Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 August; 6 (9s): 938-945

Canonical Correlation between Human Capital and Socio – Economic Indices

Alia Osman Sayed Saad
Assistant Professor in Econometrics & Social Statistics Imam Abdulrahman bin Faisal University- Dammam – Saudi Arabia

Received: 24- June -2023
Revised: 27- July -2023
Accepted: 21- August -2023

Abstract

This research reveals the association between human capital indices (mortality rate of children under 5 years, adult survival rate, Harmonized Test Scores (HTS) and expected years of school) and socio – economic indices (Life expectancy at birth, GDP per capita, and government expenditure on education). Using the canonical correlation analysis CCA, that assessing the association between the human capital indices as dependent set and socio-economic indices as predictor set. Data were fromthe World Bank (2018) for 108 countries. The Wilks's criteria show that the full model for all functions statistically significant; the full model explain 97.6% of the variance between the dependent and predictor sets. The first canonical correlation explicit a higher positive significant correlation 0.98 (P,0.000) between human capital and socio – economic indices. 74% of the human capital index variability is construed by the firstcanonical variate of socio-economic indices and 47% of the socio- economic indices' variability is construed by the first canonical variate of human capital indices in the model. Life expectancies at birth has the biggest contribution to the first socio – economic canonical variate amongother predictors. The expected years of school and probability of survival to age 5 contribute more than the adult survival rate and Harmonized TestScores (HTS) to human capital.

Keywords: Adult Survival Rate, Canonical Correlation Analysis, Structure Coefficient, Harmonized Test Scores, Human Capital, Life Expectancy at Birth, Mortality Rate.

Preamble:

Capital is a stock of dividends and return on investment for the progression in future. Human capital is the productivity of human resources which increases the salary as the productivity enhances. The productivity of human capital is correlated with abilities and lucrative information which including the knowledge, skills, and ability to perform tasks. Human capital includes education, training, skills, intelligence, and health. Investment on human resource in a country enhance the abilities which willincrease the productivity. Inability to keep up with innovation depreciatesthe human capital in long run like other assists. In global economies, human capital usually tends to migrate, which known as human capital flight.

Economic growth and the per capita income development are in related strongly with the human capital growth. therefore, some countries adducefree higher education, and this provide more skilled labor which will increase the gross domestic product and per capita income that will affect positively on the purchasing power in the country. skilled labor in the different areas such as technology, health, education, and management willenhance the economic growth, that will improve the welfare in different aspects; social, health and economic.

Adam Smith mentioned the concept of human capital as knowledge,training, talent and experience in 18th century and he explain that the education and well training increase the wealth of society. Schultz (1961) stated there is a variation in workers' wages because of differentials in education and health level as components of the human capital. Schultz imputed the variation of national output between countries to the level of investment on human capital. Becker (1962) suggested that the investmentin education and training for the labor in any organization will increase their abilities and improve the production and its quality. Liu and Wong (1981) explained that the determination of the main sectors for the investment will cumulated the physical capital; one of these sectors is the human resources which is one of the basic assets to achieve efficiency economic growth. the investment in human resources will improve the skills, the abilities and upgrade the human capital. Lucas (1988) presentedthat schooling can transform the unskilled labors into a proficient labor. Schooling is one of the main factors to develop

Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 August; 6 (9s): 938-945

the human capital that willgrowing the per capita productivity sustainably. Benhabib and Spiegel (1994) estimated a Cobb-Douglas production function using data for the physical and human capital stocks as explanatory variables. but the humancapital was unsignificant variable in Cobb-Douglas production function. therefore, they estimated author model consist of the growth productivity rate and human capital stock. they found that the growth productivity rate increase as the human capital stock increased. Fogel (1994) discussed thatthe development in public health, medical care, and the technological change between 1790 and 1980 in British contributed to 30% of the incomegrowth. Temple (1999) shows that the investment in human capital and its'stock construes the variations between countries in per capita income growth. Which agree with Barro and Lee (1996), and Sachs and Warner (1997) results. Bhargava, Jamison, Lau, and Murray (2001) argue that theadult survival rate as a health indicator has an impact on GDP growth rate.His results revealed a positive effect of adult's survival rate on GDP growthrate in low -income countries. Gyimal- Bermpong and Wilson (2004) investigated the effect of health and human capital on the growth of per capita income. They found that the health stock has a direct positive effecton per capita income level, and the additional health investment in Sub Sahara Africa has positive effect on the economic growth at the macro level. Armstrong (2006) defined that human capital is all capabilities of a human including acquired attributes or innate, whose value could beimproved by appropriate investments. Almendarez (2013) suggested that the good education with high quality will develop the economic growth.

The economics and growth studies explain that the improvement in health,human and physical capital continuously will increase the available resources productivity and leads to the technical change, which will enhance the per capita income growth, and the surplus will increase the investments and give further growth in income. Health economics studies revealed that expenditure on health care has a positive effect on income. As Schultz (1999) has explained that nations' health level will determine and ensured the development and economic growth.

Data:

The first set Z: (Z_1, Z_2, Z_3, Z_4) are the dependent variables, which are the human capital indices. Z_1 is defined as the probability of survival to age 5in a country. Z_2 is measured as the rate of adult survival in a country. Z_1 and Z_2 are the best indicators of the stock of a nation's health human capital(Shultz (1999)). While Z_3 and Z_4 measure the expected level of education and grades; where Z_3 is the Harmonized Test Scores (HTS) as a measure of education quality in a country, and Z_4 is the expected year of school.

Kraay (2018) denoted in his HCI definition to the high-quality education and good health for children today will increase the productivity in future.

The second set W: (W_1, W_2, W_3) are the predictor variables, which are thesocio-economic indices. W_1 is the life expectancy at birth stands out as health stock indictor, which denotes to the population wellbeing in a country. W_2 is defined as the annual GDP per capita. W_3 is the governmentexpenditure on education.

Canonical Correlations Analysis:

Canonical Correlation Analysis (CCA) is a multivariate method. Knapp (1978) show that CCA could conduct the univariate and multivariate classic methods. CCA explain the simultaneous correlation between two sets of predicted and dependent variables. Thompson (1991).

This research aims to achieve two goals, which are pursued in canonical correlation analysis. These goals can be introduced as:

1. Looking for the first pair of linear combinations " first canonical variable", one from the human capital index and the other from the socioeconomic index. the first canonical correlation is the maximumcorrelation between the canonical variables. the coefficients of the canonical variables are known as "canonical weights" which interpret the association between the two sets.
2. Calculate the 'structure correlations' which is the correlation betweenthe canonical variables and the original variables.

Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 August; 6 (9s): 938-945

The Canonical Correlation Analysis method generate the linear combination for two sets of variables to maximize the correlation.

If $W = (W_1, \ldots, W_p)'$ and $Z = (Z_1, \ldots, Z_q)'$ are two random vectors, the meanof these vectors are μ_W and μ_Z respectively have covariance matrices Σ_W and Σ_Z.

The combination matrix for the two vectors W and Z is $K' = (W', Z')$ and the K matrix is normally distributed; $K \sim (\mu, \Sigma)$ where:

$\mu' = (\mu'_W, \mu_Z)$ and

$$\Sigma = \begin{pmatrix} \Sigma_{WW} & \Sigma_{WZ} \\ \Sigma_{ZW} & \Sigma_{ZZ} \end{pmatrix}$$

Where the covariance matrix of W&Z is:

$$\Sigma_{WZ} = E[(X - \mu_W)(Y - \mu_Z)']$$

So, the new variables X and Y are the linear combinations of W & Z

$$X = a'W \quad Y = b'Z$$

For the sample:

