

Session Number: Parallel Session 8C  
Time: Friday, August 27, PM

*Paper Prepared for the 31st General Conference of  
The International Association for Research in Income and Wealth*

**St. Gallen, Switzerland, August 22-28, 2010**

**“Assessing Human Capital in the National Accounts Frame – Is there a  
Feedback to Theory?”**

Arto Kokkinen

For additional information please contact:

Name: Arto Kokkinen

Affiliation: European University Institute – Statistics Finland

Email Address: arto.kokkinen@eui.eu

**This paper is posted on the following website: <http://www.iariw.org>**

## Abstract

The role of human capital in economic growth has been thoroughly discussed and emphasised in a variety of modern growth theories, especially since the 1960s. At the same time, indisputable empirical proof on the positive impacts of schooling on growth has been difficult to achieve. Some studies have reported positive, some insignificant, some even negative results on the impact of schooling on growth. The reason for the controversy could be that direct measuring of human capital is not an easy task. The typical measures for human capital in empirical studies have been literacy rate, school enrolment ratio, and, as the most prominent one, average years of schooling in the working age population.

This paper explores empirically what the feedback is to theory when human capital is assessed in the National Accounts framework. While GDP and physical capital have grown exponentially in the long run, average years of schooling, and the other proxies mentioned, grow linearly. This is not necessarily the case with an estimate for human capital in the National Accounts frame. The first feedback to growth theory concerns this.

The linear growth of average years of schooling is likely at least one of the reasons why the human capital variable often enters with an exponential structure in various growth models. In empirical studies, especially in the longitudinal ones, an exponential transformation for the average years of schooling in the working ages, in accordance with Mincerian equations, has been carried out for receiving an estimate for human capital. However, without a similar transformation in the other core variables, GDP and physical capital stock, this implicitly refers to increasing returns to schooling itself. With an estimate in the National Accounts frame this type of transformation does not seem to be needed. For instance, Kendrick's estimate through accumulated costs for the stock of education and training grows exponentially, and does not refer to increasing returns to schooling. Therefore, human capital assessed in the National Accounts can probably be entered straight-forwardly in the production function, without any assumptions.

The main objective to be studied is whether human capital assessed in the National Accounts changes the view on whether physical and human capital accumulation would be the main factors of growth or whether it has been the exogenous technical change. In empirical studies, with the proxy variables for human capital, the answer has usually favoured the latter. Instead, in accordance with the lifetime labour incomes system for human capital, Jorgenson and Fraumeni have demonstrated that the accumulation of human and non-human capital accounts for a predominant share of economic growth.

The mentioned systems of Kendrick and Jorgenson & Fraumeni have broadened the National Accounts far beyond the standard GDP. Their imputed values for non-market activities have been included in the new GDP. Instead, the studies with the proxy variables have explored the connection of schooling with the standard GDP. For reaching a fair comparison for feedback a strict long run econometric analysis is done with intangible human capital by schooling from a system of national accounts in which GDP does not have to change. In this system the education expenditures have been used as investments in human capital in Finland in 1877–2000. The stock of human capital by schooling has been accumulated by the PIM method, taking into account the long graduation times, for the years 1910–2000. The role of human capital by schooling is studied together with GDP, K and hours worked by the Vector Equilibrium Correction Model.

1. INTRODUCTION.....	3
2. THEORETICAL BACKGROUND .....	5
3. DIFFERENCES IN THE MEASURES FOR H.....	7
4. THE ROLE OF REPRODUCIBLE CAPITAL IN THE GROWTH PROCESS, THE CASE OF FINLAND IN 1910–2000 .....	12
4.1 <i>The Cointegrated VAR model</i> .....	13
4.2 <i>Empirical analysis</i> .....	17
5. CONCLUSIONS .....	36
APPENDIX I: THE MODIFIED SYSTEM OF PRODUCTION OF NATIONAL ACCOUNTS WITH INTANGIBLE HUMAN CAPITAL BY SCHOOLING .....	38
DATA SOURCES.....	44
REFERENCES .....	45

## **1. Introduction**

A theory on economic growth should explain most of the variation of GDP growth. Human capital emerged in the growth theories particularly after the 1960s<sup>1</sup> as one of the logical candidates to diminish the enlarged unexplained part, the Solow residual or multifactor productivity, with the data after the Second World War.

The modern growth theory has emphasised the role of human capital as one of the most important factors in national production of goods and services and in the incomes generated in the production process (see, for instance, Romer 1986, 1987, 1990, Lucas 1988, Rebelo 1991, Mankiw, Romer and Weil 1992, Barro and Sala-i-Martin 1999, Galor 2005, Truong and Tran-Nam 2007). Human capital is seen either as an input in production together with physical capital or as an enhancing factor for technical change and labour input. While broadly speaking human capital can be given a very wide definition, especially education and knowledge accumulation have been emphasised as the most important form of human capital for growth.

At the same time, indisputable empirical proof on the positive impacts of schooling on growth has shown to be difficult to achieve. Some studies have reported positive (see e.g. Barro, 1991, Levine and Renelt, 1992, Barro and Sala-i-Martin, 1999), some insignificant or even negative results on the impact of schooling on growth (e.g. Lau, Jamison and Louat, 1991, Islam, 1995, Bosworth and Collins, 2003). The reason for the controversy could be that direct measuring of human capital is not an easy task. The typical measures for human capital in empirical studies have been average years of schooling in the working age population, school enrolment ratio and literacy rate. A problem could arise from the fact that these conventional proxy measures for human capital are not formed in accordance with the compilation process of the systematic National Accounts as the other core variables, GDP and physical capital, are.

This paper explores whether the connection of human capital and GDP would be more evident and whether the empirical feedback to theory would be different when human capital is assessed in the National Accounts framework (NA). While GDP and physical capital have grown exponentially in the long run, average years of schooling, and the other proxies mentioned, grow linearly. This is not necessarily the case with an estimate for human capital in the National Accounts frame.

---

<sup>1</sup> The discussion on human capital and growth can be dated very far back, all the way to the days of Adam Smith. Theodore Shultz (see e.g. 1961) was one of the influential authors catalyzing the discussion in the 1960s and onwards.

The main objective to be studied is whether human capital assessed in the National Accounts changes the view on whether physical and human capital accumulation would be the main factors of growth or whether it has been the exogenous technical change or multifactor productivity. In empirical studies, with the proxy variables for human capital, the answer has usually favoured the latter (see e.g. Easterly and Levine 2001, Hall and Jones 1998).

The fundamental questions of the paper are: Will the National Accounts estimates of human capital give different implications to theory than the existing conventional measures? Is there a long run equilibrium relation to be detected with GDP and human capital by schooling estimated in the National Accounts? Will the unexplained residual, multifactor productivity or the Solow residual, be diminished with a National Accounts estimate of human capital? Has there been constant or diminishing returns to scale on the reproducible capital?

The two most well-known National Accounts systems with Human Capital by Kendrick (1976) and Jorgenson & Fraumeni (1989, 1992a, 1992b) have broadened the National Accounts far beyond the standard GDP.<sup>2</sup> Their imputed values for non-market activities in accordance with their human capital estimates have been included in the new GDP. This may make the long run relation more easily achieved among the variables. Instead, the studies with the proxy variables have explored the connection of schooling with the standard GDP.

As the first part in answering the questions average years of schooling in the working age population and human capital estimates from NA are illustrated and compared in the US and in Finland. As the second part, for reaching a fair comparison with the conventional measures for feedback, a strict long run econometric analysis is done with human capital by schooling in accordance with a system of National Accounts in which GDP does not have to change. In this system the education expenditures have been used as an input in the production of human capital and finally as investments in human capital in Finland in 1877–2000. The stock of human capital by schooling has been accumulated by the PIM method, taking into account the long graduation times, for the years 1910–2000. The role of human capital by schooling, assessed in the National accounts, is studied together with GDP, K and hours worked by the Vector Equilibrium Correction Model.

The rest of the paper is structured as follows: Section 2 reviews the theoretical background. Section 3 discusses the differences in empirical measures for human capital by education. In Section 4 the role of reproducible capital in the long run growth process is examined by time series econometrics, using an estimate for intangible human capital by schooling which is formed in the National Accounts. Section 5 concludes the paper.

---

<sup>2</sup> See Aulin-Ahmavaara (2004) for a review on the most well-known approaches on including human capital in the National Accounts.

## 2. Theoretical background

One of the major discussions regarding growth theories is whether the modern economic growth could be best modelled by neo-classical or endogenous growth models. One of the questions inside this discussion is the role of human and physical capital and whether diminishing returns or constant returns to scale with respect to reproducible capital would be prevailing in the production of GDP (see e.g. Jones 2005).

Along with the early rise of endogenous growth theories, Mankiw, Romer and Weil (1992) introduced an augmented neoclassical Solow-Swan model with exogenous technology. By using enrolment in secondary schooling as a proxy for human capital, they conducted a cross-country study and reported strong empirical support for the diminishing returns to scale on reproducible capital and for the neo-classical growth model augmented by human capital. In this model, as in the original neoclassical Solow-Swan model, economic growth would finally cease without exogenous technical change, because of diminishing returns to the factors of production. In the long run investments in physical and human capital would only have a level effect on GDP, and the long run steady state growth is determined by the exogenous technology.

Mankiw, Romer & Weil (1992) were employing a constant returns to scale Cobb-Douglas production function in the following form

$$Y_t = K_t^\alpha H_t^\beta (A_t L_t)^\gamma \quad (3.1)$$

where  $\alpha + \beta + \gamma = 1$  i.e there are constant returns to scale with respect to all of the inputs<sup>3</sup>, and at the same time  $\alpha + \beta < 1$ , referring to diminishing returns to reproducible capital. The empirical counterpart of  $Y$  is GDP,  $K$  refers to physical capital and  $H$  to human capital (secondary schooling in their analysis),  $L$  is labour input and  $A$  is the level of technology. With constant returns to scale prevailing on production the equation above can be given in intensive form  $(Y_t / L_t) = (K_t / L_t)^\alpha (H_t / L_t)^\beta A_0 e^{gt}$ , where technology is assumed to advance from the initial level,  $A_0$ , with a constant average rate of  $g$  along time (t). Log-linearising the intensive form will give

$$\ln(Y_t / L_t) = \alpha \ln(K_t / L_t) + \beta \ln(H_t / L_t) + \ln A_0 + gt \quad (3.2)$$

---

<sup>3</sup> Human capital and labour input are here separate inputs.

where  $\alpha + \beta < 1$ .

In accordance with the assumption of constant returns to scale in the production function in equation (3.1),  $\alpha + \beta$  should be less than unity, referring to decreasing returns to scale on reproducible capital,  $K$  and  $H$ . In this case the production function satisfies properties for a neo-classical model. In case of  $\alpha + \beta$  will equal unity, constant (instead of diminishing) returns to scale are prevailing on the reproducible capital supporting the endogenous growth theory models. Mankiw, Romer & Weil (1992) came to the conclusion that production function consistent with their empirical results would have  $\alpha = \beta = \gamma = 1/3$ , and therefore clearly  $\alpha + \beta < 1$ .

In contrast, a branch of endogenous growth theories has suggested that there could be constant returns to scale with respect to broad reproducible capital, including both physical and human capital.<sup>4</sup> In these models, the growth would not have to come to an end without exogenously defined technology. Investments in and the accumulation of reproducible capital would be the main drivers of the growth in the long run as well.

This is the case in an alternative, endogenous growth model, which is reviewed next. Assume a Cobb-Douglas production function with constant returns to physical and human capital,  $K$  and  $H$ :

$$Y_t = AK_t^\alpha H_t^{1-\alpha} = AK_t^\alpha (h_t L_t)^{1-\alpha} = AK_t^\alpha \left(\frac{H_t}{L_t} L_t\right)^{1-\alpha}, \quad (3.3)$$

where  $0 \leq \alpha \leq 1$ . Human capital is the number of workers,  $L$ , multiplied by the human capital of the typical worker,  $h = H / L$ . Therefore, it is not only the quantitative input of labour, but the quality adjusted labour input  $Lh$ , that is important for output in this model. If the human capital of the typical worker rises steadily, the quality adjusted labour input (i.e. human capital) here grows even if the number of workers stays constant.<sup>5</sup> The model exhibits long run growth because of constant returns to reproducible capital, including both  $K$  and  $H$ , without exogenous technological progress. (See e.g. Romer 1986, Rebelo 1991, Jones and Manuelli 1990, Barro and Sala-i-Martin, 1999, pp. 38–42, 144–146, 172–174) With constant returns to scale, the model can again be rewritten in intensive form. This results in log-linear form as:

---

<sup>4</sup> Theoretical literature often refers to  $Y = AK$  type of models, in which  $K$  includes all the reproducible capital. See, for instance, Romer (1986), Rebelo (1991), Barro and Sala-i-Martin (1999).

<sup>5</sup> Strictly theoretically the assumption is that the quantity of workers,  $L$ , and the quality of workers,  $h$ , are perfect substitutes in production in the sense that only the combination,  $Lh$ , is important for output. As a consequence, a fixed number of bodies,  $L$ , is not a source of diminishing returns in the model: a doubling of  $K$  and  $H$ , keeping  $L$  fixed, will lead to a double amount of  $Y$ .

$$LN (Y_t / L_t) = \alpha LN (K_t / L_t) + \beta LN (H_t / L_t) + LN A, \quad (3.4)$$

where  $\alpha + \beta$  will equal unity.

At a first glance, the log-linearised, intensive forms of the production functions do seem quite similar. However, the difference of the latter with the neo-classical one is important: there is no exogenous technical change (or MFP growth) needed in the long run production function and  $\alpha + \beta$  equals unity. This can be tested in empirics if there is a long run equilibrium type of cointegration relation between the variables. In that case, the estimation of the parameters including statistical inference can be done straight forwardly for the level variables in 3.4 and in 3.2.

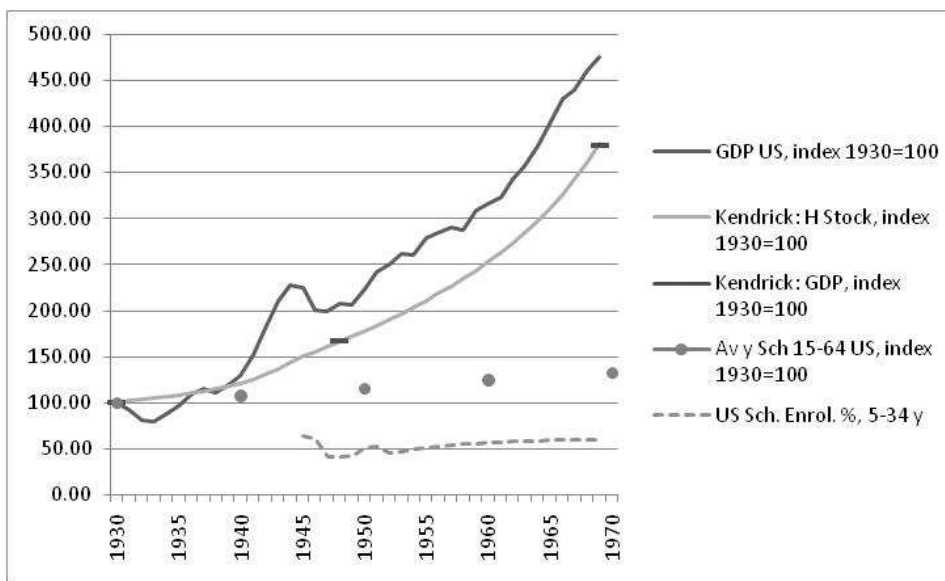
The focus of this paper is on whether the reproducible factors would be more important for the long run GDP and Labour Productivity growth than the unexplained residual, technical change or multifactor productivity, once human capital is assessed in the National Accounts frame instead of using the conventional variables. The empirical analysis in Section 4 is done for  $LN Y(t)/L(t)$ ,  $LN K(t)/L(t)$  and  $LN H(t)/L(t)$  in accordance with the equations 3.2 and 3.4, so that labour productivity is explained by physical and human capital by schooling in the labour input together with multifactor productivity or the Solow residual. For simplicity, from now on the variables will be denoted with small letters, i.e.  $y_t = LN(Y_t / L_t)$ ,  $k_t = LN(K_t / L_t)$ ,  $h_t = LN(H_t / L_t)$ .

