

LOOKING FOR THE RIGHT HUMAN CAPITAL PROXY

Rudolf Kubík

Katedra veřejných financí, VŠE v Praze, nám. W. Churchilla 4, Praha 3,
rudolf.kubik@vse.cz

Abstract: This paper aims to test different approaches of human capital stock approximation. It faces one of the main questions in explaining link between human capital and economic growth. It tries to step forward in answering what is the best proxy of human capital. It starts from Barro & Lee and Cohen & Soto datasets which are expanded by Mincerian approach to human capital measurement and educational structure of population as a human capital proxy. The original dataset covering 73 countries within 1960-1990 is being re-tested and results from panel data regression analyses are compared with expanded dataset.

Key words: Human capital, growth, panel data regression

JEL Classification: E24, J24, O47

DOI: 10.2478/v10135-009-0009-0

LOOKING FOR THE RIGHT HUMAN CAPITAL PROXY¹

Rudolf Kubík²

Introduction

Human capital as a significant determinant of economic prosperity has been a central point of attention for a long time. On the micro level, Mincer (1974) showed that one additional year of education raises earnings by 10%. On macro level the main stream represented by Barro proves human capital plays the key role in determining growth. On the other hand, there are opposite views, for example Bills and Klenow (2000) or Prichett (2001), who are pointing on overestimation of the human capital role. However, the relationship is intuitive, still there is no clear and indisputable evidence of human capital boosting an economic growth. Average years of schooling have been most widely accepted as a good proxy of human capital. Although there are studies (Cohen and Soto, 2007) arguing that the main reason why we fail in giving clear evidence is the measurement of human capital. First, how can we approximate human capital stock? Second, what is the data quality?

De la Fuente and Domenech (2006) demonstrate on a group of 21 OECD countries that the data are very unreliable and therefore the growth regressions are not relevant. Prichett (2001) demonstrates that the reason why there is not clear link between growth and education is quality and excess supply of schooling and poor institutional framework in developing countries. This invokes insignificant or low social returns on schooling. Hanushek and Kimko (2000) argue that the results are distorted also by another factor: schooling quality. Schooling quality differs considerably among countries is not usually considered in the empirical models.

Contrary to most of previous papers, this study tries not to follow at major point the main question whether there is or is not the relationship between human capital and economic growth but it aims to test what proxy of human capital fits best in the growth models. In most cases, the average years of schooling have been used for human capital interpretation (Barro & Lee 2000, Cohen & Soto 2007, Arnold & Bassanini & Scarpetta 2007, etc.). But one can hardly believe that doubling the average years of schooling (for e.g. Algeria increased years of schooling in 1980 from 1,6 years to 3,1 years in 1990) can be interpreted as doubling human capital stock in given country, therefore invoking potential double growth of economy output. In literature, there are also other possible options how to proxy human capital stock, e.g. school enrolment or pure educational structure of population which is newly employed in empirical models presented in this paper. The best proxy of the human capital will be assessed basically by statistical criterions and overall model "relevance" evaluation.

¹ Článek vznikl při řešení projektu IGA VŠE F1/30/2010 „Vliv daňových a výdajových nástrojů na mikroekonomickou a makroekonomickou efektivnost“.

² Ing. Rudolf Kubík, Katedra veřejných financí, VŠE v Praze, náměstí Winstona Churchilla 4, Praha 3, 130 67, rudolf.kubik@vse.cz

The paper is organized as follows: In the first section the theoretical background and literature overview is introduced, section II describes the methodology and data, section III introduces the empirical model and regression tests results, and finally, section IV concludes the article.

Theoretical background

Solow-Swan model is usually the starting point for growth accounting framework. It considers three main determinants of economic growth: labor L , capital K and residuum A . This residuum is usually translated as technology, knowledge or total factor of productivity TFP .

$$Y(t) = F[A(t), K(t), L(t)] \quad (1)$$

First, the residua A started to be interpreted as technology development which was either exogenous (Solow) or endogenous (Romer, Lucas). Consequently, human capital has been incorporated in the growth models and endogenous growth theory suggests that human capital is the driver for economic growth in the long run (Barro and Sala-i-Martin, 2004) as it is believed that human capital and innovations can generate non-diminishing returns to scale from capital.

