Human Development Index In Input Output Framework- An Alternative Approach

Shri Prakash
Birla Institute of Management Technology Knowledge Part – II, Greater Noida

To Be Presented At

Fifteenth International Input Output Conference

27th June-1st July, 2005

Beijing, Peoples Republic of China

Human Development Index In Input Output Framework- An Alternative Approach

Shri Prakash[1]

Abstract

The paper attempts to push forward the threshold of input output modeling to encompass HDI within its analytical framework and it enriches HDI estimation by encompassment of i) input output modeling, and ii) an alternative approach as an alternative to conventional principal component analysis and Kendall's method of ranking of rank sums.

The paper focuses on *an alternative approach to HDI to overcome* i) limitations of arbitrary determination of weights attached to different factors/variables used currently in compiling UNDP's and Planning Commission's HDI; ii) extend the coverage to variables procedures and mechanisms of human resource accumulation that are not covered either by UNDP or Planning Commission, including qualitative facets, Planning Commission's and UNDP's indexes suffer from numerous infirmities due partly to the negligence of a) competence and competency acquired from informal systems of training and learning; and b) qualitative capabilities of human resources; and iii) identify and incorporate wastage *involved in deskilling and unlearning* involved in a) unemployment, b) mal-utilization of knowledge and skills, and c) ageing and retirement. *Last factor transforms resources into non resource.*

Estimation of both direct and indirect components of HRD shall be facilitated by input output modelling. An *alternative method of factor/principal component analysis has been used to determine objectively a system of weights to be assigned to different factors/variables used in HDI*. The method incorporates a) rank correlation analysis; b) concordance analysis; and c) rank sums. Empirical part is based on 115x115 input output table for the year 1993-94. The table has been compressed into 45x45 sectors.

For the identification of variables for grouping, rank correlation analysis has been used, while the reduction of variables in initial clusters, concordance analysis has been used. Concordance coefficients have also been used for scientific/objective determination of weights to be attached to factors for incorporation into composite index. Variables used pertain to output estimates based on the i) consumption of basic goods, ii) health care, pharmaceutical health services products, and iii) education. Total output of consumption goods, determined on the basis of consumption expenditure has ben allocated among the states according to the states SDP. Sectoral Productivity has been decompose according to the proportion of sectoral employment of a) uneducated, b) professionally under qualified, and c) professionally qualified personnel.

Context

This paper seeks to develop an Input Output Model for Estimating Human Development Index to capture both direct and indirect, and instantaneous and medium run effects of processes/ sources of

[1] Dean (Research), Birla Institute of Management Technology, Plot No. 5, Knowledge Park-II, Institutional Area, Greater Noida.

I am thankful to Ms. Neelam Mishra, Ph. D. Scholar at BIMTECH, for doing major part of calculations. Usual disclaimers apply.

human resource development. Some of these processes/ sources and variables involved in HD have been overlooked in UNDP's and Planning Commission's estimates of HDI. The Commission's approach is a replica of UNDP's concept, method and approach, except that it incorporates *spatial variation of educational development in Educational Development Index (EDI).* Besides, the Commission developed this *index at sub-national levels* also (For details See, Mohanty, 2003). *The author of this paper and his associates at NIEPA have developed a Comprehensive Index of Educational Development both at the level of states and India in 1995*. The Educational Development index has been based on as many as 31 variables, including the qualitative ones (Prakash et. al, 1995). Planning Commission, probably in its eagerness to conform to UNDP framework, has overlooked that study. UNDPs and Commission's indexes have mostly the same concepts and share a common method of estimation. Errors of commission and omission are also the same.

Three basic dimensions, covered by UNDP and imitated by the Planning Commission are i) Healthy long life as measured by *life expectancy at birth*; ii) *Knowledge* as measured by literacy and enrolments in primary, secondary and tertiary education; and iii) A decent *standard of living*, as measured by per capita GDP.