As the first part of the analysis the differences in the evolution of the conventional measures and in the measures formed in the National Accounts will be reviewed. As the second part, a long run empirical analysis with standard GDP, physical capital, hours worked and intangible human capital by schooling assessed in the National Accounts frame is carried out for Finland in 1910–2000. In the long run empirical analysis the importance of the reproducible factors with respect to the results received by Mankiw et al. will be compared as Finland was included in their sample. It will also be studied whether the time trend in 3.2 will be diminished with the human capital estimate included.

### **3. Differences in the measures for H**

Figure 3.1 delineates the evolution of average years of schooling (15–64 year-old population), school enrolment ratio (at the ages 5–34) and real GDP in the US in 1930–1969. In addition, the figure gives Kendrick's estimates for the stock of human capital by education and training through

accumulated costs, and his new estimate for GDP in the U.S. (in 1930 (1929), 1948, 1969), as he has imputed foregone earnings of students as being part of investments in education, which are included in the new GDP. The foregone earnings of students formed a major part of investments in education and training and changed the level of GDP dramatically. The time frame in the figure comes from the calculation period of Kendrick. The figure shows inevitably that the proxy variables for schooling exhibit linear growth at the same time as GDP (with or without Kendrick's adjustments) and Kendrick's estimate for the stock of education and training through accumulated inputs grow exponentially.



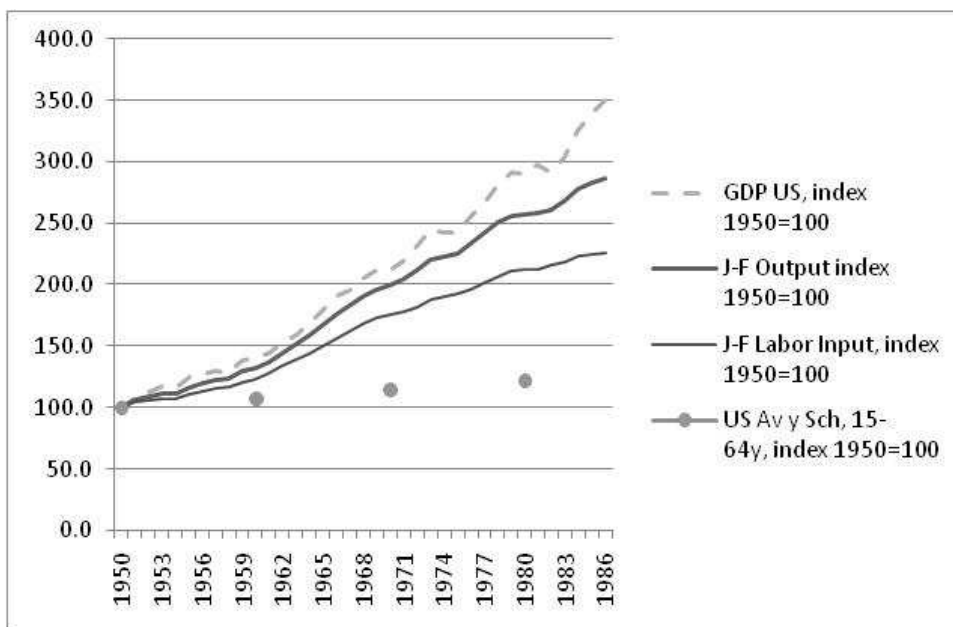
**Figure 3.1 The United States 1930–1970: Real GDP (chained 1996 billions of dollars), Kendrick’s Human Capital stock by education and training, Kendrick’s estimate for ‘new’ real GDP 1930 (1929), 1948, 1969, average years of schooling in the working age population (15–64), school enrolment % at the ages 5–34, NB: all variables except school enrolment ratio are expressed in index form, 1930=100, sources: see Data Sources**

Figure 3.2 illustrates in turn the evolution of GDP and the average years of schooling together with Jorgenson and Fraumeni’s (1992a, 1992b) estimate for quality adjusted labour input and new output of the economy, in accordance with their lifetime labour income approach for estimating the impact of investment in education on growth. In their calculations, they first estimated *educational output as the impact of education on an individual’s lifetime labour income* including labour income in market and non-market activities (time spent outside the labour market, e.g. parenting and leisure time). Therefore *the output of the education sector is defined as a measure of investment in education*. Secondly, they measured the inputs of the education sector including the outlays of educational institutions as inputs and inputs in the form of time enrolled in formal education. In their analysis, a major part of the value of the output of educational institutions accrues to students



in the form of increases in their lifetime labour incomes. By treating these increases as compensation for student time, they argue that it is possible to evaluate this time as an input into the educational process. Given the outlays of the educational institutions and the value of student time, the growth of education sector can be allocated to its sources.

Finally, they aggregated the output of education and non-education sectors of the U.S. economy to obtain a new measure of U.S. economic output for estimating the impact of investment in education on growth. The J-F output in Figure 3.2 refers to this new output. They calculated the capital input and the labour input, both in non-educational sector<sup>6</sup> and educational sectors<sup>7</sup>, and aggregated them, to allocate the growth of this new output growth of U.S. to its sources. The quality adjusted labour input in Figure 3.2 is the new labour input for U.S. economy comparable with the new output including investments in education as they have defined. The human input in this new system evolves obviously exponentially as the ‘old’ and ‘new’ output.



**Figure 3.2 The United States 1950–1986: Real GDP (chained 1996 billions of dollars), Jorgenson-Fraumeni’s new output, Jorgenson-Fraumeni’s labour input, average years of schooling in the working age population (15–64), NB: all variables are expressed in index form, 1950=100, sources: see Data Sources**

<sup>6</sup> The capital input involves weighting components of capital input by rental price. Assets were cross-classified by age, class of the asset, and legal form of organisation. Different ages were weighted in accord with profiles of relative efficiency. For the non-education sector a total of 160 components of capital input were measured separately. Hours worked for each sex were cross-classified by individual year of age and individual year of education for a total 2 196 different types of hours worked in estimating the contribution of labour input in non-education sector. Each type of hours worked was weighted by the corresponding wage rate.

<sup>7</sup> In the education sector capital input is defined as educational buildings and equipment. Labour input incorporates the value of the time teachers and other employees of educational system and student time. Intermediate goods include the purchases of educational institutions, and are included in final demand. The contributions of these inputs are obtained here as in non-educational sector by weighting the growth rates by the corresponding share of the inputs in educational output.

The result in Jorgenson and Fraumeni's (1992b) (J-F) growth accounting for the sources of growth was that the average rate of productivity growth declined from 1% p.a. to 0.5% p.a. in 1948–1986 in the U.S. economy and its contribution to growth from 31% to 17%, respectively. Labour input contributed now 61% (before 29%) and capital input 22% (before 40%) for growth. Therefore the quality adjusted labour input accounted now considerably more than before and accounted for most of the growth. *Together capital input and labour input accounted now for almost all of the growth.* They concluded that the accumulation of human and non-human capital accounts for the predominant share of economic growth.

On the basis of all of above, one could argue that assessing human capital in the National Accounts framework by either with accumulated inputs or with lifetime labour incomes, will give more weight on human capital and on reproducible capital than the traditional measures. However, both Kendrick and Jorgenson and Fraumeni systems included a substantial amount of imputed non-transaction based flows not included in the standard GDP. At the same time, the conventional measures have been used to explain the standard GDP.

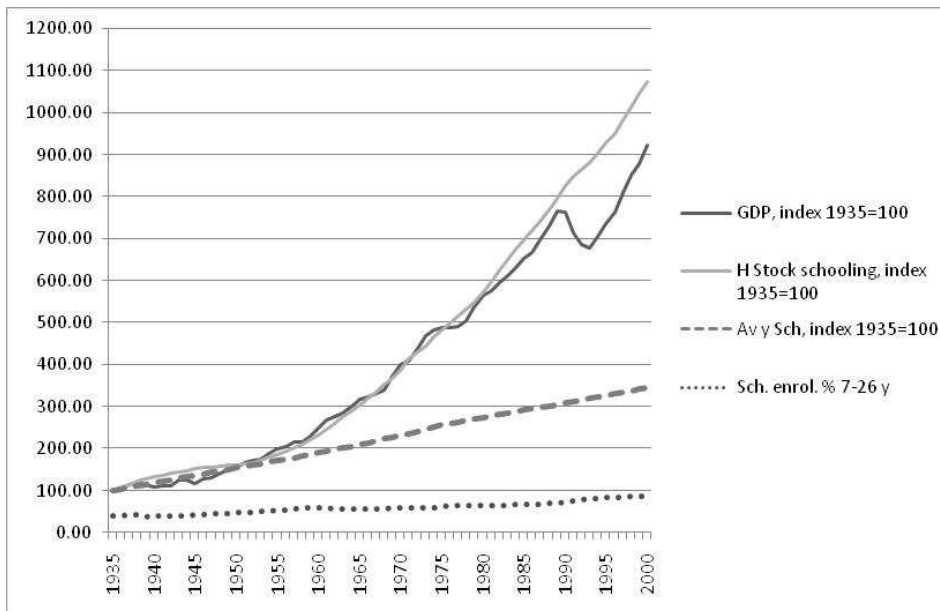
It could be argued that the imputed foregone earnings have affected the evolution of Kendrick's estimate, as the market wages used as a reference for valuation themselves have grown exponentially. Similarly, the imputed compensation for the time in the non-market activities, valued again on the basis of market wages, and the projections for the rest of the lifetime incomes in the J-F system have affected both the evolution of output and investment in education substantially. The question whether human capital by schooling assessed in the National Accounts frame can be given more weight in explaining the standard GDP based on market activities remains open.

To perceive whether this would be the case, Figure 3.3 depicts the evolution of a National Accounts estimate on intangible human capital by schooling based on paid monetary transactions on education for Finland in 1935–2000 together with real GDP and the conventional schooling measures (the average years of schooling in 16–64 year-old population, school enrolment ratio at the ages 7–26 for Finland). Here, the estimate for human capital by schooling is formed through accumulated volume of monetary inputs in education in accordance with a modified<sup>8</sup> system of National Accounts including human capital by schooling in which the GDP does not have to change. The long graduation times in education have been taken into account and the stock of

---

<sup>8</sup> Human capital is excluded from the asset boundary of the international Standard of National Accounts (SNA) in the 1993 version, currently empirically applied in the EU. The revised SNA2008 to be implemented in some years of time excludes it from the core accounts as well, but proposes it as an additional satellite account to the core system. It is argued in the SNA2008 (par 3.48) that "Human capital is not treated by the SNA as an asset. It is difficult to envisage "ownership rights" in connection with people, and even if this were sidestepped, the question of valuation is not very tractable."

human capital by schooling is accumulated by the volume of expenditures up to the time when a person has graduated from her highest education (see Appendix I). In this case the conventional measures and the national accounts estimate can be compared with the same standard GDP, which makes the examination exact. Intangible human capital by schooling through the volume of accumulated costs based on paid transactions on education is growing exponentially and very similarly to GDP, while average years of schooling are not.



**Figure 3.3 Finland 1935–2000: Real GDP (constant 2000 ref. year Millions of Euro), Intangible Human Capital stock by schooling, average years of schooling in the working age population (16–64), school enrolment % at the ages 7–26, NB: all variables except school enrolment ratio are expressed in index form, 1935=100, sources: see Data Sources**

The growth theory models aim to reflect the empirical reality. The linear growth of average years of schooling is likely at least one of the reasons why the human capital variable often enters with an exponential structure in various growth models. In empirical studies, an exponential transformation for the average years of schooling at working ages, e.g. in accordance with Mincerian equations, has been carried out for receiving an estimate for human capital.<sup>9</sup> In the longitudinal studies, GDP and physical capital have been log-linearised while average years of schooling not, referring again to an exponential transformation on the schooling variable in a non-logarithmic form (see e.g. Self and Grabowski 2003). However, without a similar transformation in the other core variables, GDP and physical capital stock, this refers implicitly to increasing returns to schooling itself in the

<sup>9</sup> For instance in the form  $H = e^{\phi u} L$ . In this formulation  $u$  is the fraction of an individual's time spent learning skills, approximated e.g. by average years of schooling, and  $\phi$  is a positive constant, in turn approximated by an overall average wage increase rate for an additional year of schooling (e.g. 0.10) in accordance with Mincerian equations. (See e.g. Jones 2002, 54-56, Bils and Klenow 2000)

production of GDP. With an estimate in the National Accounts frame this type of transformation does not seem to be needed. Therefore, the human capital variable assessed in the National Accounts can probably be entered straight-forwardly in the production function, without any assumptions. This is a first feedback for theory models of assessing human capital in the NA frame.

To summarise, the NA estimates of human capital do grow exponentially in the long run as GDP does with exhibiting an evolution much more similar to GDP than the conventional measures. Therefore, assessing human capital in the National Accounts might suggest: 1) Human Capital could have a more straight-forward relationship with GDP than assessed with average years of schooling in the working age population. The exponential structure of H in entering the production function might not be needed with national accounts estimates. 2) There seems to be a long run steady state equilibrium relationship between the evolution of human capital and GDP, and possibly with exponentially growing physical capital. 3) Human capital could get a higher weight in explaining the evolution of GDP than what the conventional measures have suggested. Together with physical capital, the unexplained residual, multifactor productivity or the Solow residual could be diminished significantly in the production function.

These suggestions will be econometrically tested in Section 4 with the data for Finland in 1910–2000 as the variable for intangible human capital by schooling was constructed in the National Accounts frame for Finland in such a way that the standard GDP did not change. The methodology for assessing intangible human capital by schooling for the econometric analysis in the Finnish case is explained in Appendix I. The Finnish case allowed for forming enough long time series for long run econometric analysis.

#### ***4. The role of reproducible capital in the growth process, the case of Finland in 1910–2000***

As noted in Section 3, the assessment of human capital in the National Accounts framework as GDP and physical capital may change the empirical implications to theory compared with other measures. The role of intangible human capital by schooling assessed in the National Accounts frame will be scrutinised in depth in this section by Vector Equilibrium Correction models in the Cointegrated Vector Autoregressive model framework (Engle and Granger, 1987, Johansen, 1996).