While L and K are easy for interpretation, TFP could be in general interpreted as residual, which is explaining the rest of the economic growth not possible to explain by capital or labor. Human capital could be considered as a qualitative dimension of labor L , technology could be perceived as qualitative dimension of physical capital K .

One of the first authors aiming to discover the qualitative aspect of labor L was Denison (1962), Jorgenson and Griliches (1968) and Ho and Jorgenson (1999) who worked out the structure of factors influencing the labor quality. These studies were performed on US data from labor market. They consider mainly sex, age, level of education and self-employment status.

Based on growth accounting framework presented for example by Barro, Sala-i-Martin (2004) the GDP growth can be desegregated into three main parts:

1. population growth,
2. labor input growth (total number of hours worked),
3. labor productivity growth (e.g. measured as GDP per hour worked).

Labor productivity can be further decomposed:

- capital productivity (so called capital deepening) measured as a ratio of amount of physical capital per hour worked,
- labor force quality (LFQ),
- total factor of productivity (TFP).

Ho and Jorgenson (1999) conclude that TFP can explain in average 30% of growth in developed countries within years 1947–1995. They also conclude that within 1947–1973 there is significantly lower contribution of TFP to growth as this period is known as so called productivity slowdown period.

Tang and MacLeod (2006), for instance, concluded that ageing decreases the labor productivity. Older workers are on average less productive than younger workers and labor force ageing has a negative impact on productivity growth in whole Canada.

Aaronson and Sullivan (2001) testify that there is a robust link between labor quality and economic cycle. They argue that the labor quality decreases when there is an economic expansion. In years of prosperity also less-qualified workers are entering the labor market and are able to assert themselves. Overall labor force quality therefore is worsening. Such results are confirmed also by European Central Bank study (ECB, 2006) which presents interesting results:

- Productivity growth in European countries within 1983-2004 can be explained from one third by human capital quality growth (approximated by educational level and labor market experience). Labor force quality has been annually increasing in average by 0.6%.
- Significant increase in quality began in the 1990s when higher amount of people with tertiary education started to enter the labor market. Consequently, in the second part of the 1990s, there was a slowdown of this accrual due to economic growth and overall increase of employment which brought also workers with lower human capital.
- Considering the labor force quality obviously decreases the importance of *TFP* for economic growth.
- Slightly negative correlation between labor force quality and GDP confirms anticyclical labor force quality behavior.

Methodology and data

Most widely acknowledged interpretation of human capital is the average years of schooling of economically active population (first introduced by Benhabib and Spiegel 1994). This interpretation is broadly employed in empirical papers. Despite some methodology updates done mainly by Cohen and Soto (2007), these measures are still rather quantitative approximation of human capital. But there are quite few more possible human capital proxies possible to employ in the models. Among others:

- Enrollment rates,
- Discounted value of the wage premium due to education,
- International test scores by students,
- International adult literacy tests,
- Estimates of the market value of human capital,
- Estimates of educational attainment.

Due to lack of data and the extent of this study, I expand the Barro-Lee and Cohen-Soto datasets by only two further human capital approximations: pure educational structure of population and Mincerian approximation.

According to Mincerian approach also applied by Prichett (2001), human capital is defined as exponential function,

$$h = \exp(r \cdot S) - 1 \quad (2)$$

where h is human capital stock per worker, r is return to education (set at 0,1 by Prichett) and S is the average years of schooling. This Mincerian formula will be applied in the regression models to proxy the human capital (section III.). Variable S is thus approximated either by years of schooling (Barro & Lee or Cohen & Soto data) or by percentage share of population with given level of education (OECD data).

Both sources Cohen and Soto [CS] and Barro and Lee [BL] proxy the human capital stock as a population share on educational attainment multiplied by the appropriate length in years of each educational category. Cohen and Soto (2007) update this approach and apply methodology, where the years of schooling are interpreted as weighted average of different age groups:

$$yS_t = \sum_{g=1}^G l_t^g yS_t^g \quad (3)$$

where l_t^g is the population share of group g in population of 15 years of age and above and yS_t^g is the number of years of schooling of group g and is defined as,

$$yS_t^g = \sum_j a_t^g D_j \quad (4)$$

where a_t^g interpretes the share of group g with educational level j and D_j is the corresponding duration of education in years.