Where the mean and covariance of the sample is:

$$\bar{K} = \begin{pmatrix} \bar{W} \\ \bar{Z} \end{pmatrix} \qquad S = \begin{pmatrix} S_{WW} & S_{WZ} \\ S_{ZW} & S_{ZZ} \end{pmatrix}$$

Where:

$$\bar{W} = (1/n)\sum_{i=1}^{n} W_i \quad and \quad \bar{Z} = (1/n)\sum_{i=1}^{n} Z_i$$

$$S_W = \frac{1}{n-1}\sum_{i=1}^{n} (W_i - \bar{W})(W_i - \bar{W})'$$

$$S_Z = \frac{1}{n-1}\sum_{i=1}^{n} (Z_i - \bar{Z})(Z_i - \bar{Z})'$$

$$S_{WZ} = S'_{ZW} = \frac{1}{n-1}\sum_{i=1}^{n} (W_i - \bar{W})(Z_i - \bar{Z})'$$

Canonical Variates and Correlations:

The new variables X&Y are the canonical variates X = α' W and Y = β' Z,the sample has these properties:

$$\hat{Var}(X) = \alpha' S_W \alpha$$

$$\hat{Var}(Y) = \beta' S_Z \beta$$

$$\hat{Cov}(X,Y) = \alpha' S_{WZ} \beta$$

(X_1, Y_1) is the sample first pair canonical variates with linear combinationvectors {α1, β1} that maximize the correlation between the canonical variates:

$$\hat{Corr}(X,Y) = \frac{\hat{Cov}(X,Y)}{\sqrt{\hat{Var}(X)}\sqrt{\hat{Var}(Y)}} = \frac{\alpha^F S_{WZ} \beta}{\sqrt{\alpha^F S_W \alpha}\sqrt{\beta^F S_Z \beta}} \quad \ldots\ldots\ldots\ldots (8)$$

Subject to:

- The first pair canonical variates X_1 and Y_1 have a unit vector that give us the maximum correlation in equation (8).
- The L^{th} pair (X_k, Y_k) have a unit variance matrix that give us the maximum correlation in equation (8) between all options which areuncorrelated with the previous L-1 canonical variables pairs.

The Calculation of the Canonical Variates and Correlation:

The L^{th} pair of canonical varieties estimated by:

$$\hat{X}_L = \hat{\alpha}_L S_W^{1/2} W \quad \text{and} \quad \hat{Y}_L = \hat{\beta}_L S_Z^{1/2} Z$$

Where:

- $S_W^{-1/2} S_{WZ} S_Z^{-1} S_{WZ} S_W^{-1/2}$ is the L^{th} eigenvector $\hat{X}_L$.
- $S_Z^{-1/2} S_{WZ} S_W^{-1} S_{WZ} S_Z^{-1/2}$ is the L^{th} eigenvector $\hat{Y}$.

Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 August; 6 (9s): 938-945

Y XY X XY Y L

And the estimated Lth canonical correlation is:

$\hat{Corr}(X, Y) = \hat{\rho}$

Where:

- $S^{-1/2}$ S^{-1} S $S^{-1/2}$ or $S^{-1/2}$ S S^{-1} S $S^{-1/2}$ Give us the

X XY XY X Y XY X XY Y Y

Lth eigenvalue $\hat{\rho}$.
L

- $\hat{X} = \hat{\alpha}' W$ where $\hat{\alpha} = [\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_p]$ and the p canonical variates in W are $\hat{X} = (\hat{X}_1, \ldots, \hat{X}_p)'$.
- $\hat{Y} = \hat{\beta}' Z$ where $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2, \ldots, \beta_q]$ and the q canonical variates

in Z are $\hat{Y} = (\hat{Y}_1, \ldots, \hat{Y}_q)'$.

- We are used the first p canonical variates in Z If $p \leq q$.

Redundancy Index:

Redundancy Index is the percentage of variance in the actual variables in the first set in the canonical function that explained by the canonical variatein the other set in that function. Dattalo (2014).