## 4.1 The Cointegrated VAR model

The VAR model with two lags in the unrestricted form in levels is given by<sup>10</sup>:

$$\mathbf{x}_t = \mathbf{\Pi}_1 \mathbf{x}_{t-1} + \mathbf{\Pi}_2 \mathbf{x}_{t-2} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim iid N_p(0, \Omega) \quad (4.1.1)$$

The Cointegrated VAR model (CVAR) in the vector equilibrium correction form can be derived by subtracting  $\mathbf{x}_{t-1}$  from both sides of 4.1.1, and can be expressed equivalently in terms of likelihood as  $\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \mathbf{\Gamma} \Delta \mathbf{x}_{t-1} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t$ . The reduced form with the deterministic components estimated in this paper is expressed in 4.1.2:

$$\Delta \mathbf{x}_t = \mathbf{\Pi} \mathbf{x}_{t-1} + \mathbf{\Gamma} \Delta \mathbf{x}_{t-1} + \boldsymbol{\mu}_0 + \boldsymbol{\alpha} \boldsymbol{\beta}'_1 \mathbf{t} + \mathbf{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim iid N_p(0, \Omega) \quad (4.1.2),$$

where  $\mathbf{x}_t = (y_t, k_t, h_t)'$  in the empirical analysis following, with small letters in the variables referring to a variable divided by number of hours worked and expressed in natural logarithms,  $\mathbf{\Pi} = -(\mathbf{I} - \mathbf{\Pi}_1 - \mathbf{\Pi}_2)$  and  $\mathbf{\Gamma} = -\mathbf{\Pi}_2$ ,  $\boldsymbol{\mu}_0$  is a vector of constants,  $\mathbf{D}_t$  a vector of dummies,  $t$  is the time trend restricted to cointegration relations. In the case of  $I(1)$ -analysis the rank of the coefficient matrix  $\mathbf{\Pi}$  can be used to test the number of stationary cointegration relations (which is the rank,  $r$ , of  $\mathbf{\Pi}$ ) between the levels of the variables and the number of unit roots, i.e. common trends (with  $p$  variables, the number of common trends is  $p-r$ ).

If there exists  $r$  cointegration relations, the matrix  $\mathbf{\Pi}$  has a reduced rank, and  $\mathbf{\Pi}$  can be expressed as  $\boldsymbol{\alpha} \boldsymbol{\beta}'$ , where  $\boldsymbol{\alpha}, \boldsymbol{\beta}$  are  $p \times r$ . The transposed vector  $\boldsymbol{\beta}'$  includes the long run cointegration coefficients and the vector  $\boldsymbol{\alpha}$  adjustment coefficients for the variables under review. The constant,  $\boldsymbol{\mu}_0$ , appears unrestrictedly in the model generating a time trend in the Moving Average (MA) form. The time trend in the Autoregressive representation of the model above is restricted to the cointegration relations, thus denoted by  $\boldsymbol{\alpha} \boldsymbol{\beta}'_1 \mathbf{t}$ , implying to a time trend in at least one of the cointegration relations.

It is worth noticing that the point estimates of the parameters are exactly the same in 4.1.1 and 4.1.2 in the sense that the parameter estimates of 4.1.1 can be simply calculated from

---

<sup>10</sup> The VAR model with lag length of two is reviewed since the lag length was inferred to be two by both Schwarz and Hannan-Quinn criteria in the following empirical analysis.

4.1.2 by  $\mathbf{\Pi} = -(\mathbf{I} - \mathbf{\Pi}_1 - \mathbf{\Pi}_2)$  and  $\mathbf{\Gamma} = -\mathbf{\Pi}_2$ . The latter form is favoured in estimation since statistical significance on the parameters of the level variables can only be inferred by 4.1.2 in the case of cointegration.

Inverting the VAR model gives the Moving Average (MA) representation defining the pushing forces of the system or the common stochastic trends. The MA form, assuming no dummies for the moment, is given by:

$$\mathbf{x}_t = \mathbf{C} \sum_{i=1}^t \boldsymbol{\varepsilon}_i + t\mathbf{C}\boldsymbol{\mu}_0 + \mathbf{C}^*(L)(\boldsymbol{\varepsilon}_t + \boldsymbol{\mu}_0 + \boldsymbol{\alpha}\boldsymbol{\beta}_1 t) + \mathbf{X}_0 \quad . \quad (4.1.3)$$

where  $\mathbf{C} = \boldsymbol{\beta}_\perp (\boldsymbol{\alpha}'_\perp \boldsymbol{\Gamma} \boldsymbol{\beta}_\perp)^{-1} \boldsymbol{\alpha}'_\perp$  or  $\mathbf{C} = \tilde{\boldsymbol{\beta}}_\perp \boldsymbol{\alpha}'_\perp$ ,  $\mathbf{X}_0$  contains both the initial value,  $\mathbf{x}_0$ , of the process  $\mathbf{x}_t$  and the initial value of the short-run dynamics  $\mathbf{C}^*(L)\boldsymbol{\varepsilon}_0$ . Equation (4.1.3) shows that the evolution of the level variables  $\mathbf{x}_t$  can be described by stochastic trends  $\mathbf{C} \sum_{i=1}^t \boldsymbol{\varepsilon}_i$  (which can also be denoted as  $\tilde{\boldsymbol{\beta}}_\perp \boldsymbol{\alpha}'_\perp \sum_{i=1}^t \boldsymbol{\varepsilon}_i$ ), linear time trend cumulated by the constant  $\boldsymbol{\mu}_0 t$  (multiplied by  $\mathbf{C}$ ) and stationary stochastic components  $\mathbf{C}^*(L)\boldsymbol{\varepsilon}_t$ , and initial values (Juselius, 2006, p. 256).

For given  $\boldsymbol{\alpha}$  and  $\boldsymbol{\beta}$  one can find the orthogonal complements,  $\boldsymbol{\alpha}'_\perp$  and  $\boldsymbol{\beta}_\perp$  of dimension  $p \times (p - r)$  and of full rank so that  $\text{rank}(\boldsymbol{\alpha}'\boldsymbol{\alpha}_\perp) = p$ ,  $\text{rank}(\boldsymbol{\beta}'\boldsymbol{\beta}_\perp) = p$ . These orthogonal complements can be used to decompose the long run impact matrix  $\mathbf{C}$  in the MA form as shown above. When the time trend is restricted to the cointegration relations, it will appear in the Moving Average representation in the stationary part, and hence, is not affecting the non-stationary part. The time trend of the levels of the variables is induced by the unrestricted constant in the Autoregressive form of the model.

The decomposition of the  $\mathbf{C}$  matrix is similar to the  $\mathbf{\Pi}$  matrix: however, in the AR representation  $\boldsymbol{\beta}$  determines the common long-run relations and  $\boldsymbol{\alpha}$  the loadings, whereas in the moving average representation  $\boldsymbol{\alpha}'_\perp$  determines the common stochastic trends and  $\tilde{\boldsymbol{\beta}}_\perp$  their loadings. The non-stationarity in the process  $\mathbf{x}_t$  originates from the cumulative sum of the combinations

$\alpha'_{\perp} \sum_{i=1}^t \varepsilon_i$ . In the case of an I(1)-process the number of such combinations is  $p - r$ . The common driving trends are defined as the variables  $\alpha'_{\perp} \sum_{i=1}^t \varepsilon_i$ . (Johansen 1996, Juselius 2006)

It is worth noticing that the matrices  $\alpha_{\perp}$  and  $\beta_{\perp}$  can be directly calculated for given estimates of  $\alpha$ ,  $\beta$ , (and  $\Gamma = -(\mathbf{I} - \Gamma_1 - \Gamma_2 - \dots - \Gamma_{k-1})$ ), with lag length  $k$ , see Johansen (1996) Chapter 4). This means that the common stochastic trends and their weights can be found either based on unrestricted  $\hat{\alpha}$ ,  $\hat{\beta}$  or on restricted estimates  $\hat{\alpha}^c$ ,  $\hat{\beta}^c$ . The CATS program used later for conducting the empirical analysis uses the latest estimates of  $\alpha$ ,  $\beta$  as a basis for the calculations in the moving average representation.

The notion of cointegration relations  $\beta'x_t$ , and the notion of common trends  $\alpha'_{\perp} \sum_{i=1}^t \varepsilon_i$  are two sides of the same coin, as are the adjustment coefficients  $\alpha$  and the loading coefficients  $\tilde{\beta}_{\perp}$ . The cointegrated VAR model provides a general framework within which one can describe economic behaviour in terms of forces pulling towards equilibrium, generating stationary behaviour ( $\beta'x_t$ ), and forces pushing away from equilibrium, generating non-stationary behaviour ( $\alpha'_{\perp} \sum_{i=1}^t \varepsilon_i$ ).

Let us open the notation in the case of one cointegration relation in the Autoregressive representation for the first equation, for  $\Delta y_t$ , of the model with the variables in the following analysis (three variables  $y_t, k_t, h_t$ ;  $p = 3, r = 1, p - r = 3 - 1 = 2$  common trends, all variables  $\sim I(1)$ ).<sup>11</sup>

$$\Pi \mathbf{X}_{t-1} = \alpha \beta' \mathbf{x}_{t-1} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} (\beta_1 y_{t-1} + \beta_2 k_{t-1} + \beta_3 h_{t-1}).$$

It is often useful to normalize the cointegration relation by the coefficient of one of the variables. If we normalize on  $y_t$ , the equation for the first variable,  $\Delta y_t$ , can be given the usual equilibrium correction form

$$\Delta y_t = \alpha_1 \beta_1 (y_{t-1} + \frac{\beta_2}{\beta_1} k_{t-1} + \frac{\beta_3}{\beta_1} h_{t-1}) + \Gamma_{11} \Delta y_{t-1} + \Gamma_{12} \Delta k_{t-1} + \Gamma_{13} \Delta h_{t-1} + \varepsilon_{1t}, \quad \text{or}$$

$$\Delta y_t = \bar{\alpha}_1 (y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) + \Gamma_{11} \Delta y_{t-1} + \Gamma_{12} \Delta k_{t-1} + \Gamma_{13} \Delta h_{t-1} + \varepsilon_{1t}, \quad ,$$

<sup>11</sup> See *CATS in RATS, version 2*, by J.G. Dennis, Hansen H., Johansen S. and Juselius K., Estima 2005.

where the bars above the alphas and betas refer to the normalisation, in which the  $\beta$  parameters are divided by the beta-parameter of the normalised variable (and  $\alpha$  :s multiplied by it).

The cointegration relation in the parenthesis above is stationary, which is why it is often interpreted as the long-run equilibrium for the levels  $\mathbf{x}_t = (y_t \ k_t \ h_t)'$ . If  $\beta'X_t \neq 0$ , it is interpreted as a long-run disequilibrium error and for fixed lags, the loading  $\bar{\alpha}_1$  captures its effect on  $\Delta y_t$ . The growth of GDP per hours worked is explained above i) by the stationary cointegration or equilibrium correction relation of itself with physical capital per hours worked and human capital per hours worked, and ii) by the one lag differenced values of itself and the other two variables mentioned.

The whole CVAR model in the Vector Equilibrium Correction (VEC) form in the case of one cointegration relation is given in (4.1.4). In this three-equation system, the growth rates of  $\Delta y_t$ ,  $\Delta k_t$  and  $\Delta h_t$  (differenced variables expressed in natural logarithms) are each explained at the same time by the stationary cointegration relation, with each having their own adjustment parameter,  $\alpha_i$ , and by the growth rate with one lag of the growth rate of each variable:

$$\begin{cases} \Delta y_t = \bar{\alpha}_1(y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) & + \Gamma_{11} \Delta y_{t-1} + \Gamma_{12} \Delta k_{t-1} + \Gamma_{13} \Delta h_{t-1}) + \varepsilon_{1t} \\ \Delta k_t = \bar{\alpha}_2(y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) & + \Gamma_{21} \Delta y_{t-1} + \Gamma_{22} \Delta k_{t-1} + \Gamma_{23} \Delta h_{t-1}) + \varepsilon_{2t} \\ \Delta h_t = \bar{\alpha}_3(y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) & + \Gamma_{31} \Delta y_{t-1} + \Gamma_{32} \Delta k_{t-1} + \Gamma_{33} \Delta h_{t-1}) + \varepsilon_{3t} \end{cases} \quad (4.1.4)$$

As noted above with the number of variables equalling to 3,  $p = 3$ , one cointegration relation,  $r = 1$ , the number of common stochastic trends,  $p - r = 3 - 1 = 2$ , refers to two common trends which are combinations of the cumulated residuals  $\mathbf{a}'_{\perp,1} \sum_{i=1}^t \varepsilon_i$ ,  $\mathbf{a}'_{\perp,2} \sum_{i=1}^t \varepsilon_i$  with loadings to the variables  $\tilde{\beta}'_{\perp,1}$  and  $\tilde{\beta}'_{\perp,2}$ .



## 4.2 Empirical analysis

The data should be reasonably long for time series econometrics modelling. In Finland the general government sector has provided all of the education services in 1975–2000 and almost all in the previous history.<sup>12</sup> This enabled to gather the data for the estimation of the evolution of the monetary flows paid for education in Finland backwards until 1877 and apply them with the Historical National Accounts data set, available in Finland from 1860 onwards. Consequently, the Finnish data allow for the econometric testing.

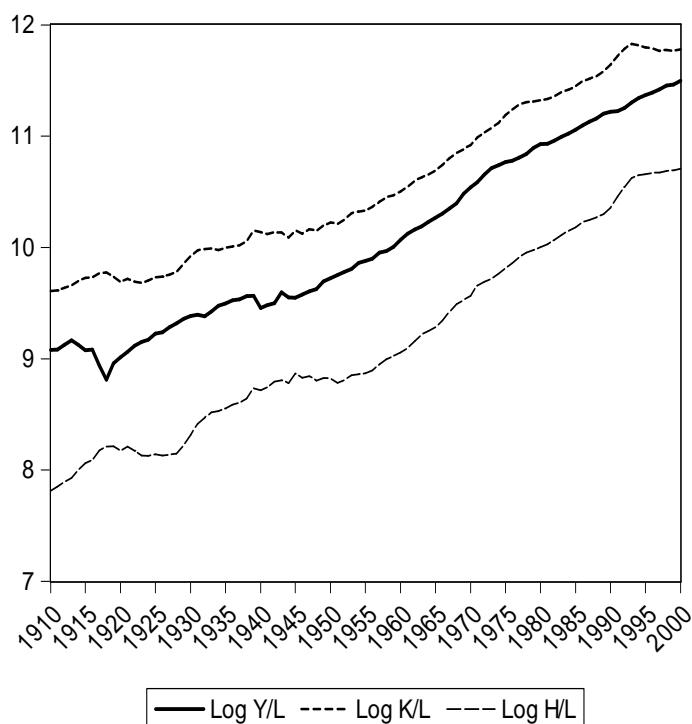
GDP per person can be expressed as a product  $(\text{GDP}/\text{hours worked}) \times (\text{hours worked}/\text{person})$ . The rise in the former component explains the most of the long run growth of GDP per capita. The long run implications of the growth theories are often derived in intensive form for labour productivity growth with variables expressed in proportion to labour input or to efficient units of labour.