I apply basically four approaches to approximate human capital in regression models. First, I am following average years of schooling by [BL]. Second, I use updated methodology of average years of schooling introduced by [CS]. Third, pure educational structure of population data introduced by OECD is employed [OECD] and the fourth approach applies Mincerian methodology on [BL], [CS] and [OECD] data.

The data set covers 73 countries within 1960-1990 in ten-year intervals. These countries can be divided in six major groups: Africa (17), Asia (14), Europe (16), North America (2), South and Latin America (22), Australia & Oceania (2). On average, African countries reach the highest percentage share of people with no schooling and lowest average years of schooling. On the contrary, in terms of average years of schooling, the most educated continent is North America. Following table 1 there is not very significant surge in average years of schooling in the developed countries. Most rapid growth is in African and Asian countries. We can also observe considerable gap between Europe and North America/Australia.

Table 1: Average years of schooling by continents (Barro-Lee data)

	1960	1970	1980	1990
Africa	1,46	1,70	2,42	3,14
Asia	2,09	2,63	3,53	4,87
Australia & Oceania	9,50	9,73	10,73	10,65
Europe	5,86	6,49	7,43	8,29
North America	8,52	9,30	11,07	11,25
South & Latin America	3,25	3,59	4,49	5,36

Source: own calculations

The presented dataset was combined basically from three sources:

- Penn World Tables 6.3 for GDP per worker¹ data,
- Cohen and Soto 2007² panel data set contains following variables:
 - Years of schooling of population aged 25 and over calculated according to methodology introduced by:
 - ♦ Cohen & Soto [CS],
 - ♦ Barro & Lee [BL].
 - Capital per worker,
- OECD educational database panel dataset contains following variables³:
 - Years of schooling of population aged 25 and over,
 - Percentage share of population aged 25 and over:
 - ♦ Without schooling,
 - ♦ With completed higher education.

Model specification

According to Cohen and Soto (2007) I estimated a simple production function based on augmented Solow model,

$$Y_{it} = A_{it} K_{it}^{\alpha} H_{it}^{1-\alpha} \quad (5),$$

where Y is the aggregate output of country i in year t , A is total factor productivity, K is aggregate physical capital and H is aggregate human capital stock calculated as

$$H_{it} = h_{it} L_{it} \quad (6),$$

where h is human capital per worker and L is total labor force. After dividing (4) by L and logarithming

$$\log(y_{it}) = \log(A_{it}) + \alpha \log(k_{it}) + (1 - \alpha) \log(h_{it}) \quad (7),$$

¹ „Worker“ is a census definition based on economically active population

² Data available at <http://soto.iaecsic.org/Data.htm>

³ The OECD dataset contains not only data for OECD members! It covers 95 countries around the world and is available at <http://www.oecd.org/dataoecd/33/13/2669521.xls>

Based on the data, I am able to build an unbalanced panel for 73 countries within 1960–1990 (as the data goes in ten-year intervals the time series length is 4). Since the dataset consists of randomly selected countries across the world, it indicates to prefer random effects to fixed effects model. This assumption is confirmed by Breusch-Pagan and Hausman statistics which is reported by each model.

The following variables are considered in the models:

- Y log GDP per worker as a dependent variable,
- K log physical capital stock per capita as an independent variable,
- $TY25$ average years of schooling of 25 years and older population following Cohen and Soto methodology as an independent variable,
- $BL25$ average years of schooling of 25 years and older population following Barro and Lee methodology as an independent variable,
- $NOSCHO25$ percentage share of population age 25 and over with no schooling as an independent variable,
- $HIGH25C$ percentage share of population 25 and over with completed higher education as an independent variable.

Summary statistics of the variables are summarized in the following table:

Table 2: Summary statistics of variables

Variable	Mean	Minimum	Maximum	Std. deviation
TY25	4.91570	0.204970	12.4440	3.06772
BL25	4.48887	0.170000	12.0000	2.76463
Y	8.96281	6.63959	10.5087	0.926777
K	9.27668	5.66164	11.5468	1.38169
NOSCHO25	0.346959	0.000000	0.978161	0.299536
HIGH25C	0.0509800	0.000000	0.349253	0.0596586

Source: own calculations

Empirical results

This section shows results of data performance in growth regressions. Following equation (6), I estimate an econometric model

$$\log(y_{it}) = \beta_0 + \beta_1 \log(k_{it}) + \beta_2 \log(h_{it}) + \eta_i + \tau_t + \varepsilon_{it} \quad (8)$$

where τ and η are time and country specific effects and ε are residuals.