Limitations of Conventional Index

Numerous variables, which affect human development are overlooked e.g. expected life depends on adequate consumption of food, nutrition, child and mother care, specially maternity child bearing, health, sanitation, potable water supply, hygiene. Moreover, per capita GDP as a measure of standard of living is not appropriate because increase in GDP need not necessarily reflect a higher standard of living. Personal distribution of income and wealth, besides the level of GDP, level and structure of taxation, the determinant of disposable income, number and quality of goods available for inclusion in consumption basket, and level and structure of prices are important determinants of the standards of living. *Per capita income is only a crude proxy*. Compilation of human poverty index (HPI) offers to extent some corrective to this proxy. But poverty focuses on i) relative lower level of living, and ii) the relative extent of deprivation. It overlooks the level of fulfillment, and prosperity, the counterpart of deprivation as reflected by the general consumption standard of living. Incorporation of consumption explicitly will meet the requirements of adequacy and representativeness as a measure of welfare of the population as a whole. The poor may derive relatively less benefit as the trickle down effect of development may not operate. A proportion of GDP, is saved, while the rich may spend a lot on luxuries or foreign tours etc.. Then income does not take public provision of the social services into account, which have an important bearing on the standard of living.

The basic idea of the study is to i) include additional variables to have a wider coverage in HDI an ii) integrate alternative to principal component analysis in commodity production model. An attempt has been made to include the following facets of life which have a direct bearing upon human resource development:

I. the following variables relating to expected life and health have been considered: i) birth rate; ii) death rate; iii) number of doctors; iv) population served per doctor; v) number of PHC's and CHC's and sub centres; vi) average rural population served; vii) auxiliary health workers; vii) number of hospitals; viii) number of beds. Output of pharmaceutical industry and proportion of population covered by preventive and promotive measures of health also need inclusion;
II. numerous variables pertaining to education and training; and

III. variables related to the mitigation of poverty and promotion of welfare such as public distribution of goods of basic needs, measure of inequality of income distribution, and range and quantity of consumer goods.

Planning Commission, Government of India, has developed Human Development Index, including education index, to supplement rather than supplant UNDP index. Both these indexes suffer from *numerous infirmities* due to i) *exclusion of some processes and sources of human resource development, ii) arbitrarily determined weights, and iii) exclusion of important variables from the index.* Limitations and shortcomings of these indices are *both conceptual and methodological*. UNDP's HDI neglects a) *certain elements of competence and competency acquired through i) heritage and nature's gifts to man, ii) formal training, and iii) non-formal systems of training and learning*, specially *learning by doing and on the job experience;* b) *qualitative capabilities* of human resources. All these are the recognized sources of human capital formation. UNDP's approach is not consistent in so far as it takes account of outcome(s) of some process/source and some determinants, e.g. determinants say of expected age. It also overlooks variables of vocational, higher and professional education, the dominant determinant of human capital, c) finally, *level and pattern of utilization of available human resources is totally neglected.* Thus, it focuses only on the most rudimentary elements of human resource development, compromising the representative character of the index. The study first explains the importance of the omitted variables in human resource development.

Accumulation of human resources involves the transformation of population into resource, which needs investment. Transformation mitigates supply constraints: a) *Rising incomes with economic growth loosens financial constraint both at macro and micro levels, accelerating private demand and promoting public investment in health and education;* b) *Increased supply eliminates the need for qualificational down gradation to overcome shortages;* and c) *Rising stock and diversification of manpower with the growth of education encompass the changes in level, kind, content and quality of learning.* These changes reflect higher capability profile of younger manpower, affecting both flows and stock of human resources (Prakash, 1995, 2000, Cf. Schultz, 1966). As the supply constraints are mitigated, twin processes of upgradation of the base of human capital comes into operation: i) practioners, personnel without formal education and/or training, are not recruited any more at the entry level; ii) qualifications at entry as well as higher levels are upgraded, a) raising the minimum level of education required; and b) substitution of general by vocational or professional education. Taking only enrolements in different levels of education along with literacy into account for HDI misses the core element of aspect of educational processes, and hence, human capital accumulation in the dynamic state. The wastage, as represented by drop outs and failure lowers both the quantity and quality of human resource development (For details, See Prakash, 1995, Prakash and Chowdhury, 1994). Variables like pass and leavers' ratios, attendance rates, minimum level of learning acquired are important determinants of human capital accumulation rates and their quality.