Structure coefficients:

Measure the simple linear correlation between the canonical variates and its independent variables. Wickramasinghe, (2019).

Results:

In this research, we seek to quantify the association between human capitaland some of the socioeconomic indices. CCA is a suitable technique to determine the correlation and its power among the two sets of the indices.firstly, find the pair linear combinations that maximize the correlation. secondly, we determine the pair of linear combinations with maximum correlation between other pairs which are uncorrelated with the initially selected pair and so on.

Using the World Bank data (2018) for 108 countries; we measured the canonical correlation between the human capital indices: probability of survival to age 5, adult survival rate, expected years of school, and harmonized test scores as dependent variables, and socioeconomic indices:life expectancy at birth, GDP per capita and government expenditure on education as predictor variables.

Copulatively, the full model across all functions was statistically significant using the Wilks's $\Lambda = 0.0243$ criterion, $F = 68.593$, and p-value
$= 0.000$. Wilks's Λ is the unexplained variance by this model, 1- Λ yields the explained variance of the full model. Therefore, the full model can explain a substantial portion, about 97.6% of the variance shared betweenthe human capital and socio-economic indices.

The hierarchical arrangement of functions could be tested by the Dimension Reduction Analysis. As the result show above, the functions (1-3) were highly significant since the criterion of Wilks's $\Lambda = 0.0243$, and theF - statistic =

68.593 with p-value = 0.000. And the Functions 2–3 were also statistically significant with criterion of Wilks's Λ= 0.697, and the F -statistic = 6.72 with p-value = 0.000. The test of Function 3 was also statistically significant as the criterion of Wilks's Λ show that its value is 0.899 with F- statistic = 5.767, and its p-value = 0.003.

Table (1): Canonical Variates – Correlation and Redundancy Index

Canonical variate	Canonical Correlation	Square Canonical Correlation	Redundancy Indexfor W variables	Redundancy Indexfor Z variables
1	0.98	0.9604	0.7395	0.4705
2	0.47	0.2209	0.0026	0.0375
3	0.32	0.1024	0.0012	0.0348

CCA is used to investigate the relation between human capital and socio- economic indices. Table (1) show the correlation of each successive function was 0.98, 0.47, and 0.32. However, the redundancy index of all functions except the first one is about zero. Therefore, the first canonical correlation is noteworthily in the context of this study.

The first redundancy index for the predictor's variables (W) is 74%, whichmeans that the first canonical variate of socio-economic indices explains 74% of the human capital indices in the model. Analogously, the first canonical variate of human capital indices explains 47% of the socio- economic indices in the model as the first Redundancy Index for the dependent's variables (Z) show. Each Redundancy Index denote that the canonical variates for the socio – economic indices explain more variationin human capital indices in the first function than the canonical variate forthe human capital indices explained in the set of socio – economic indices.Dattalo (2014).

The first canonical variate pair is:

$X_1 = 0.135\, W_1 + 0.0001\, W_2 - 0.009\, W_3$

$Y_1 = 10.836\, Z_1 + 6.46\, Z_2 + 0.001 Z_3 + 0.0168\, Z_4$

The first canonical variate (standardized coefficients) pair is:

$X_1 = 1.06\, (W_1 - \overline{W_1}) / S_{W_1} + (-0.09)\, (W_2 - \overline{W_2}) / S_{W_2} + (-0.02)\, (W_3 - \overline{W_3}) / S_{W_3}$

$Y_1 = 0.34\, (Z_1 - \overline{Z_1}) / S_{Z_1} + 0.04\, (Z_2 - \overline{Z_2}) / S_{Z_2} + 0.07\, (Z_3 - \overline{Z_3}) / S_{Z_3} + 0.60\, (Z_4 - \overline{Z_4}) / S_{Z_4}$

Table (2): Structure coefficients (Canonical Loadings)