In the empirical analysis that follows, a time trend was originally included in both of the cointegration relations to start with the most unrestricted model. Figures 4.2.1 a) and b) demonstrate the evolution of the levels of the variables in natural logarithms and their growth rates along time for Finland in 1910–2000. The first phase of modelling is to determine the rank of PI. Before of that, it is required that the residuals are non-autocorrelated white noise. For this, the non-normal large shocks to the variables are needed to be modelled by deterministic dummies.

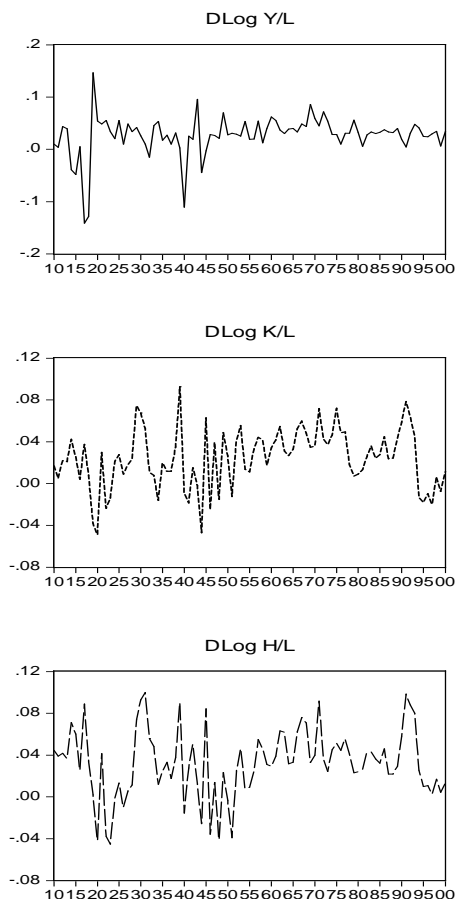
All of the variables are trending upwards and the inclusion of a time trend (possibly to be interpreted as the long run average Solow residual or Multi Factor Productivity, MFP) in the cointegration relations seems to be feasible, which refers to a possibility that the series have both deterministic and stochastic trends. Looking at the first differences of the variables it becomes obvious that there are two periods containing probably outliers: the end of WWI around 1917-1919 (including the year for the Russian revolution and Finland gaining independence, 1917, the civil war year 1918 in Finland and the beginning of the recovery 1919) and the years of WWII 1939–1945 (in Finland the wars ended in 1944 and 1945 is the first year of recovery).

---

<sup>12</sup> The state and local authorities have produced practically all of the primary and secondary education in Finland in the time period under review. From the late 1960s all of the professional/vocational and university education has been organised by the public sector as well. In the case of private organisers of professional education in the late nineteenth and early twentieth century, the state and the municipalities were financing them in the form of subsidies. The University of Helsinki, the biggest and the oldest university has been owned by the state. Education in the few privately organised higher education institutes from 1910 onwards were subsidised by the general government.



a) Levels (later vector  $x_t$ )



b) Growth rates ( $\Delta x_t$ )

**Figures 4.2.1 a) Physical capital per number of hours worked (Log K/L), GDP per number of hours worked (Log Y/L) and human capital per number of hours worked (Log H/L) in Finland in 1914–2000, in natural logarithmic form. b) The LN growth rates of the same variables, sources: see Data Sources**

There seems to be a structural break by the end of WWII, in 1944 or 1945, which is particularly detectable in the labour productivity growth. The growth rate of  $h_t$  accelerates from the 1950s to the early 1970s with a blip around 1970–1971. The accelerating growth is connected with the large after the war age groups entering the schools together with expanding education. The blip could be associated with the comprehensive school reform initiated at this time, with expanding the compulsory education to nine years in Finland. One additional interesting aspect can be noticed from Figure 4.2.1b): the growth rates of  $k_t$  and  $h_t$  seem to be surprisingly similar when considering how differently the estimate for H has been cumulated taking into account the long graduation times from all of a person’s formal education (for instance 20 years for a person graduating from the university).

In the end, after testing the residuals for several outlier combinations a level shift restricted to the cointegration relations was set to 1944 as WWII ended in Finland in 1944, to separate the development before and after WWII in the long run relations of the variables and to take into account the recovery from the war with a lot of physical and human capital destroyed.. For the other outliers in the WWII era, a transitory impulse dummy was used for 1939–1943, and a permanent impulse dummy was set to 1940, which, together with 1944 were perhaps the most destructive war years for Finland. A permanent impulse dummy was set for the first year after the war, 1945, to describe the first recovery year. To account for the turbulent end of the WWI period including a civil war in 1918, a permanent impulse dummy was located at 1917 and a transitory impulse dummy for the years 1918–1919.

The dummies included are specified below:

- A level shift dummy for 1944:  $D_{s1944_t} = 1$  for  $t \geq 1944$ , zero otherwise;
- A permanent impulse dummy for the year 1917,  $D_{pt} = 1_{1917}$ , zero otherwise;  
and for the year 1945,  $D_{pt} = 1_{1945}$ , zero otherwise and 1971,  $D_{pt} = 1_{1971}$ , zero otherwise.
- A transitory intervention dummy for the years 1918–1919:  $D_t = -1_{1918}, 1_{1919}$ , zero otherwise; for the war period 1939–1943:  $D_t = -1_{1939}, 0_{1940}, 0_{1941}, 0_{1942}, 1_{1943}$ , zero otherwise.

After the inclusion of these deterministic variables, the residuals of the baseline model in equation 4.1.2 were independently, identically and normally distributed white noise, see Table 4.2.1, as is assumed in the VAR model. The LM-tests with 1 to 2 lags of no multivariate residual autocorrelation could not be rejected with p-values 0.34 and 0.47. A little sign of minor ARCH effects of multivariate residuals was detected as LM-test p-value of no ARCH effects with one lag was 0.041, however, the hypothesis of no ARCH effects with two lags could not be rejected with a p-value of 0.09. The joint normality test was accepted with a p-value of 0.33.

*The unrestricted baseline VAR model* was first estimated, and the results are shown in Table 4.2.2. It can be detected that there are statistically significantly adjusting parameters (with  $|t| > 3.51$  which is the 5% critical value in the Dickey-Fuller distribution) in the two of the alpha-vectors implying to two cointegration relations. The multivariate residuals of the model are assumed to be independently, identically and normally distributed with mean zero and a constant variance. The misspecification tests given below show that this assumption cannot be rejected after including the deterministic dummies defined above. The cross correlation between the residuals of  $\Delta k_t$  and

$\Delta h_t$  is very high, 0.88, which is not a surprise looking at the growth rates of the variables. This can be interpreted as a strong sign of capital and skills being complements.

**Table 4.2.1 Misspecification tests for the unrestricted VAR(2) ,**  
t-ratios in brackets [ ], dummies included

<i>Multivariate tests</i>			
Residual autocorrelation:	<i>p</i> -value		<i>p</i> -value
$LM_1 : \chi^2(9)$	0.34	$LM_2 : \chi^2(9)$	0.47
Normality:			
$LM : \chi^2(6)$	0.33		
ARCH:			
$LM_1 : \chi^2(36)$	0.041	$LM_2 : \chi^2(72)$	0.089
<i>Univariate residual std. errors and cross-correlations :</i>			
		$\mathcal{E}_{\Delta y}$	$\mathcal{E}_{\Delta k}$
		0.017	0.018
			$\mathcal{E}_{\Delta h}$
			0.020
		$\mathcal{E}_{\Delta y}$	1
		$\mathcal{E}_{\Delta k}$	0.052
		$\mathcal{E}_{\Delta h}$	-0.099
			0.880
			1
<i>Trace correlation</i>	0.72		

**Table 4.2.2 The unrestricted baseline VAR(2) model in the VECM form ,**  
t-ratios in brackets [ ]

	$y_t$	$k_t$	$h_t$	$D_{s1944}$	$t$
$\beta'_1$	3.6	-16.8	18.1	3.7	-0.31
$\beta'_2$	16.6	-16.8	6.0	1.6	-0.27
$\beta'_3$	-9.3	17.2	-1.6	-0.24	-0.15
	$\alpha_1$	$\alpha_2$	$\alpha_3$		
$y_t$	0.003 [0.14]	<b>-0.008</b> [-4.22]	-0.003 [-1.62]		
$k_t$	<b>-0.0084</b> [-4.43]	0.0029 [1.54]	-0.005 [-2.69]		
$h_t$	<b>-0.015</b> [- 6.92]	0.0030 [1.40]	0.0029 [1.37]		

### Determining the cointegration rank (rank of PI)

As mentioned above, when the process is  $I(1)$ , the number of unit roots equal to  $p - r$ , which is the number of stochastic trends in the system. There are four different ways to gather information for making inference on the rank of PI,  $r$ . First, as discussed above, the significance of the coefficients (in accordance with the DF-distribution) in the alpha-vectors in the unrestricted VAR model can be used in inference of the rank. Secondly, a trace test (possibly with a small sample correction) with simulated critical values (adjusted for dummies, etc.) can be carried out in the program CATS.

**Table 4.2.3 Trace test and the characteristic roots**

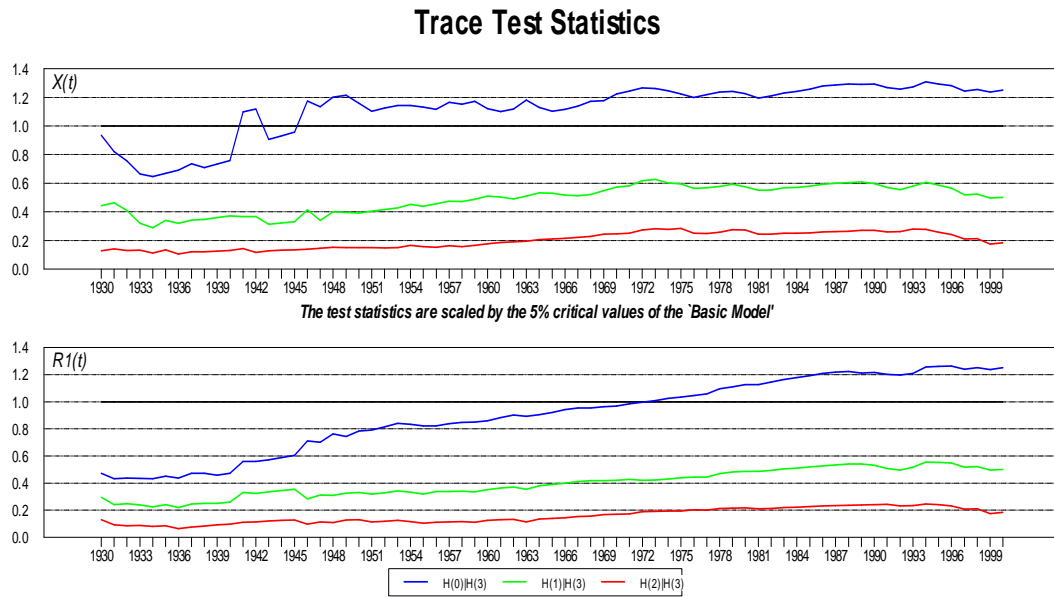
Trace refers to simulated asymptotic distribution with respect to deterministic components in the model  
 \*small sample Bartlett corrected, refers to the model without deterministic components

$r$	$p - r$	Trace	Trace*	Modulus of the 5 largest characteristic roots				
		$p$ -value	$p$ -value					
0	3	0.0000	0.0001	1	1	1	0.58	0.58
1	2	0.0327	0.0558	1	1	0.76	0.76	0.31
2	1	<b>0.1675</b>	<b>0.2231</b>	1	<b>0.81</b>	0.73	0.73	0.40

Table 4.2.3 describes that with deterministic components (dummies) in the analysis, the trace test suggest the rank of PI to be two ( $p$ -value\*=0.22). Rank equal to one, ( $r = 1$ ), with a small sample corrected  $p$ -value\* =0.0558, is on the border of acceptance. The trace test without dummies was clearly rejecting  $r = 1$  and suggesting  $r = 2$ , as well as the simulated  $p$ -values above without the small sample Bartlett correction. Rank equal to zero implying to no cointegration relations was statistically highly significantly rejected.

As a third step, the rank can be further detected by reducing the rank of  $\Pi$  in the model, and hence imposing an increasing number of unit roots ( $p-r$ ) in the model. With the feasible rank the next largest root of the characteristic function after the unit roots should be clearly less than 1. This enables to suggest a choice for  $r$  as well.

As Table 4.2.3 illustrates, with 1 stochastic trend (unit root) in the model, and therefore rank = 2, the next largest root gets a value of 0.8, which can obviously be considered a non-unit root (the difficult area of testing a unit root is in the range of values between 0.95 – 1). With two stochastic trends, implying to rank equal to one, the next root will get a value of approximately the same size as in the previous case, 0.76. These observations support a choice of  $r = 2$ .



**Figure 4.2.2 The recursively estimated trace test statistics**

Fourthly and finally, the trace test statistics were estimated with recursive samples without the deterministic components in the model (see in Figure 4.2.2 especially the lower graph, the test with a model from which the short run effects have been concentrated out). The graph of the test is implying undoubtedly at least  $r = 1$ , as the test for  $r = 1$  surpasses the critical value of the 95% quantile of the appropriate asymptotic distribution (this 5% critical value is scaled to one in the graph). The trace test for  $r = 2$  does not cross the critical value, however it exhibits a linear growth over time, and with longer data would probably surpass the critical value. Very strictly interpreting the results, the recursively estimated trace test would indicate  $r = 1$ . However, if the second linear combination of the levels of the series would be non-stationary, the recursive trace test should not display linear growth with longer data. This is again supporting  $r = 2$ , and the typical inference here would be  $r = 2$ .

With all of the above analysis, it can be inferred that the cointegration rank of the model should most likely to be determined to two, with some indication of  $r = 1$  being on the borderline of acceptance. Therefore, the focus is first on the analysis of two cointegration relations and the hypotheses on which of the variables form the two equilibrium correction relations will be next tested. After this the case of  $r = 1$  will be reviewed.

*Analysis with two cointegration relations between  $y_t$ ,  $k_t$  and  $h_t$*

The analysis with  $r = 2$  is based on the three-equation system with the time trend and the level shift (both restricted to cointegration relations) included in each of the long run relations to begin with. In the case of two (or more) cointegration relations, a lot of work may occur for identifying the model by testing hypothesis.<sup>13</sup> This means setting and testing restrictions on the  $\bar{\beta}$ -coefficients.