Human Resource Development

Human resource development depends upon a) the quantity and quality of inputs used in building human capability, b) process(es) of capability building and its quality dimension, and c) outcome/output. UNDP and Planning Commission neglect some elements of all the above three factors and variables associated with these. This study addresses some of these issues and highlights some variables overlooked in the conventional indexes.

Conventional indexes fail to i) distinguish human resources from population, and ii) differentiate between human resource and human capital. Despite the widespread networking of institutions of technical and professional education, colleges and universities, imparting general education, polytechnics and other institutions of vocational studies, non-formal system still remains an important source of supply of lower and lower middle technical and professional manpower. Non formal sector institutions, impart training and build knowledge base of a large segment of population, specially in areas relating to emerging occupations/ professions relating to emerging areas of knowledge. ***De-skilling and un-learning effect of non-utilisation and/or under-utilisation, or mis-matched use of knowledge and skills are also overlooked*** (For Conceptualization and Estimation, See Prakash, 2000, For Deskilling Effect, See M. Jain, 2003). For example, one estimate shows that a minimum of 3-4 years' continuous primary schooling is needed for literacy. But prolonged non-use of the ability to read, write and/or numeracy *pushes the person back into illiteracy* (Prakash et.al, 1994). *This applies to other skills and levels of knowledge/ education also.* Then, past stock of human resources, brought forward into the current year, *requires adjustment for wastage due to ageing, death and retirement and non or under-utilization of acquired capability.*

The current index of HRD also overlooks the difference between i) autonomous process of development of man's capabilities, which materialize even if minimal subsistence for physical survival is available; and ii) non-autonomous process of human development based on investment in man. The *latter process supplements, reinforces and accelerates the outcome of autonomous process on the one hand, and it magnifies and diversifies the effect on autonomous development of human resources on the other.* Autonomous process and some components of non-autonomous process, such as investment in food, potable water supply, nutrition, clothing, housing, hygiene, sanitation, health and health care, focus largely on protective, preventive and curative measures of health on the one hand, and the building of *physical strength, stamina and endurance on the other.* None of these enters into UNDP's index of HD. This study has, therefore, explored several variables pertaining to health and health care on the one hand, and component of sectoral output catering to the consumption needs of human resources employed in various production sectors of living conditions. As against per capita income as a proxy the consumption variable (determined by IO model on the basis of private final consumption expenditure directly affect the building of strength, stamina, endurance and other components of physical capability directly and intellectual capability indirectly on the one hand, it also relates to welfare on the other.

Variables like birth and death rates and expected age which UNDP assigns high importance, need supplementation by variables pertaining to control over maternity related deaths, infant mortality, morbidy, fecundity, and child and mother care measures. These variables affect status of health. Investment in family welfare, including control over population growth, are the important means of HRD. ***Effect of all these factors either on population growth or human resource development are not captured by expected age at birth,*** *notwithstanding the inclusion of birth and death rates in UNDPs HDI.* Variables like expected age do not reveal the quality of life. ***Living a long life does not necessarily mean living a healthy or productive or happy life****.* Frequency of sickness and its duration affect both regularity/punctuality and productivity at the work place.