W's variables	canonical variates: X_1	Z's variables	canonical variates: Y_1
Life expectancy at birth	0.99763	Probability of Survival to Age 5	0.94518
GDP per capita	0.64951	Adult Survival Rate	0.84681
Government expenditure on education	0.20253	Harmonized TestScores	0.79468

Journal for Re Attach Therapy and Developmental Diversities
eISSN: 2589-7799
2023 August; 6 (9s): 938-945

		Expected Years of School	0.97260

The first canonical variate X_1 can be interpreted as health and economic status with positive weight in this variate and education expenditure with negative coefficient. First canonical variate Y_1 have positive weights withall the dependent variables in the model, which represent health and education human capital stock in a country. The first pair of canonical variates are correlated strongly. The first canonical variates explain 96.5%of the variance shared between the human capital and socio-economic indices. Cooley and Lohnes (1971).

Since the life expectancy at birth is correlated very strongly and positivelywith the first canonical variate U_1, the life expectancy at birth contributed more to the first canonical variate U_1 than GDP per capita and government expenditure on education. GDP per capita has a moderate and positive correlation with the first canonical variate U_1, and the first canonical variate is correlated weakly with percentage of government expenditure oneducation. The first canonical variable U_1 explain 49% of the variance in predictor variables. Table (2).

The first canonical variate V1 is correlated strongly and positively with alldependent variables in the model. But expected years of school andprobability of survival to age 5 contributed more to the first canonical variate V1 than the adult survival rate and Harmonized Test Scores (HTS)relatively. The first canonical variate V1 explain 77% of the variance in dependent variables. Table (2).

Discussion:

This study investigates the association between human capital indices andsocio – economic indices using CCA. Instead of using separated linear regression models for each human capital indicator, we used the CCA in view the fact that it measures the effects of multiple predictor variables onmultiple dependent variables simultaneously in one model. Since the CCAinvests all the information in the two groups of variables and arrives at anefficient estimate of the relationship between the two groups, the CCA presents a more efficient model for measuring the effects of socio- economic indicators on human capital indicators. CCA used the two sets predictor and dependent set to limit the inefficiencies multiple testing thatmay occur to reduce type one error. Furthermore, the latent variablemethod averts the multicollinearity problem to arise, Liu, Drane, and Wu (2009). The result of CCA gives a global view of the association between indicators of human capital and socio-economic. We found that 97.6% of the variance shared between the human capital and socio-economic indices. In addition to providing an assessment of the association betweentwo sets of variables, the application of CCA helped in determining the interaction between the two sets that might contribute to the relationship based on the maximum conical correlation of the first pair of canonical variates, which explains 96.5% of the variance shared between the humancapital and socio-economic indices. In addition to, the first redundancy index for socio-economic indices explains 74% of the human capital indices in the model and the human capital indices explains 47% of the socio-economic indices in the model. As Sherry and Henson (2005),structure coefficients are answers to the question "what variables are contributing to the relationship between the variables set across the functions?". The life expectancy at birth has got the highest structure coefficient, which is a common and important predictor for human capitalprogress. The second variable among the predictors sets (based on its structure coefficient) is GDP per capita. Thus, the country with high GDPper capita has a positive and significant effect on the human capital. The health economics literature indicates a positive effect of income on expenditure on health care Gyimal- Bermpong and Wilson (2004). In this study, the variable of government expenditure on education got the third place among the predictors. The expected years of school and probability of survival to age 5 have got higher structure coefficient than adult survivalrate and Harmonized Test Scores (HTS) relatively, which are common andimportant for the socio – economic indices development. The expected years of school and Harmonized Test Scores (HTS) reveal the quality of education in a country which are the key to the development of human capital Lucas (1988), and economic growth (Wong, (1981) and Almendarez (2013)). Variation of human capital among different countriesand its investment explain the variation between countries in economic growth (Barro and Lee (1996), Sachs and Warner (1997), and Temple (1999)). Probability of survival to age 5 and adult survival rate determine the health capital stock level in a country; they have a positive and direct impact on GDP per capita growth.