The three-equation model with rank equal to two is presented below. The first cointegration relation is normalised on  $y_t$ , the second one is chosen to be normalised on  $h_t$  at least to start with, referring to the possibility that either  $k_t$  or  $y_t$  and possibly the time trend and level shift could explain the long run development of it. The short run part of the equations, one differenced lag of all of the variables, is denoted briefly with  $\Gamma\Delta\mathbf{x}_{t-1}$ .

$$\begin{cases} \Delta y_t = \bar{\alpha}_{11}(y_{t-1} + \bar{\beta}_{12}k_{t-1} + \bar{\beta}_{13}h_{t-1} + D_s 44 + g_1 t) + \bar{\alpha}_{12}(\bar{\beta}_{21}y_{t-1} + \bar{\beta}_{22}k_{t-1} + h_{t-1} + D_s 44 + g_2 t) + \Gamma\Delta\mathbf{x}_{t-1} + \varepsilon_{1t} \\ \Delta k_t = \bar{\alpha}_{21}(y_{t-1} + \bar{\beta}_{12}k_{t-1} + \bar{\beta}_{13}h_{t-1} + D_s 44 + g_1 t) + \bar{\alpha}_{22}(\bar{\beta}_{21}y_{t-1} + \bar{\beta}_{22}k_{t-1} + h_{t-1} + D_s 44 + g_2 t) + \Gamma\Delta\mathbf{x}_{t-1} + \varepsilon_{2t} \\ \Delta h_t = \bar{\alpha}_{31}(y_{t-1} + \bar{\beta}_{12}k_{t-1} + \bar{\beta}_{13}h_{t-1} + D_s 44 + g_1 t) + \bar{\alpha}_{32}(\bar{\beta}_{21}y_{t-1} + \bar{\beta}_{22}k_{t-1} + h_{t-1} + D_s 44 + g_2 t) + \Gamma\Delta\mathbf{x}_{t-1} + \varepsilon_{3t} \end{cases}$$

Before proceeding to testing hypothesis, the baseline VAR is estimated with rank two and the *general model specific data properties can be tested*. These model specific data properties refer to general tests whether some of the variables could be excluded from the model, whether there are stationary (or trend stationary when a time trend is included) variables in the analysis, *whether some of the variables are the primarily purely adjusting pulling forces and some define the pushing forces*. The two latter tests can be conducted by imposing restrictions on alpha vector, i.e. whether some of the variables are always adjusting to the shocks of all other variables and whether some do not adjust at all to shocks in other variables but cannot be excluded from the cointegration relations.

According to these general tests, none of the variables could be excluded, none were stationary with the time trend and the level shift included and all of the variables proved to be adjusting at least to one variable in the model. Instead, as the results with the baseline VAR for  $r = 2$  below indicate,  $y_t$  is only adjusting to  $k_t$  and  $h_t$  in one relation, in which  $k_t$  and  $h_t$  are not

<sup>13</sup> The problem is that any linear combination of the two cointegration relations will preserve the stationarity property. In other words, when the cointegration rank is larger than one there is an identification problem: it is the space spanned by  $\beta$  and not  $\beta$  itself which is uniquely determined. The econometric programs (CATS is used here) provides procedures for testing structural hypotheses on the cointegration space. These procedures allow the user to impose and test hypothesis by identifying restrictions on the cointegration vectors.<sup>13</sup>

adjusting at all. In the other cointegration relation  $k_t$  and  $h_t$  are the only variables adjusting. This suggests a test of a known alpha vector where  $y_t$  is the pulling force and only adjusting to the other variables to maintain the long run cointegration equilibrium, whereas the other variables form the pushing forces of the system. This test of a unit vector in alpha on  $y_t$  could not be rejected with a p-value 0.45. *The Granger causality seems to run from  $k_t$  and  $h_t$  to  $y_t$  in the long run.*

**Table 4.2.4 The unrestricted VECM with  $r = 2$  ,**  
t-ratios in brackets [ ]

	$y_t$	$k_t$	$h_t$	$D_{s1944}$	$t$
$\beta'_1$	0.196	-0.92	1	0.20	-0.0169
$\beta'_2$	1	-1.01	0.36	0.09	-0.0161
	<b><math>\alpha_1</math></b>	<b><math>\alpha_2</math></b>			
$y_t$	0.046 [0.14]	<b>-0.13</b> [-4.16]			
$k_t$	<b>-0.15</b> [-4.27]	0.049 [1.48]			
$h_t$	<b>-0.27</b> [-6.85]	0.049 [1.38]			

**Table 4.2.5 The long run structure with unit vector in alpha on  $y$ ,**  
LR test of restricted model  $\chi^2(1)$ , p - value [0.45],  
t-ratios in brackets [ ]

	$y_t$	$k_t$	$h_t$	$D_{s1944}$	$t$
$\beta'_1$	1	-0.88	0.35	0.096	-0.0195
$\beta'_2$	0.096	-0.86	1	0.20	-0.0155
	<b><math>\alpha_1</math></b>	<b><math>\alpha_2</math></b>			
$y_t$	<b>-0.12</b> [-4.50]	0.00 [0.00]			
$k_t$	0.00 [0.00]	<b>-0.15</b> [-4.29]			
$h_t$	0.00 [0.00]	<b>-0.26</b> [-6.84]			



**Table 4.2.6 The MA representation corresponding to a unit vector in alpha on y,**  
t-ratios in brackets [ ]

	$\mathcal{E}_{yt}$	$\mathcal{E}_{kt}$	$\mathcal{E}_{ht}$	
$\sigma_{\varepsilon_i}$	0.0241	0.0402	0.0323	
Common trend (CT) weights				
$\alpha'_{\perp}$	-0.00	1	-0.58	
Loadings to CT				
$\tilde{\beta}'_{\perp}$	2.36	3.94	3.17	
The Long Run Impact matrix C				
	$\mathcal{E}_{yt}$	$\mathcal{E}_{kt}$	$\mathcal{E}_{ht}$	$t \cdot \gamma_{0,i}$
$y_t$	0.00 [0.00]	<b>2.36</b> [2.92]	<b>-1.365</b> [-2.08]	0.0301
$k_t$	0.00 [0.00]	<b>3.94</b> [2.92]	<b>-2.28</b> [-2.08]	0.0260
$h_t$	0.00 [0.00]	<b>3.17</b> [2.92]	<b>-1.83</b> [-2.08]	0.0350

The MA representation with a unit vector in alpha is given in Table 4.2.6. It shows that, as expected here, the common stochastic trend is defined by the combination of the cumulated residuals of  $k_t$  and  $h_t$ , and  $y_t$  has not had any long run effects, only transitory effects, to the evolution of the levels

of the variables in the system. The common stochastic trend is defined here as  $\left[ \sum_{i=1}^t \varepsilon_{ki} - 0.58 \sum_{i=1}^t \varepsilon_{hi} \right]$

with positive loadings for each variable  $\tilde{\beta}_{\perp} = [2.4 \quad 3.9 \quad 3.2]'$ . An intuitive description would be that the economy and each of the variables have grown when an innovation in  $k_t$  has been larger than 0.58 of an innovation in  $h_t$ .

*The purpose of the following analysis is to identify the two cointegration relation in the system in such a way that they are distinct from each other, and the system can be equally, in terms of likelihood of the model, presented with a reduced number of parameters.*

The main hypotheses to be tested in the case of two cointegration relations are: Is there a long run relation between  $y_t$ ,  $k_t$  and  $h_t$  to be found referring to a Cobb-Douglas type production function in equations 3.2 and 3.4? If there is, are the coefficients of  $k_t$  and  $h_t$  summing up to one or less than one implying either to constant or decreasing returns to scale of the broad reproducible capital? Has there been a long run relation between  $k_t$  and  $h_t$  and technological

*progress? How have  $k_t$  and  $h_t$  been interacting with each other in accordance of technological progress?*

In order to identify the system and test statistically the significance of the parameters, the parameters of the variables need to be restricted so that the two cointegration relations can be distinguished from each other.<sup>14</sup> Several combinations of restrictions on  $\bar{\beta}$ -coefficients were tested to identify the system. Statistical tests and graphical inspections to analyse the mean reverting properties of the cointegration relations with respect to different restrictions were conducted. The analysis of the finally chosen identified cointegration relations and the identified final long run structure will be presented next.

Starting with the first hypothesis, the stationarity of the long run relation between  $y_t$ ,  $k_t$  and  $h_t$  was tested by imposing a homogenous restriction so that the parameters of  $k_t$  and  $h_t$  would sum up to unity, when the relation is normalised on  $y_t$ . The level shift and the time trend were restricted to zero in this relation. The likelihood ratio (LR) test could obviously not reject this hypothesis with a p-value of 0.395, indicating that the likelihood of the estimated model did in practice not change with the imposed restriction. Therefore, *we will have to accept the hypothesis of the other long run relation being the production function of the type in equation 3.4 with constant returns to scale when intangible human capital by formal education is assessed in the National Accounts frame. This implies for non-decreasing returns to scale on broad reproducible capital, human and physical capital.*

After this, the hypothesis of the capitals and technological progress together with the imposed level shift in 1944, was imposed as a second cointegration relation in the model. With exactly these two cointegration relations, the model could not be rejected with a p-value of 0.35.

According to the estimated model in Table 4.2.7, *the long run development of each of the variables has been endogenous*: In the first – constant returns to scale production function – relation labour productivity,  $y_t$ , has been induced by a weighted sum of  $k_t$  and  $h_t$ . *Physical capital in proportion to labour input has a weight of 0.53 and human capital of 0.47, referring to almost identical long run average contributions to the labour productivity development.* Labour productivity shows to have been adjusting to the shocks of the other variables statistically significantly with a t-ratio of -3.1 (here the normal t-ratio value |1.98| with 5% risk level should be exceeded for statistical significance). At the same time,  $k_t$  and  $h_t$  have not been adjusting in the same relation to the shocks of labour productivity growth and to shocks of the other type of capital in the labour input. Thus, it can be obviously referred to as the production function relation.

---

<sup>14</sup> See Juselius, K. (2006) *The Cointegrated VAR model: Econometric Methodology and Macroeconomic Applications*.

**Table 4.2.7 The final identified long run structure,**

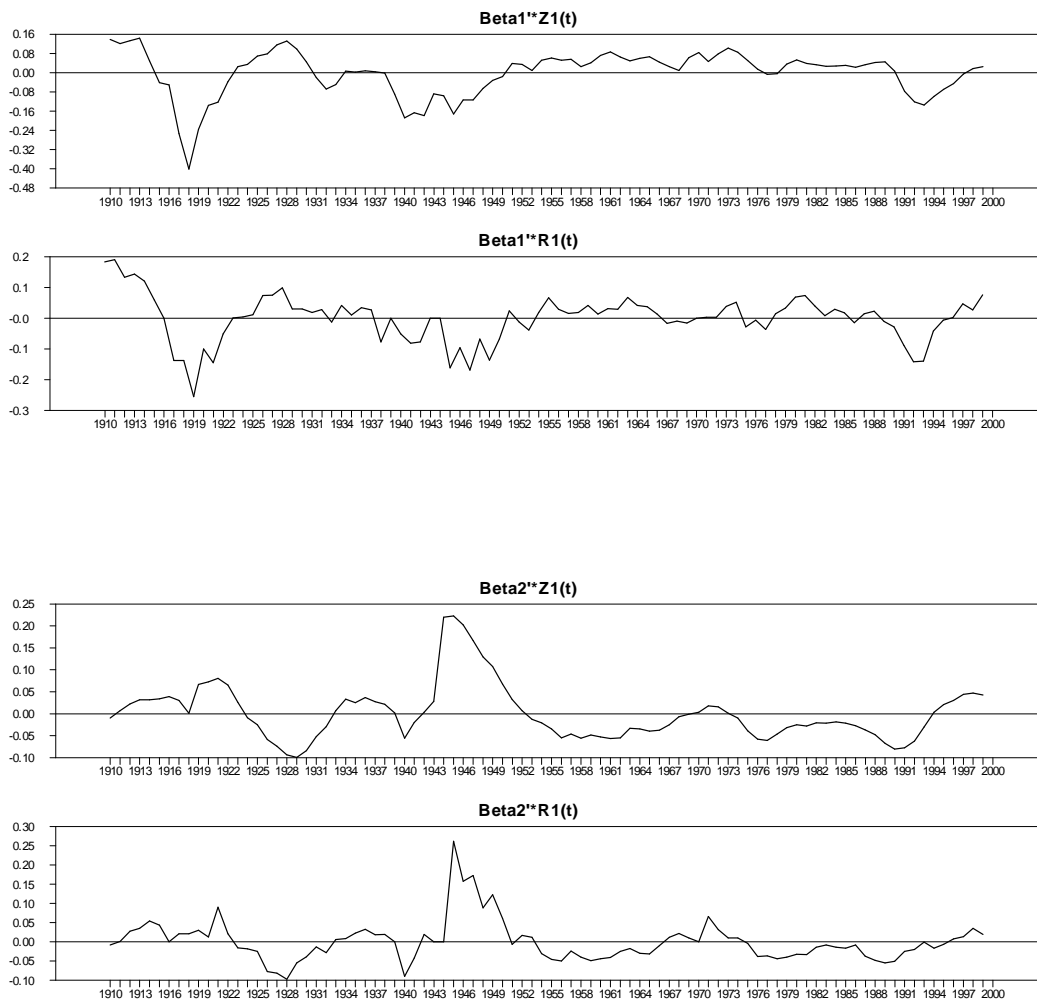
LR test of restricted model  $\chi^2(2)$ , p - value [0.35],

t-ratios in brackets [ ]

	$y_t$	$k_t$	$h_t$	$D_{s1944}$	$t$
$\beta'_1$	<b>1</b>	<b>-0.53</b> [-5.47]	<b>-0.47</b> [-4.93]	0	0
$\beta'_2$	0	<b>-0.74</b> [13.3]	<b>1</b>	<b>0.176</b> [ 5.81]	<b>-0.0155</b> [ 8.53]
	<b><math>\alpha_1</math></b>	<b><math>\alpha_2</math></b>			
$y_t$	<b>-0.11</b> [-3.10]	-0.08 [-1.68]			
$k_t$	0.05 [ 1.34]	<b>-0.12</b> [-2.53]			
$h_t$	0.01 [0.30]	<b>-0.25</b> [-4.94]			

Instead, in the second relation  $k_t$  and  $h_t$  have been adjusting to each other, which is making the development of the variables in this empirical model endogenous. As one type of capital has grown the other type of capital must have grown, in accordance with the time trend, to keep the relation  $h_t = 0.74k_t + gt$  or  $h_t - 0.74k_t - gt = 0$  stable. Since, in non-logarithmic form  $(H_t / L_t) / (0.74K_t / L_t) = H_t / 0.74K_t$ , the relation describes the evolution of capitals themselves. As both are adjusting in the relation, physical and human capital have grown with respect to each other along with technological change. Human capital has grown faster, and in order for the relation to stay stable, the technological progress, the time trend is needed for enhancing the effects of physical capital. This is realistic since the technology embodied in the equipment and machinery has no doubt advanced in  $k_t$ , and with the same costs the capital goods produced later are of higher quality. The technological progress and the growth of  $K_t$  has created demand for human capital. At the same time more human capital has induced a possibility to higher technology capital goods and promoted interest in investing in production equipments in Finland. Consequently, in the empirical model above, *a positive increase in human capital has stimulated physical capital to grow and vice versa*. After 1944 the whole relation between the capitals has shifted by 18% in favour for human capital,  $t \geq 1944$ :  $[h_t + 0.18 = 0.74k_t + gt]$ . This is probably due to the destruction of physical

capital in WWII and to schooling expansion that has provided more possibilities to use higher and higher technology equipment and machinery adopted from the evolving world technology frontier. The latter has created an incentive to invest in continuously advancing physical capital. In addition, much more physical capital was destroyed in 1939–1944 than human capital, and the shift in that time also reflects an exogenous shock in the relative relation of the capitals. The rate of the average growth of technological progress, interpreted here to be embodied in the physical equipment, is estimated to be less than 1.6 per cent per year in the model.