Age and gender structure, upon which impinges *the dependency ratios,* are equally important from the *view point of growth of both intellectual and physical capability* on the one hand, and participation rates in economic and educational activities on the other. Matrimony and child birth affect female participation in workforce and after a particular age, one has to withdraw from the workforce, *when a productive asset becomes a non-resource*. These two factors *affect the stock of human resources and their availability for utilization.* Gender and age structure of population directly affect participation in

productive activities, whereas unemployment affects utilization of available resources. Similarly, one becomes eligible to partake in education, while after crossing a specific age, most participants go out of education. Then, empirical evidence suggests that younger workers are better educated and more qualified than the older ones (Prakash, 1995). This is an outcome of the following factors: As an economy moves from lower to maturer stages of development a) Educational development occurs concurrently; b) Investment in nutrition, hygiene, sanitation, health and health care, both in absolute and relative terms, rises; and c) Qualitative improvement gets greater attention as a part of the consolidation of gains of quantitative expansion. Still girls and children belonging to SC and ST mainly suffer from various handicaps such as dropping out of school prematurely and/or remaining totally out of school (Prakash et al., 1996). Gender inequality becomes important from this view point. Some estimate of the degree of gender and spatial inequalities needs incorporation in the index. In my view, GDI of UNDP needs to be supplemented, if not substituted by GII, and SII, the gender and spatial inequality index. Net rather than gross enrolments take care of some facets. Why are these variables not used in HDI/EDI by UNDP? A comprehensive composite index needs incorporation of all the pivotal factors of supply and demand. This study has incorporated as many as 37 variables relating to education, 9 variables relating to health and 9 related to output and consumption of basic consumer goods.

Learning by Doing Professionals

The products of non-formal system of learning and training acquired by actually doing the job also end up as professionals. Such professionals are defined as ***Practioners in Manpower Accounts of India.*** Manpower Accounts *show their occupation profile* according to sector of their employment. Number of such professionals and positions they occupy in occupational hierarchy reflect *the past degree and direction of down-gradation of qualifications* to overcome shortages of specific high and middle level manpower for particular occupations/ *sectors. This process also highlights the importance of non-formal learning, training and job experience. This also manifests the level and structure of educational development and its flexibility and responsiveness to the emerging need for new knowledge.*

Change in the proportion of practicing professionals in total may highlight the coverage of demand-supply gap, while productivity may highlight the contribution of this source and mechanism of HRD to economic growth. Incidentally, Guilds and Professional Associations used to impart training for skill building in a non-formalised system not very long ago. Even today, *the large army of mechanics, even those of high priced machines and such high priced and white goods as cars and domestic gadgets, are either self-employed or employed in big organizations,* ***learn their trades and acquire skills through non-formal streams of human resource building. Lower and lower middle manpower is mostly accumulated by this process of HRD.*** Vertical mobility takes many of them very high up in occupational hierarchy.

Formal system of education and training focuses only on high and middle/lower middle level manpower and it is extremely *slow to respond to emerging changes in manpower needs of the economy* in the wake of technological upgradation of production base, which necessitates new skills, occupations and knowledge. *Private initiatives of non formal system(s) are not only flexible but these also fill up the gaps in formal system timely.* In fact, a very large proportion of software operators and average level computer operators/experts in the wake of information technology revolution have emerged largely from non-formal system of learning/training initially. Success of NIIT is just one example. This source of HRD is totally missing in UNDP's accounting. Due to paucity of time, I have not been able to access these data. The updated version of the paper shall use these data.

Training for Filling Precept-Practice Gaps

In-house training focuses on the a) induction of chosen personnel into the culture and work environment of employer organizations; and b) filling up of gaps in the knowledge, skills and training received and that required for particular jobs to be assigned to the employed professionals; and c) there *is in-service training* imparted to human resources of a company or organization from time to time to update their knowledge and skill base. University Grants Commission has established a wide spread net work of Academic Staff Colleges, which organize a number of orientation and subject based refresher courses each years for college and university teachers. AICTE also sponsors Faculty Development Programmes. ***Such training constitutes an important instrument of human resource development to meet the changing knowledge and skill needs of manpower employed by corporates and other organizations.*** Both these facets of training involve huge investment in augmenting human capability. The number of training institutes, training programmes conducted, persons trained thereof and their annual budgets may highlight the importance of training needs of human resources and its cost. *Sizeable investment is made in such sources of human resource development both in the developed and developing countries.* How does one then overlook these in estimating HRD index? All this investment is totally overlooked by UNDP's HDI. Updating exercise of this study is expected to take care of this aspect.