Conclusion:

Inequality in human capital between countries refer to socio – economic indicators variation. The main socio-economic indicator associated with human capital indices is the life expectancy at birth. Life expectancy at birth is the main cumulative indicator for the health capital stock in the country. Life expectancy at birth and GDP per capita have an effective contribution to the human capital indices.

References:

1. Almendarez, L. (2013). Human capital theory: Implications for educational development in Belize and the Caribbean. Caribbean Quarterly, 59(3-4), 21-33.
2. Armstrong, P. 2006. A Handbook of Human Resource Management Practice, 10th edition, Amazon.
3. Becker, G. S. 1962. Investment in human capital: a theoretical analysis,Journal of Political Economy, vol. 70, pp. 9–49.
4. Benhabib, J. and Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data, Journal of Monetary Economics, vol. 34, no. 2, pp. 143–173.
5. Bhargava, A. Jamison, D. T. Lau, L. J. and Murray, C. J. L. (2001) Modeling the effects of health on economic growth, Journal of Health Economics, vol. 20, no. 3, pp. 423–440.
6. Barro, R. J.,&Lee, J.W. (1996). International measures of schooling years and schooling quality. American Economic Review, 86(2), 218–223.
7. Cooley, W. W., & Lohnes, P. R. (1971). Evaluation research in education. New York: Irvington.
8. Dattalo, P. (2014). A demonstration of canonical correlation analysis with orthogonal rotation to facilitate interpretation. Unpublished manuscript, School of Social Work, Virginia Commonwealth University, Richmond, Virginia.
9. Fogel, R. W. (1994). Economic growth, population theory, and philosophy: The bearing of long-term processes on the making of economic policy. American Economic Review, 84(3), 369–395.
10. Gyimal- Bermpong, K. and Wilson, M. (2004). Health human capital and economic growth in Sub-Saharan Africa African and OECD countries, The Quarterly Review of Economics and Finance, 44, 296- 320.
11. Henson, R. K. (2000). Demystifying parametric analyses: Illustrating canonical correlation as the multivariate general linear model. MultipleLinear Regression Viewpoints, 26, 11–19.
12. Liu J, Drane W, Liu X, Wu T (2009) Examination of the relationshipsbetween environmental exposures to volatile organic compounds and biochemical liver tests: application of canonical correlation analysis. Environmental research 109: 193–199.
13. Lucas Jr, R. E. (1988). On the mechanics of economic development, Journal of Monetary Economics, vol. 22, no. 1, pp. 3–42.
14. Kraay, A. (2018). Methodology for a World Bank human capital index.The World Bank.
15. Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. Psychological Bulletin, 85, 410–416.
16. Liu, P., and Wong, Y. (1981), Human Capital and Inequality inSingapore, Economic Development and Cultural Change, 29, (2), 275-
17. 93
18. Sachs, J., & Warner, A. (1997). Fundamental sources of long-rungrowth. American Economic Review, 87(2), 184–188.
19. Sherry, A., & Henson, R. (2005). Conducting and interpreting canonicalcorrelation analysis in personality research: A user-friendly primer. Journal of Personality Assessment, 84, 35–46.
20. Schultz, T. W. 1961. Investment in human capital, The American Economic Review, vol. 51, no. 1, pp. 1–17.
21. Schultz, T. P. (1999, Summer). Health and schooling investments in Africa. Journal of Economic Perspectives, 13(3),67–88.
22. Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. Measurement and Evaluation in Counseling and Development, 24, 80–93.
23. Wickramasinghe, N.D., (2019). Canonical correlation analysis: an introduction to a multivariate statistical analysis. Journal of the Collegeof Community Physicians of Sri Lanka, 25(1), pp.37–38.