**Figures 4.2.3 The cointegration relations with  $r = 2$ , CRS-production function in intensive form with human and physical capital as a first relation (Beta 1, upper graphs), capitals and time trend in the second relation (Beta 2, lower graphs)**

The long run equilibrium correction (or cointegration) relations  $\mathbf{B}_1 : [y_t - 0.53k_t - 0.47h_t = 0]$  and  $\mathbf{B}_2 : [h_t - 0.74k_t - 0.18D_{s1944} - 0.016t = 0]$  are sketched in Figure 4.2.3. The upper graph ( $\mathbf{B}'_1\mathbf{Z}_1(t)$ )

in both figures refers to the  $\beta'X_t$  relation in the model in equation 4.1.2 with the short run effects of the lagged differences  $\Gamma\Delta X_{t-1}$  included in the model. The lower graphs ( $B_1'R_1(t)$ ) refer to a model from which the short run effects have been concentrated out. The  $R_1$  model is of more importance in identifying the long run structure. However, the mean reverting properties should be present in both graphs, and the evolution of the relation should not differ significantly between the models. Both relations exhibit a mean reverting stationary behaviour, with similar evolution in both,  $Z_1$  and  $R_1$ , models.

**Table 4.2.8 The MA representation corresponding to the identified long run structure, t-ratios in brackets [ ]**

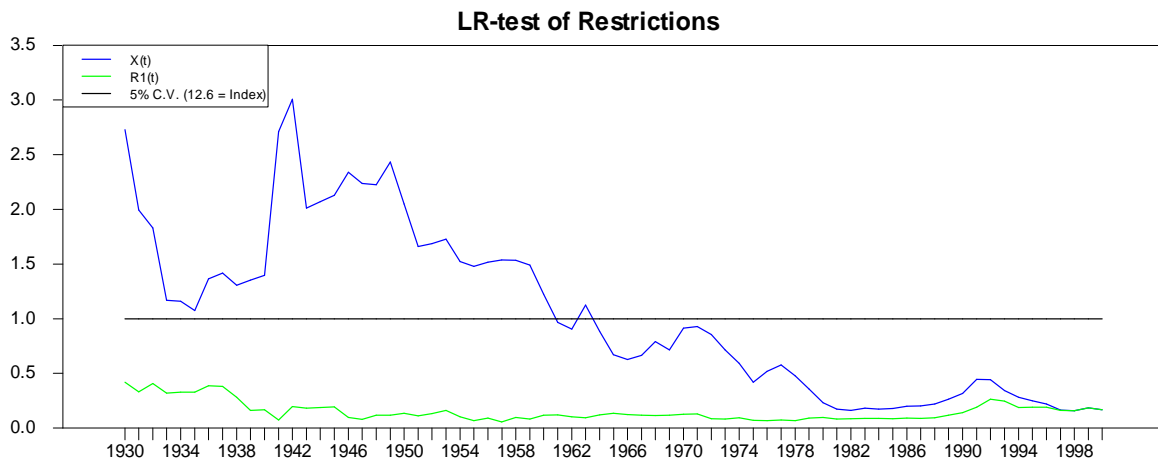
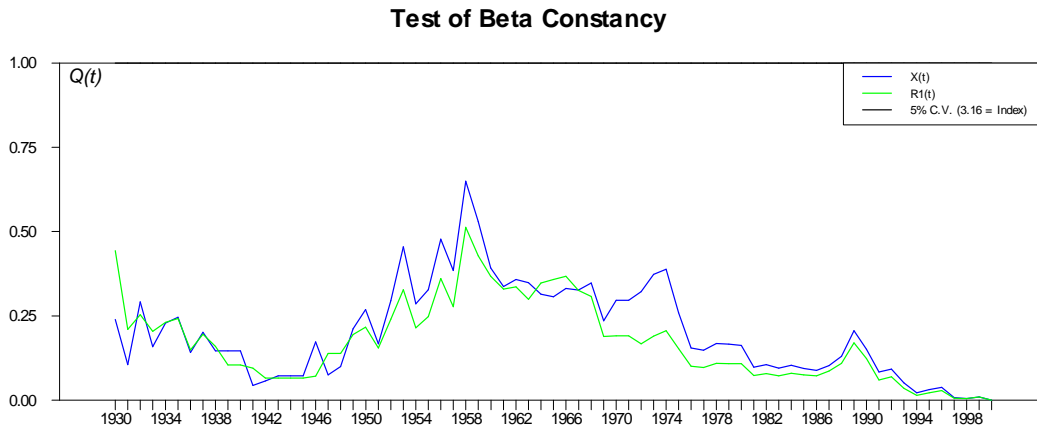
	$\varepsilon_{yt}$	$\varepsilon_{kt}$	$\varepsilon_{ht}$	
$\sigma_{\varepsilon i}$	0.0237	0.0270	0.0200	
Common trend (CT) weights				
$\alpha'_{\perp}$	0.38 [1.55]	<b>1</b>	<b>-0.58</b> [-5.98]	
Loadings to CT				
$\tilde{\beta}'_{\perp}$	<b>1.75</b> [2.54]	<b>2.00</b> [2.54]	<b>1.48</b> [2.54]	
The Long Run Impact matrix C				
	$\varepsilon_{yt}$	$\varepsilon_{kt}$	$\varepsilon_{ht}$	$t \cdot \gamma_{0,i}$
$y_t$	<b>0.66</b> [3.19]	<b>1.75</b> [2.53]	<b>-1.01</b> [-2.08]	0.0312
$k_t$	<b>0.75</b> [3.19]	<b>2.00</b> [2.53]	<b>-1.15</b> [-2.08]	0.0272
$h_t$	<b>0.55</b> [3.19]	<b>1.48</b> [2.53]	<b>-1.47</b> [-2.08]	0.0356

The Moving Average representation of the identified model is given in Table 4.2.8. The main message of the MA form with regard to the common trend is the same as in the case of a unit vector in alpha on  $y_t$ : the common stochastic trend is defined by the combination of the cumulated shocks

to  $k_t$  and  $h_t$  by  $\left[ \sum_{i=1}^t \varepsilon_{ki} - 0.58 \sum_{i=1}^t \varepsilon_{hi} \right]$ . The proportion in non-logarithmic form,  $K/L / H/L = K/H$ , and

the innovations in  $K$  and  $H$  seem to be the source for the pushing force in the model as the alpha

orthogonal coefficient for  $y_t$  is not statistically significant. All of the variables have a positive loading with respect to the common trend,  $\tilde{\beta}_\perp = [1.75 \ 2.0 \ 1.48]'$ .



**Figures 4.2.4 Recursively estimated tests for the constancy of the model long run parameters and identifying restrictions**

Finally the constancy of the model  $\beta$ -parameters and the constancy of the restrictions imposed above on  $\beta$  in the long run was examined with recursive tests. The parameters of the long run relations have been stable over the whole estimation period and certainly below the critical value of rejecting the constancy (scaled to 1 in the figure), as can be seen in the upper graph in Figures 4.2.4. Similarly, it becomes obvious from the lower graph that the identifying restrictions on the cointegration relations would have been accepted in all of the recursively estimated sample periods

along the time frame. Attention should here be paid to the lower line, referring to the stability of the long run  $\mathbf{R}_1$  model, from which the short run effects have been concentrated out.

*Analysis with one cointegration relation between  $y_t$ ,  $k_t$  and  $h_t$*

The analysis of only one cointegration relation will be reviewed here first because the analysis of the cointegration rank gave some signs of  $r = 1$  being on the borderline of acceptance. Secondly, we can use this analysis to ensure without any doubts that the first relation above can be named as a production function relation.

With one cointegration relation and the other variables normalised on GDP per number of hours worked ( $y_t$ ) the vector equilibrium correction model gets the form:

$$\begin{cases} \Delta y_t = \bar{\alpha}_1(y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) & + \Gamma_{11} \Delta y_{t-1} + \Gamma_{12} \Delta k_{t-1} + \Gamma_{13} \Delta h_{t-1}) + \varepsilon_{1t} \\ \Delta k_t = \bar{\alpha}_2(y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) & + \Gamma_{21} \Delta y_{t-1} + \Gamma_{22} \Delta k_{t-1} + \Gamma_{23} \Delta h_{t-1}) + \varepsilon_{2t} \\ \Delta h_t = \bar{\alpha}_3(y_{t-1} + \bar{\beta}_2 k_{t-1} + \bar{\beta}_3 h_{t-1}) & + \Gamma_{31} \Delta y_{t-1} + \Gamma_{32} \Delta k_{t-1} + \Gamma_{33} \Delta h_{t-1}) + \varepsilon_{3t} \end{cases}$$

In the analysis above it was shown that the level shift 1944 was not present in the first cointegration relation. Therefore, a permanent impulse dummy was set to 1944, to account for the level shift in the Moving Average form in the evolution of the levels of  $\mathbf{x}_t$ . This produces the same effect on the evolution of the level series as a level shift restricted to cointegration relation plus a difference of it unrestricted included in the model (see Juselius 2006, pp. 104–109).

After the estimation of the baseline VAR, the parameters of  $k_t$  and  $h_t$  were restricted to sum up to the coefficient of  $y_t$  and for testing that the production function would be exactly of the type  $y = k^\alpha h^{1-\alpha}$ , the time trend was restricted to zero. The estimation results (below with t-values in parenthesis) of this model with restricting the parameters of k and h summing up to one in the cointegration relation argue that GDP per hours worked could be explained solely by physical and human capital per hours worked. The LR test on the restriction was accepted with a p-value of 0.24. Along with the homogenous restriction on the parameters, the constant returns to scale hypothesis with respect to broad capital could not be rejected. Together with this, it was tested that the time

trend (the Solow residual) could be left out of the production relation when human capital by formal education is included.

Table 4.2.9 shows that all the variables are statistically obviously significant in this long-run relation. Physical capital in the labour input gets a parameter value of -0.54 and human capital in the labour input only a slightly smaller value of -0.46, implying that with a simple production function approach they would explain each half of the long run GDP per hours worked in Finland in 1910–2000. More importantly, technical change or the Solow residual could be excluded from the model, with LR-test the p-value of 0.24 in the case of one cointegration relation. Here, with only one long run stationary relation in the ECM model, all of the variables adjust statistically significantly to the long run equilibrium between the variables: labour productivity adjusts with a pace of 6.7% each year to get back to the equilibrium of the disequilibrium caused by shocks, physical capital per hours worked by 11% and human capital per hours worked by 14% a year. What we could not hypothesise in this model, is that the capitals may have adjusted to the development on each other and not on  $y_t$ .

**Table 4.2.9 The identified long run structure with  $r = 1$**

LR test of restricted model  $\chi^2(2)$ , p - value [0.24], p-value\* [0.46]

t-ratios in brackets [ ]

	$y_t$	$k_t$	$h_t$	$t$
$\beta'_1$	<b>1</b>	<b>-0.538</b> [-5.86]	<b>-0.462</b> [-5.04]	0
<b><math>\alpha_1</math></b>				
$y_t$	<b>-0.07</b> [-2.61]			
$k_t$		<b>0.11</b> [ 4.15]		
$h_t$			<b>0.14</b> [4.56]	

The cointegration relation is sketched in Figure 4.2.5 (the lower of them is estimated by reducing the short run effects from the relation). The relation is obviously stationary and the biggest shocks to the system caused by wars (1917-1919 and 1939-1945) and depressions (early 1930s and early 1990s) are the causes for momentary disequilibria between the variables.



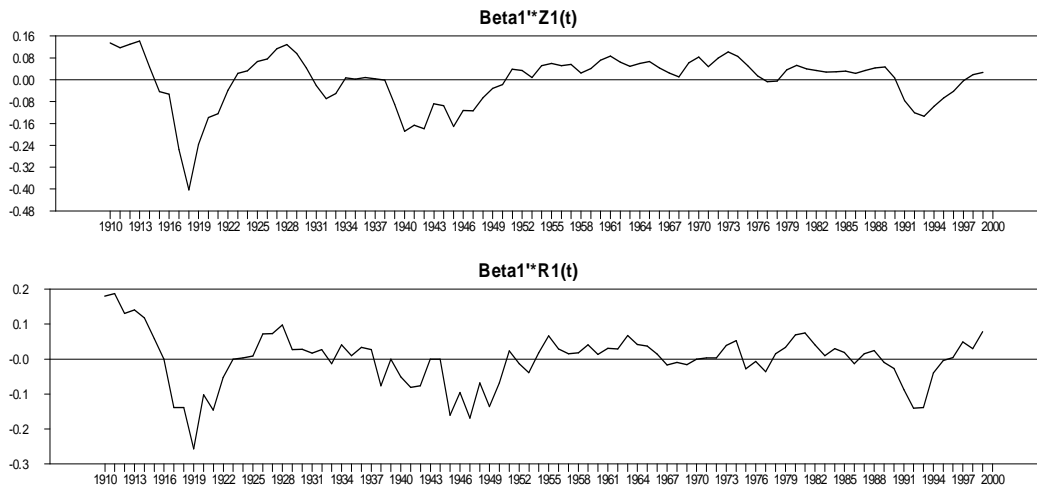


Figure 4.2.5 The cointegration relation with  $r = 1$ , only  $y$ ,  $k$  and  $h$  in the model

Table 4.2.10 The MA representation corresponding to the identified long run structure with  $r = 1$

t-ratios in brackets [ ]

	$\varepsilon_{yt}$	$\varepsilon_{kt}$	$\varepsilon_{ht}$
$\sigma_{\varepsilon i}$	0.0211	0.0297	0.0222
Common trends (CT) weights			
$\alpha'_{\perp,1}$	<b>1.63</b> [2.09]	<b>1</b>	0.00 [ 0.00]
$\alpha'_{\perp,2}$	<b>2.15</b> [2.26]	0.00 [ 0.00]	<b>1</b>
Loadings to CT:s			
$\tilde{\beta}'_{\perp,1}$	0.91 [1.78]	<b>2.50</b> [3.49]	-0.95 [1.77]
$\tilde{\beta}'_{\perp,2}$	-0.30 [-0.86]	<b>-1.60</b> [-3.22]	<b>1.20</b> [3.24]

The Long Run Impact matrix C

	$\varepsilon_{yt}$	$\varepsilon_{kt}$	$\varepsilon_{ht}$	$t \cdot \gamma_{0,i}$
$y_t$	<b>0.83</b> [5.34]	0.91 [1.78]	-0.30 [-0.85]	0.031
$k_t$	<b>0.65</b> [2.98]	<b>2.50</b> [3.49]	<b>-1.60</b> [-3.22]	0.027
$h_t$	<b>1.03</b> [6.31]	-0.95 [-1.77]	<b>1.20</b> [3.24]	0.035

The MA form, given in Table 4.2.10, shows that the two stochastic trends, with  $r = 1$ , are defined by the innovations on  $y_t$  and  $k_t$  and on  $y_t$  and  $h_t$ , stating that the growth in labour productivity growth itself has pushed the system to grow together with  $k_t$  and  $h_t$ .