The main focus of presented paper is on different approximation of h . The next table presents the results of 6 models with 6 various approximations of human capital stock:

1. Barro & Lee data [BL] – h approximated as average years of schooling;
2. Barro & Lee data according to Mincerian approach [mBL] - h approximated as average years of schooling recalculated by Mincerian formula;
3. Cohen & Soto data [CS] – h approximated as weighted average years of schooling (updated Cohen & Soto methodology);¹

¹ For further details about Cohen & Soto updates to methodology please refer to Cohen and Soto (2007).

4. Cohen & Soto data according to Mincerian approach [mCS] - h approximated as average years of schooling recalculated by Mincerian formula;
5. Educational structure of population based on [OECD] data - h approximated as percentage share of population with tertiary or no education;
6. Educational structure according to Mincerian approach [mOECD] - h approximated as percentage share of population with tertiary or no education recalculated by Mincerian formula.

Table 3: Panel test regression results

	[BL]	[mBL]	[CS]	[mCS]	[OECD]	[mOECD]
	(1)	(2)	(3)	(4)	(5)	(6)
const	4.315 *** (0.222)	4.747 *** (0.277)	4.221 *** (0.23)	4.591 *** (0.29)	4.749 *** (0.398)	4.630 *** (0.453)
log k	0.516 *** (0.026)	0.508 *** (0.027)	0.530 *** (0.027)	0.521 *** (0.029)	0.488 *** (0.036)	0.488 *** (0.037)
log average years	0.179 *** (0.041)	0.164 *** (0.036)	0.141 *** (0.042)	0.132 *** (0.037)		
log % no school					-0.082 *** (0.029)	-0.082 *** (0.029)
log % higher					0.032 (0.029)	0.031 (0.029)
Akaike criterion	262	275	275	275	268	268
Hannan - Quinn criterion	266	279	279	279	274	274
Breusch-Pagan test	198.4	198.7	206.6	207.2	170.9	170.8
Hausman test	8.3	8.5	8.5	8.7	14.8	14.9
N	261	261	262	282	229	229

Note: *** indicates significance at 1% level, ** indicates significance at 5% level, * indicates significance at 10% level, std. error in parenthesis. "N" is number of observations.

Source: own calculations

The table 3 suggests interesting findings:

- 1) Physical capital contribution to output is around 50%, which is slightly under the generally acknowledged value (around 60%).
- 2) Based on Akaike and Hannan-Quinn criterion, all tested models report similar quality and relevance (the lower value the better).
- 3) Significant relationship between education and growth is reported in all models.
- 4) Human capital approximation according to Mincerian approach reports slightly lower contribution of education to growth. Microeconomic evidence suggests that additional year of education raises earnings by 10% (Mincer, 1974). Years of schooling reports 13.2%-17.9% contribution to growth.
- 5) Approximation of human capital as educational structure of labor force offers interesting finding: Tertiary schooling contribution has not been confirmed even on 10% significance level. However, there is robust negative relationship between growth and percentage share of population without schooling. Additional increase in percentage share of people without schooling can hinder the growth.

Conclusion

This paper did not try to provide an answer to an eminent question of human capital and economic growth relationship. It aimed to shed light on approximation of human capital stock and suggest suitable quantitative proxy for growth models.

Consistently with most of previous researches, the results indicate significant positive contribution of human capital to economic growth. It suggests that results are more consistent with previous papers once recalculation by Mincerian approach is employed. However, there is no significant difference if Barro & Lee or Cohen & Soto data are used.

On the other hand, different results are obtained when educational structure is used to proxy the human capital stock. On contrary to other researches, a positive link between percentage share of tertiary educated population and growth has not been confirmed but results confirm a robust negative effect of share of population with no schooling on growth. This suggests conclusion that the “right“ educational structure might be the goal for educational policy. Therefore we should not focus on permanent increase of tertiary educated people but rather diminish the uneducated share of population and improve the quality of schooling.

In general, I can not conclude which approximation of h fits best in the growth model, although educational structure offers additional interesting information. Further analyses with longer time-series should be performed and the schooling quality considered. In the forthcoming analyses additional variables for human capital proxy will be employed.

