous Threshold LSTR Approach, Dimitris Christopoulos, Peter McAdam and Elias Tzavalis
19-21 Limit Theory for Martingale Transforms with Heavy-Tailed Multiplicative Noise, Stelios Arvanitis and Alexandros Louka
20-21 Optimal taxation with positional considerations, Ourania Karakosta and Eleftherios Zacharias
21-21 The ECB's policy, the Recovery Fund and the importance of trust: The case of Greece, Vasiliki Dimakopoulou, George Economides and Apostolis Philippopoulos

2022

01-22 Is Ireland the most intangible intensive economy in Europe? A growth accounting perspective, Ilias Kostarakos, KieranMcQuinn and Petros Varthalitis
02-22 Common bank supervision and profitability convergence in the EU, Ioanna Avgeri, Yiannis Dendramis and Helen Louri
03-22 Missing Values in Panel Data Unit Root Tests, Yiannis Karavias, Elias Tzavalis and Haotian Zhang
04-22 Ordering Arbitrage Portfolios and Finding Arbitrage Opportunities, Stelios Arvanitis and Thierry Post
05-22 Concentration Inequalities for Kernel Density Estimators under Uniform Mixing, Stelios Arvanitis
06-22 Public Sector Corruption and the Valuation of Systemically Important Banks, Georgios Bertsatos, Spyros Pagratis, Plutarchos Sakellaris
07-22 Finance or Demand: What drives the Responses of Young and Small Firms to Financial Crises? Stelios Giannoulakis and Plutarchos Sakellaris
08-22 Production function estimation controlling for endogenous productivity disruptions, Plutarchos Sakellaris and Dimitris Zaverdas
09-22 A panel bounds testing procedure, Georgios Bertsatos, Plutarchos Sakellaris, Mike G. Tsionas

10-22 Social policy gone bad educationally: Unintended peer effects from transferred students, Christos Genakos and Eleni Kyrkopoulou
11-22 Inconsistency for the Gaussian QMLE in GARCH-type models with infinite variance, Stelios Arvanitis and Alexandros Louka
12-22 Time to question the wisdom of active monetary policies, George C. Bitros
13-22 Investors' Behavior in Cryptocurrency Market, Stelios Arvanitis, Nikolas Topaloglou and Georgios Tsomidis
14-22 On the asking price for selling Chelsea FC, Georgios Bertsatos and Gerassimos Sapountzoglou
15-22 Hysteresis, Financial Frictions and Monetary Policy, Konstantinos Giakas
16-22 Delay in Childbearing and the Evolution of Fertility Rates, Evangelos Dioikitopoulos and Dimitrios Varvarigos

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/