*What if human capital per hours worked did not exist as an empirical variable in the analysis? Could it be excluded from the model? How would the rate of the technical change (or Solow residual) change in the long run?* These questions can be answered by imposing a hypothesis  $\bar{\alpha}_1 = \bar{\beta}_3 = 0$ . The results are shown in Table 4.2.11.

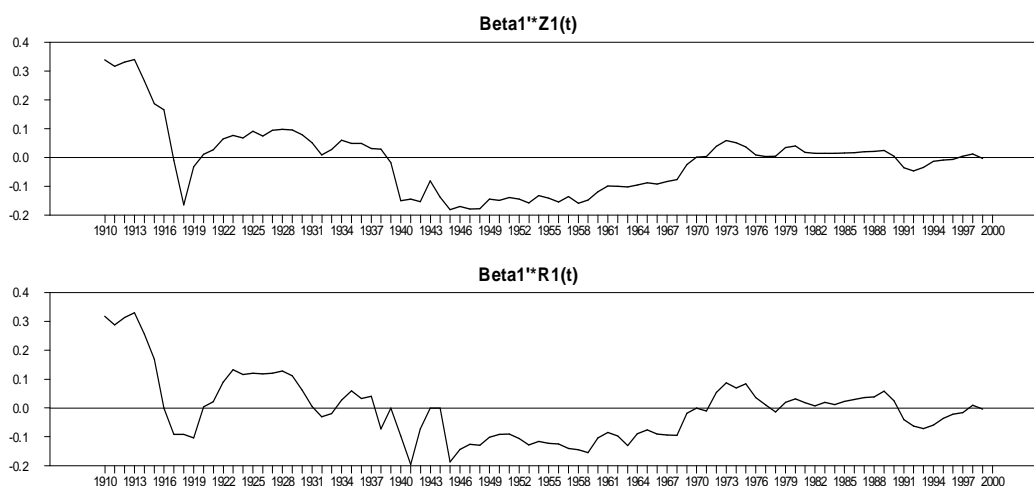
The LR-test, comparing the Log-Likelihoods of the models with and without  $h_t$ , does not support excluding human capital from the model. The growth of the estimated time trend (or the “Solow residual”) interpreted as technological progress would be 2.4% a year when human capital per hours worked was excluded from the model. If still this model was fitted to the data,  $k$  would not be adjusting statistically significantly (t-value -0.64) to the long-run equilibrium, which would imply for it being weakly exogenous in the model and therefore the determinant of long-run labour productivity growth together with the time trend. *Without having human capital assessed in the NA in the model, it would be possible to end up with a traditional neo-classical Solow-Swan explanation for labour productivity growth, with a large Solow residual.* The cointegration graph below is perhaps not exhibiting strong mean reverting properties throughout the sample and could not give strong support to a stationary behaviour.

**Table 4.2.11 The long run structure without h in the model,  $r = 1$**

LR test of restricted model  $\chi^2(2)$ , p - value [0.0027], p-value\* [0.0341]

t-ratios in brackets [ ]

	$y_t$	$k_t$	$h_t$	$t$
$\beta'_1$	<b>1</b>	<b>-0.263</b> [-5.86]	0	<b>-0.0243</b> [-3.89]
<hr/>				
	$\alpha_1$			
$y_t$	<b>-0.07</b> [-4.23]			
$k_t$	-0.12 [-0.65]			
$h_t$	0 [0.00]			



**Figure 4.2.6 The cointegration relation with  $r = 1$  and only  $k$  and  $t$  in the model**

To conclude the analysis with one cointegration relation between the variables, *GDP per hours worked could be explained solely by physical and human capital per hours worked*. This relation replicates a constant returns to scale production function in intensive form  $Y / L = A(K / L)^\alpha (H / L)^{1-\alpha}$  referring to endogenous growth models of the type  $Y = AK$ , where  $K$  is defined as broad reproducible capital including human capital. All of the variables are endogenous in the model and adjust to the development of the other variables. This implies that the empirical evolution of the variables has been truly endogenous in the model. The time trend or the Solow residual could be left out of the production relation when human capital by formal education, assessed in the National Accounts, is included in the model. This is signalling for the possibility of an endogenous technological progress induced by human and physical capital accumulation.

Without a human capital variable, constructed as in this study, the conclusion would be in favour of exogenous neo-classical growth. Therefore the building of the human capital variable in the National Accounts frame, as GDP and physical capital, can change our insight of the growth process and may give support to endogenous growth theories with constant returns to scale on reproducible capital.

## **5. Conclusions**

One of the major discussions regarding growth theories is whether the modern economic growth could be best modelled by neo-classical or endogenous growth models. One of the questions inside this discussion is the role of human and physical capital and whether diminishing returns or constant returns to scale with respect to reproducible capital would be prevailing in the production of GDP (see e.g. Jones 2005).

In empirical studies, with the conventional measures for human capital, such as average years of schooling in the working age population, the answer has favoured the latter (see e.g. Easterly and Levine 2001, Hall and Jones 1998). It has also shown to be difficult to achieve indisputable evidence on the positive impact of schooling on GDP growth.

This paper explored whether the connection of human capital and GDP would be more evident and whether the empirical feedback to theory would be different when human capital is assessed in the National Accounts framework (NA). The main objective to be studied was whether human capital assessed in the National Accounts could change the view on whether physical and human capital accumulation would be the main factors of growth or whether it has been the exogenous technical change (or multifactor productivity).

As the first part average years of schooling in the working age population and human capital estimates from NA were illustrated and compared in the U.S. and in Finland along the 20<sup>th</sup> century. The graphical inspection showed that the National Accounts estimates for human capital by education of Kendrick and Jorgenson and Fraumeni for the U.S. and of this paper for Finland do grow exponentially and follow closely the evolution of GDP, while the conventional measures grow linearly. Therefore, a variable for human capital by schooling in the National Accounts implies to more weight on human capital in explaining GDP, and gives the feedback that human capital by schooling can be entered without an exponential structure into a growth theory model.

Jorgenson and Fraumeni's (1992b) conducted a growth accounting analysis resulting in that together capital input and labour input accounted now for almost all of the growth of their adjusted output in accordance with their lifetime labour income system. They concluded that the accumulation of human and non-human capital accounts for the predominant share of economic growth, which gives implications to endogenous growth. Instead, Mankiw, Romer and Weil (1992) reported strong support for the neo-classical Solow-Swan model augmented by human capital. In their analysis they were employing secondary school enrolment as a proxy with respect to the growth of the standard GDP.

Both the systems of Kendrick and Jorgenson and Fraumeni have included a substantial amount of imputed non-transaction based flows both in investments in human capital and in their GDP. It could be argued that this has influenced the results. At the same time, the conventional measures have been used to explain the standard GDP. The question whether more weight on the reproducible capital can be given in explaining the standard GDP based on market activities remains open.

For reaching a fair comparison of National Accounts estimate with the conventional measures for the discussion on the feedback to theory, intangible human capital by schooling was assessed in the modified National Accounts system in which GDP does not have to change. An exhaustive Vector Equilibrium Correction analysis was conducted on the evolution of the long run labour productivity, physical capital per hours worked and human capital per hours worked for Finland in 1910–2000.

The results showed that labour productivity was adjusting to the innovations in physical and human capital in the labour input to maintain the long run steady state equilibrium between the variables. Instead, human and physical capital in proportion to hours worked were not adjusting to the shocks in labour productivity. Therefore, the direction of Granger causality has been from human and physical capital in the labour input to labour productivity.

The pushing force for the labour productivity growth was the proportion of the innovations of physical capital to human capital. Physical capital connected to the level of technology and human capital have been growing in response to each other in the second long run equilibrium relation. The result suggest that technology can be seen embodied in physical capital giving support to vintage capital models.

According to the results the Solow residual or multifactor productivity can be excluded from the production function in intensive form when human capital by schooling is assessed in the National Accounts frame. Constant returns to scale are prevailing in the long run production with respect to broad reproducible capital, including human and physical capital. Therefore, the results give support in the simplest form to the  $y_t = Ak_t$  type of endogenous growth models with  $k_t$  referring to broad reproducible capital. The long run average elasticity of GDP is suggested to be approximately 0.5 with respect to both physical and labour input adjusted by human capital by schooling.

In the essence, the feedback to theory of assessing human capital by schooling in the National accounts is that human and physical capital are the main determinants of the long run standard GDP growth. Human capital is one of the most important factors for long run growth.

## ***Appendix I: The modified system of production of National Accounts with intangible human capital by schooling***

Human capital is excluded from the asset boundary of the international Standard of National Accounts (SNA) in the current empirically applied 1993 version (SNA1993). The revised SNA2008 to be implemented in some years excludes it from the core accounts as well, but proposes it as an additional voluntary satellite account outside the core system.

In the System of National Accounts (SNA) non-financial assets are either produced assets or non-produced assets. Following Aulin-Ahmavaara's comprehensive work, in order to treat human capital as an asset it has to belong to either of these categories (Aulin-Ahmavaara 2002). At the same time, learning new skills and knowledge requires inputs. Thus, in Aulin-Ahmavaara's words, "If human capital is wanted to be seen as an asset, it has to be produced." (Aulin-Ahmavaara, 2002, p. 3) But production of human capital falls outside the production boundary of the SNA. Therefore, including human capital inside of National Accounts necessitates moving its production inside the production system where output and other produced assets (e.g. physical capital) are produced.

It is worth clarifying that excluding human capital in the SNA is not an accident. It is a logical consequence of the definition of production in the system. The issue of leaving human capital out is dealt in detail in the current applied version SNA93 and in the in the revised SNA2008 to be implemented in empirical work in some years. The SNA2008 (par 3.48) states "*Human capital is not treated by the SNA as an asset. It is difficult to envisage "ownership rights" in connection with people, and even if this were sidestepped, the question of valuation is not very tractable.*" However, economists have often requested to include human capital inside the system. National Accounts are also constantly criticised of not incorporating the most important factors for modern economic growth.

The valuation problem mentioned above may have something to do with the most well-known proposed systems including human capital, also suggesting flow variables without monetary transactions to be accounted either as investments or as services of human capital (see Kendrick, 1976, Jorgenson and Fraumeni, 1989, 1992a, 1992b). For instance, treating foregone earnings of students as investments in the core system would make GDP to include this same amount, for which no transactions occurred. It should be added to the balance sheet of households as well. While foregone earnings can be justified from the input-output view, it may be unjustified from the point of view institutional sector accounts, which aim at giving the financial position (the net lending/borrowing to/from other institutional sectors) of the institutional sectors and how they

have financed their production and investments. If foregone earnings were added in the output (/or as an input) of the household sector (/of the sector where educational institutions belong in each country), this would have to change the disposable income of the household sector (/the respective sector) even without any transactions occurred. At the same time, thinking of GDP per capita as a measure for material living standard, in the long run analysis the inclusion of foregone earnings in the core accounts would not perhaps make sense, since no one has eaten with foregone earnings.

Without arguing that foregone earnings of students would not be important in education decisions, this paper introduces a system concentrating on paid monetary flows as investments in formal education for two particular reasons: First, the aim here has been to explore whether schooling has had a role in the standard GDP growth. All the empirical studies with the proxy variables for human capital have sought for the connection with the standard GDP. Second, if foregone earnings are added to both investments and to GDP, it results in an approx. 20-30% increase in the level of GDP and hence would make a much bigger part of investments (Kendrick, 1976). When human capital stock, accumulated by such investments would be studied with GDP, the long run cointegration relation would be empirically much easier to achieve. Therefore, at least to start with, it would be good to make sure that the connection of education can strictly be found with the standard GDP. Inversely, after finding a long run cointegration relation as in this paper, the relation is probably to be found as well with adding the same figure to investments (accumulated to the corresponding stock) and to GDP. Next, the modifications to the system in this paper will be introduced in comparison with SNA.

The system of production in the SNA and the revisions to the production system of it in this paper are shown below. The description of the system of production of the SNA (the basic equations 1–5a) and 6) is based on the representation of Aulin-Ahmavaara (2002) <sup>15</sup>. The system is simplified in the sense that taxes and subsidies are ignored and a simple geometric rate of depreciation is assumed. The revisions to the production system are shown with the bolded variables and with equation 5b).

In the original system of production of SNA the first equation (1) defines the supply and demand in the economy in a time unit (in a year or in a quarter): Output (O) is the sales revenues producers get when selling the products they have produced. Imports (M) include the value of goods and services imported to the country. The use or demand of these products is on the right hand side of the first equation (1): Part of these products has been re-used in the production of

---

<sup>15</sup> Aulin-Ahmavaara was originally describing the system of production of the SNA93, however, to my best knowledge the reasoning here does not change in the SNA2008. The biggest change here in the SNA2008 is that research and development expenses will be subtracted from intermediate uses and treated as investments.

other producers as instant intermediate inputs ( $U$ ) or as investments ( $I_K$ , typically e.g. machines and buildings that are used for more than one year in the production of the buyer-producer). A good part of them is used as private final consumption ( $C$ ) and government final consumption ( $G$ ) or exported ( $E$ ). All of the variables are expressed in monetary terms, in current prices, or in fixed prices, where the price changes have been deflated.

The second equation (2) shows how GDP (or value added) can be calculated through output minus intermediate inputs or through net-demand, i.e.  $C + I + G + (E-M)$ . The third equation (3) emphasises that the value of output can also be calculated through incomes generated in the production process, namely through intermediate inputs plus the compensations for labour ( $W$ ) and capital ( $R$ , operating surplus of the producers). As a consequence, in the fourth equation (4) the same GDP can be derived by the incomes generated in the production process as the sum of compensation for labour and for capital. The original fifth equation (5a) describes the accumulation of physical capital: investments in physical capital increase the accumulated stock and the depreciation decreases the value of the stock. The labour input (equation 6) is treated as an exogenous variable as households decide whether they are available in the labour market and how much they are willing to work.

In order to include intangible human capital by schooling in produced assets the production system is revised in this paper (the bold variables and equation 5b). In equation 1, the education expenditures are deducted from final consumption (in the Finnish case from general government) expenditures and reclassified as intermediate inputs (used in the learning process of students). The new skills the students have embodied in a year are treated as produced human capital by schooling ( $O_H$ ) and then as investments in human capital ( $I_H$ ). The former is added to the nation's output and the latter is added together with physical capital investments ( $I_K$ ). The produced human capital by schooling and the investments in human capital by schooling (to be used finally for more than a year in the production process) are valued, at fixed prices, through the actual monetary flows paid in the economy, i.e. their value is equal to education expenditures. This means that the accounts are balanced and GDP does not change in the new equations 2 and 4 as the value of produced human capital ( $O_H$ ) and investments in human capital ( $I_H$ ) equals the value of education expenditure that is added to intermediate inputs and subtracted from final consumption (in the Finnish case from general government) expenditure. Time spent at school (and outside while doing homework) is seen here to be used in the production of human capital, however, since no one has paid the students, no monetary value is given to it. The depreciation of human capital by schooling of the labour force is assumed to be included in their wages and salaries as part of the compensation for the skills accumulated and used in the labour market. It is worth noticing that in equations 3 and



4 the compensation for labour includes the compensation for skills and knowledge by education used in the production. In a modern economy the compensation for skills and knowledge should count for a bigger part than the compensation for physical labour work.