References

- AARONSON, D., SULLIVAN, D. (2001). *Growth in Worker Quality, Economic Perspectives*. Federal Reserve Bank of Chicago: Chicago
- AGHION, P., HOWITT, P. (1998). *Endogenous growth theory*. Boston: MIT.
- ARNOLD, J., BASSANINI, A., SCARPETTA, S. (2007). *Solow or Lucas? Testing growth models using panel data from OECD countries*. OECD Economic department, OECD economic studies. WP 592
- BARRO, R., SALA-I-MARTIN, X. (2004). *Economic Growth*. Boston: MIT.
- BARRO, R., LEE, J.W. (2000). *International data on educational attainment: Updates and implications*. Center for international development at Harvard University: Harvard Working paper 42, 2000.
- BENHABIB, J., SPIEGEL, M.M. (1994). The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-country Data. *Journal of Monetary Economics*. 34(2), Pp. 143-173.
- BILS, M., KLENOW, P. (2000). Does schooling cause growth? *American economic review*. 90(5). Pp. 1160-83.

- COHEN, D., SOTO, M. (2007). Growth and human capital: Good data, good results. *Journal of economic growth*. 12(1). Pp. 51-76.
- DE LA FUENTE, A. (2003). *Human capital in a global and knowledge-based economy part II.: assessment at the EU country level*. Final report. CSIC.
- DE LA FUENTE, A., DOMENECH, R. (2006). Human Capital in Growth Regression: How Much Difference Does Quality Data Make? *Journal of the European Economic Association*, 4(1). Pp. 1-36.
- DENISON, E.F. (1962). *The sources of economic growth in the United States and the alternatives before Us*, New York.
- ECB. (2006). *Growth in euro area labor quality*. Working paper no. 575/2006. Frankfurt: ECB.
- HANUSHEK, E., KIMKO, D., (2000). Schooling, labor force quality, and the growth of nations. *The American Economic Review* 90(5), 1184-1208.
- HO, M., JORGENSEN, D. (1999). *The quality of the U.S. Workforce – 1948-1998*. Harvard University.
- JORGENSEN, D. GRILICHES, Z. (1968). The Explanation of Productivity Change. *Review of Economic Studies* 34, pp. 249-280.
- KRUGER, A., LINDHAL, M. (2001). Education for growth: Why and for whom? *Journal for economic literature*. Vol. 39(4). Pp. 1101-36.
- LOENING, J. L. (1996). *Effects of primary, secondary and tertiary education on economic growth – Evidence from Guatemala*. World Bank. Working paper no. 3610/1996.
- MINCER, J. (1974). *Schooling, Earnings, and Experience*. Columbia U. Press.: New York.
- PRICHETT, L. (2001). Where has all the education gone? *World bank economic review*. 15(3). Pp. 367-391.
- HESTON, A., SUMMERS, R., ATEN, B. (2006), *Penn World Table Version 6.2*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- SODERBOM, M. (2001). *Trade and human capital as determinants of growth*. University of Oxford, Working paper 10/2001.
- SOTO, M. (2002). *Rediscovering education in growth regressions*. OECD: Paris.
- SOTO, M. (2008). *The casual effect of education on aggregate income*. OECD: Barcelona.
- TANG, T. MacLEOD, C. (2006). Labor force ageing and productivity performance in Canada. *Canadian journal of economics*. Vol. 39, No. 2, 5/2006.

LOOKING FOR THE RIGHT HUMAN CAPITAL PROXY

Author: Rudolf Kubík

Katedra veřejných financí, VŠE v Praze, nám. W. Churchilla 4, Praha 3,
rudolf.kubik@vse.cz

Abstract: This paper aims to test different approaches of human capital stock approximation. It faces one of the main questions in explaining link between human capital and economic growth. It tries to step forward in answering what is the best proxy of human capital. It starts from Barro & Lee and Cohen & Soto datasets which are expanded by Mincerian approach to human capital measurement and educational structure of population as a human capital proxy. The original dataset covering 73 countries within 1960-1990 is being re-tested and results from panel data regression analyses are compared with expanded dataset.

Key words: Human capital, growth, panel data regression

JEL Classification: E24, J24, O47

DOI: 10.2478/v10135-009-0009-0