Both conceptually and empirically, *distinguishing the means and mechanisms of HRD from their outcomes* and *differentiating actual from potential outcome/gain* is needed. Students, h*uman resource in the making,* and labour force participation rates, preferably age and gender wise, are also important indicators of human resource development. These two variables together manifest accumulation/flow and stock components of human resource development. Besides, incidence of unemployment, casual and marginal employment, level and pattern of optimum, under and over utilization of available resources facilitate *assessment of actual use and possible deskilling effect of non-utilization.* Then experience also contributes to development of skill, knowledge and understanding of professional needs of the jobs. In this *illustrative and experimental exercise,* time constraint has forced the neglect of some of these variables as a factor of HD.

Alternative Principal Component Analysis

The study proposes to use evolve an *alternative method of factor/principal component analysis for estimating HDI, which takes into account both the quantitative and qualitative aspects.* It also *determines weights objectively.* Alternative method of component analysis had been developed in the context of *quality of education on behalf of Regional Office of UNESCO, Bangkok,* Thailand (Prakash et al, 1995). This study extends that method to *cover non-educational components* of human resource development also. *Synthesis of this model into input output model of the economy has also been attempted. Thus, the study incorporates HDI into input Output analysis to push forward the threshold of IO models and endow the compilation of HDI with a highly sophisticated methodological base.*

Data Base

Data of Indian economy have been used for empirical illustration of the conceptual and methodological base of the alternative approach. Input Output Table of Indian Economy, 1993-94, the last table compiled and made available by Planning Commission, constitutes the base of empirical prognostication of the input output model. Information about health, education and employment has been gathered from such other sources as Census 1991, Economic Survey, Fact Book of Manpower, IAMR and Reports of NSSO.

Method and Model

Complementary sub-models have been developed:

i) A partially decomposable input output model of production;

ii) A model, designed for identifying and clustering of both ordinally and cardinally measured variables into different factors of human resource development; and

iii) A model to determine weights of different factors/ sub-indices to be used in composite HDI.

Results of application of all three models are synthesized and integrated to determine state wise one single composite index of human development.

I. Model of Production System

Want of universal mutuality of interrelations or inter dependencies in developing economies make a decomposable model suit the analysis better. Education and health, in so far as supply of material goods is concerned, have only unidirectional relationships with commodity production. Similarly, commodity production sectors of the economy draw their human capital from education but supply the material inputs to education and training. This also makes decomposable model appropriate for this study. It will also simplify the estimation of Leontief inverse.

Economy has been classified into 2 groups of sectors:

i) sectors pertaining to factors of human resource development;
ii) commodity production sectors, producing goods and services.

The triangularised decomposable input output model is given below:

$$\begin{bmatrix} X^1 \\ X^2 \end{bmatrix} = \begin{bmatrix} A_{11}X^1 & A_{12}X^2 \\ 0 & A_{22}X^2 \end{bmatrix} + \begin{bmatrix} F^1 \\ F^2 \end{bmatrix} \qquad \text{……………………………………………………… (1)}$$

The following solutions hold

$$X^2 = (I-A_{22})^{-1} F^2 \qquad \text{…………………………………..………………. (2)}$$

$$X^1 = (I-A_{11})^{-1} A_{12} (I-A_{22})^{-1} F^2 + (I-A_{11})^{-1} F^1 \qquad \text{………………………………………………….. (3)}$$

Where X^1 is a column vector of output (X^1_1 X^1_2 X^1_3……X^1m); F^1= $F^1_{1,}$ $F^1_{2,}$ F^1_3………F^1m) and f^2 are column vectors w of final use, (F^2_{m+1} F^2_{m+2} F^2_{m+3}………F^2n) and A is input coefficients matrix, partitioned into 3 sub matrices.