In the new equation 5b, intangible human capital by schooling, H, is accumulated with the perpetual inventory method. When the long graduation times are taken into account, as in the calculations for this paper, the accumulation is done when a person finally enters the labour market, by all investments  $I_{Ht}$  (education expenditures) up to that time. The accumulation is decreased by the rate of depreciation of human capital by schooling with the assumption of geometric age-efficiency profiles. The physical capital of the educational institutions is here left to physical capital stock. The broad capital accumulation in the revised system will include physical capital and intangible human capital by schooling.

$$1. [O+O_H] + M = [U + \textit{education expenditure}] + C + [I_K + I_H] + [G - \textit{education expenditure}] + E$$

$$2. GDP = [O+O_H] - [U + \textit{education expenditure}] = C + [I_K + I_H] + [G - \textit{education expenditure}] + E - M$$

$$3. [O + O_H] = [U + \textit{education expenditure}] + W + R$$

$$4. GDP = [O + I_H] - [U + \textit{education expenditure}] = W + R$$

$$5a) \frac{dK}{dt} = I_{Kt} - \delta_K K_t$$

$$5b) \frac{dH}{dt} = I_{Ht} - \delta_H H_t$$

$$6. L = \bar{L}$$

Where

O = gross output, U = intermediate uses / intermediate inputs, C = private final consumption,

G = General government final consumption expenditure

$I_K$  = gross physical capital formation,  $I_H$  = gross human capital by schooling formation

E = exports, M = imports

W = labour compensation, R = operating surplus (or mixed income)

K = physical capital stock,  $\delta_K$  = rate of depreciation of physical capital

H = human capital stock by schooling,  $\delta_H$  = rate of depreciation of human capital by schooling

L = labour input

Geometric depreciation rates were used because they typically combine the age-price/age-efficiency and the retirement profile for a cohort of assets. As shown in the OECD's manual for measuring capital (OECD, 2009), various age-efficiency profiles for individual assets, when combined with retirement profiles for entire cohorts, generate profiles that are more or less convex to the origin so that the geometric model can be used as an approximation to a combined age-efficiency/retirement pattern. Furthermore, the productive capital stock and the net capital stock coincide in the case of geometric depreciation rates because age-price and age-efficiency profiles coincide. The depreciation rates for each type of education were approximated by using the declining balance method. Hulten and Wykoff (1996) have made a suggestion for converting an average service life of a cohort,  $T^A$ , into a depreciation rate, with formula  $\delta = R / T^A$ , where R is the declining-balance rate. Under the double declining balance formula, R is set to equal 2, but generally it would be best to turn to empirical estimation results for the shape of the geometric depreciation pattern. Recently, Baldwin et al. (2007) have reported econometric estimates of declining balance rates for traditional capital in the range between 2 and 3. In this paper, the long run retirement average was set to be at age 65 and the graduation ages depending on the type and level of education, and on the year in the history of the education system. The accumulation was done separately for people entering the population at the working ages with primary, lower secondary, upper secondary, professional/vocational and university education.

The average age for entering the labour markets with basic education (for the years after 1975 comprehensive and before it primary education) was set to 16, with upper secondary education to 19 and with university education to 28, yielding the average service lives for the mentioned types of education 49, 46 and 37 years, respectively. The rounded depreciation rates were set by calibrating with the declining balance formula: 5% depreciation for basic education, 5.5% for upper secondary and 7.5% for university education,<sup>16</sup> giving the respective declining balance rates 2.45 per cent, 2.73 per cent and 2.8 per cent, falling in the range between 2 and 3. The decline rate in basic and upper secondary is assumed to be lower than in the university (and professional) education because of the basic knowledge and skills giving nature of this type of education. For university and professional education the declining balance rate is assumed to be somewhat faster (and similar in respect to each other) because they include specialisation directed to the labour market at the time and the evolution of labour markets has been fast in connection with fast transformation of the society and rapid technological change in Finland.

---

<sup>16</sup> With these geometric depreciation rates the basic and upper secondary education investment has lost 92% of its value in its average service life and university education investments 94% in their average service lives.

For the professional/vocational education different paths for entering the schools and completing the education were considered: For the years up to mid-1950s the typical way to professional education was considered through lower secondary or primary education at the average age of 16. Most of the lines on this type of education were lasting for two years, however, 20% of the students were considered to have continued a third year (e.g. in upper engineering schools). After 1955 until 1990 an additional 4<sup>th</sup> year was set to be available to 25% of the third class students. After the polytechnic reform in the early 1990s the proportion of the 4-year professional education was increased gradually to 70% of students. Together with similar declining balance rate to university education and in accordance with the service lives of 47, 46 and 45 years, the depreciation rates were set to 6%, 6.1% and 6.2%.

*The entire stock reflects people in the working ages with different education along time, taking into account the volume of the resources put to education each cohort with different educational path have used.* To achieve this, the number of students in each type of education was collected, and the number of students in each year (first, second, third etc.) of each type of education was estimated.

Intangible human capital by schooling (or the accumulated knowledge and skills) is used finally in the process of production and income generation in the labour market in the working ages incorporated in the hours worked. The wages and salaries that are accounted in GDP (see the above equation 4) are paid for people participating in the production. In a modern economy an increasing part of their income is compensation for their skills and knowledge used in production. This does not imply in any way that the accumulated human capital by schooling – especially of people working in research and development – could not affect technical change separately. Rather, the reasoning here implies the existence of both channels on how human capital enhances economic growth.

## Data Sources

### Data for U.S. in Figure 3.1 and 3.2:

Carter, Susan B., Sigmund G. Scott, Michael R. Haines, Alan L. Olmstead, Richard Sutch, Gavin Wright (Eds.) (2006). *The Historical Statistics of the United States*. New York: Cambridge University Press, 2006; real GDP for the U.S., school enrolment rate at the ages of 5–34

Jorgenson, Dale W., Fraumeni, Barbara M. (1992b). *Investment in Education and U.S. Economic Growth*, Scandinavian Journal of Economics, 94. Supplement, pp. 51–70; J-F Output, J-F Labour input

Kendrick, John W. (1976). *The Formation and Stocks of Total Capital*, Columbia University Press for NBER, New York; Kendrick's stock of H by education and training

Godo, Yoshihisa and Yujiro Hayami (1999). Accumulation of Education in Modern Economic Growth; Comparison of Japan with the United States, *ADB Working Paper No. 4*, Tokyo: Asian Development Bank Institute; average years of schooling in the working age population for the U.S.

### Finnish data in Figures 3.3 and 4.2.1

(own calculations, the sources below; Intangible human capital by schooling, average years of schooling in the working age population 16–64, school enrolment rate at the ages 7–26):

Hjerppe, Riitta (1996). *Finland's Historical National Accounts 1860-1994: Calculation Methods and Statistical Tables*. University of Jyväskylä, Department of History, Suomen historian julkaisuja 24. Jyväskylä 1996, 177 pp.

Jalava, Jukka and Matti Pohjola (2007): “The roles of electricity and ICT in growth and productivity: case Finland”, *Pellervo Economic Research Institute Working Papers*, N:o 94 (April 2007); Physical capital

Kivinen, Osmo (1988). *The systematisation of education: basic education and the state school doctrine in Finland in the 19th and 20th centuries*, (in Finnish), Turun yliopisto, Turku

Klemelä, Kirsi (1999). *Ammattikunnista ammatillisiin oppilaitoksiin: ammatillisen koulutuksen muutoutuminen Suomessa 1800-luvun alusta 1990-luvulle*, Turun yliopisto, Turku

Statistics Finland, *Official National Statistics (ONS): National Accounts*, publications: 1948–1965, database 1975–2004, December 2005, Statistics on state's financial accounts in *Statistical yearbooks 1879–2008*, ONS on *Educational Institutions 1971–2009*; *Population Statistics*; Havén Heikki (ed.), Education in Finland, *ONS 1998:1*;

Vattula, Kaarina (Ed.,1983) *Suomen taloushistoria 3. Historiallinen tilasto*. (in Finnish: Finland's economic history 3, Historical Statistics). Tammi, Helsinki 1983

## References

- Aulin-Ahmavaara, Pirkko (2002). *Human Capital as a Produced Asset*. Session Paper Number 1, Session 6A, 27th General Conference of the International Association for Research in Income and Wealth, Stockholm, Sweden August 18-24, 2002.
- Aulin-Ahmavaara, Pirkko (2004). Moving Human Capital Inside the Production Boundary. *The Review of Income and Wealth*, series 50, number 2, 213-228.
- Barro, Robert J. (1991). Economic Growth in a Cross Section of Countries, *Quarterly Journal of Economics*, vol. 106, 1991:2, 407–43.
- Barro, Robert J., Sala-i-Martin, Xavier (1999). *Economic Growth*, First MIT Press edition, The MIT Press, Cambridge, Massachusetts, London, England
- Baldwin, John, Guy Gellatly, Tanguay Marc (2007). “Depreciation Rates for the Productivity Accounts”, *The Canadian Productivity Review*, Statistics Canada Research Paper.
- Bils, Mark and Klenow, Peter (2000). “Does Schooling Cause Growth or the Other Way Around?”, *American Economic Review*, 90 (December): 1160–83
- Bosworth, Barry P., Collins, Susan M. (2003). The Empirics of Growth: An Update. *Brookings Papers on Economic Activity*, 2:2003.
- Easterly, William and Ross Levine (2001). It’s Not Factor Accumulation: Stylized Facts and Growth Models. *The World Bank Economic Review*. Vol. 15, No. 2, 177–219
- Engle, R., and C. Granger (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55, 251-276
- Galor Oded (2005). From Stagnation to Growth: Unified Growth Theory. In Aghion P., Steven N. Durlauf (Eds), *Handbook of Economic Growth*. Elsevier B.V., Amsterdam
- Galor Oded and Omer Moav (2004). From Physical to Human Capital Accumulation: Inequality in the Process of Development, *Review of Economic Studies*
- Hall Robert E. and Charles I. Jones (1998). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics*, Vol. 114, No. 1, (Feb., 1999), pp. 83-116
- Hulten, Charles R. and Frank C. Wykoff (1996). ”Issues in the Measurement of Economic Depreciation: Introductory Remarks”. *Economic Inquiry* 34, pp 10-23
- Islam, N. (1995) Growth Empirics: A Panel Data Approach. *Quarterly Journal of Economics*, 110, 1127-70
- Johansen, S. (1995). Identifying restrictions of linear equations. With applications to simultaneous equations and cointegration. *Journal of Econometrics*, 69(1), 111-132

- Johansen, S. (1996). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, 2. edn. Advanced Texts in Econometrics, Oxford University Press: Oxford.
- Jones, Larry E. and Manuelli, Rodolfo E. (1990). “A Convex Model of Equilibrium Growth: Theory and Policy Implications”, *Journal of Political Economy*, 98, 5 (October), pp. 1008–1038
- Jones, Charles I. (2002). *Introduction to Economic Growth*, second edition, W.W. Norton and Co., New York
- Jones, Charles I. (2005). “Growth and Ideas”, in *Handbook of Economic Growth*, Volume 1B, Handbooks in Economics 22, Elsevier B.V., Amsterdam, Netherlands, pp. 1064–1111
- Jorgenson, Dale W., Fraumeni, Barbara M. (1989). The Accumulation of Human and Nonhuman Capital, 1948-84. In R. E. Lipsey and H. S. Tice (eds.) *The Measurement of Saving, Investment and Wealth*, Studies in Income and Wealth, Vol. 52, Chicago, IL, University of Chicago Press. Reprinted in D.W. Jorgenson (1995). *Productivity, Volume 1: Postwar U.S. Economic Growth*, Cambridge, MA and London, England, The MIT Press
- Jorgenson, Dale W., Fraumeni, Barbara M. (1992a). The Output of Education Sector. *Output Measurement in the Service Sector*, in Z. Griliches, (ed.) Studies in Income and Wealth, Vol. 55, Chicago, IL, University of Chicago Press. Reprinted in D.W. Jorgenson (1995). *Productivity, Volume 1: Postwar U.S. Economic Growth*, Cambridge, MA and London, England, The MIT Press
- Jorgenson, Dale W., Fraumeni, Barbara M. (1992b). Investment in Education and U.S. Economic Growth, *Scandinavian Journal of Economics*, 94. Supplement, pp. 51–70.
- Juselius, K. (2006). *The Cointegrated VAR model: Econometric Methodology and Macroeconomic Applications*. Advanced Texts in Econometrics, Oxford University press: Oxford.
- Kendrick, John W. (1976). *The Formation and Stocks of Total Capital*, Columbia University Press for NBER, New York
- Lau, L. J., D.T. Jamison and F.F. Louat (1991). Education and Productivity in Developing Countries: An Aggregate Production Function Approach. Policy Research Working Paper # 612, World Bank
- Levine, R. and D. Renelt (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review*. 82, 942-63
- Lucas, Robert E. Jr. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics*, 22, 1 (July), 3-42.
- Mankiw, N. Gregory & Romer, David & Weil, David N. (1992). A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics*, vol. 107, 1992:2, 407–37.
- OECD (2009). *Measuring Capital: OECD Manual 2009*, ISBN 978-92-64-02563-9, OECD
- Romer, Paul M. (1986). Increasing Returns and Long-Run Growth, *Journal of Political Economy*, vol. 94, 1986:5, 1002–37.

Romer, Paul M. (1987). Growth Based on Increasing Returns Due to Specialization, *American Economic Review*, vol. 77, 1987:2, 56–62.

Romer, Paul M. (1990). Endogenous Technological Change, *Journal of Political Economy*, vol. 98, 1990:5, pt. II, S71–S102.

Rebelo, Sergio (1991). “Long-Run Policy Analysis and Long-Run Growth”, *Journal of Political Economy*, 99, 3 (June), 500–521

Schultz, Theodore W. (1961). Investment In Human Capital. *The American Review of Economics*, Vol. LI, March 1961, Number one, 1-17.

Self, Sharmistha and Grabowski, Richard (2003). “Education and long-run development in Japan”, *Journal of Asian Economics*, 14 (2003), 565–580

SNA93 (1993). *System of National Accounts 1993*, Commission of the European Communities-Eurostat, International Monetary Fund, OECD, United Nations, Brussels/Luxembourg, New York, Paris, Washington D.C

SNA2008 (2009). *System of National Accounts 2008*, Commission of the European Communities-Eurostat, International Monetary Fund, OECD, United Nations, New York

Truong, Công N. and Binh Tran-Nam (2007). Endogenous Growth and Publicly Funded Knowledge Accumulation. *Review of Development Economics*, 11(2), 421-435