Employment

Solution values of X^1 and X^2 from relations 2 and 3 are used in order to determine employment in terms of F^1 and F^2. Employment involved in the production of given output is given by

$\hat{N}^1 = L^1 \hat{X}^1$ …… (4)

$\hat{N}^2 = L^2 \hat{X}^2$ …….. ……... (5)

where $\hat{N}^1$ and $\hat{N}^2$ are row vectors of estimates of sectoral employment. L^1 and L^2 are row vectors of sectoral employment coefficients, a_{0i} where $a^1_{0i} = N^1_{0i} / X^1_i$, that is, labour required per unit of output, $X^1_{i;}$ $\hat{N}^1 = \sum_i^m N^1_{oi}$ is total employment in the first sub-group of sectors. Substituting for X^1 and X^2, we get

$\hat{N}^1 = L^1 (I - A_{11})^{-1} [A_{12} (I - A_{22})^{-1} F^2 + F^1]$ …… (6)

$\hat{N}^2 = L^2 (I-A_{22})^{-1} F^2$ …… (7)

Vectors of labour coefficients of relations 4 and 5 have also been decomposed into three sub-vectors according to level/type of education: L^{11}, L^{12}, L^{13}, and L^{21}, L^{22}, L^{23}. Equations 6 and 7 may be replaced by similar equations to furnish estimates of N^{11}, N^{12}, N^{13}, N^{21}, N^{22} and N^{23} that correspond to L^{1k}, L^{2p}, k,p=1,2,3.

Productivity

Contribution of human capital to output and its growth can be derived from factor productivity. So the productivity has to be estimated. But three different concepts of productivity are distinguished: *productivity for growth; productivity for welfare, and productivity for competition* (Cf. Juan & Febrero).

The gross factor productivity[1], $\hat{P}$ is given by

$\hat{P}^1 = (N^1)^{-1} X^1$ …… (8)

$\hat{P}^2 = (N^2)^{-1} X^2$ ……..(9)

where $\hat{P}^1$ and $\hat{P}^2$ are sector wise column vectors of productivity for two groups of sectors, $p^1_j = \hat{X}^1_j / N^1_{0j,}$ where $l^1=L^1(I-A_{11})^{-1}$, $l^2=L^2(I-A_{22})^{-1}$, and $l^1=(l^1_{01}, l^1_{02,}\ldots\ldots l^1_{0m})$, $l^2=(l^2_{0m+1,}\ldots\ldots l^2_{0n})$.

Use of solution rather than observed values of X in the estimation of employment and productivity focuses on consideration of both (i) direct and indirect requirement of labour for production, (ii) direct and indirect requirement of capital, iii) inclusion of both direct indirect components of net output in productivity. But the aversion of double counting warrants that either i) $\hat{X}^1$ and $\hat{X}^2$ are used to estimate the elements of the vector L of labour coefficients, or alternatively ii) $l^1=L^1(I-A_{11})^{-1}$ and $l^2 = L^2 (I-A_{22})^{-1}$ are used. It shall still be better to use RHS of relations 8 and 9 for estimating productivity. Capital requirements are embodied in final demand from which solution values of X are derived; capital comprises of Gross Fixed Capital Formation (GFCF) and Change in Stocks, and (iii) requirements of growth that are met out of surplus. Matrix A takes into account what Sraffa calls *"Requirement of Self Replacement of commodity production by commodities in the Economy"* at a given level, whereas capital coefficients take care of growth.

Growth Concept of Productivity may be used to evaluate the growth effect of human capital, *which relates surplus of output over inputs* used in production *to human and physical capital requirements*. The surplus output constitutes the supply pool for allocation among consumption, investment and exports. *The consumption component of productivity relates both to human capital accumulation and welfare.*

II. Model of Clustering/ Factoring Variables

Thurstone introduced the concept of factor analysis in 1931. *Factor is defined as a linear combination of two or more variables.* Principal component/ factor analysis may, therefore, be defined as the method of compressing or combining two or more variables to define a new variable, called factor/ component. *Factors/ components reduce the number of variables involved in the system*. Since linear combinations are based on the structure of relationships, *the factor analysis may also be defined as the method of detection of structure of relationships to facilitate classification and clustering of variables according to twin principles of maximizing* i) *homogeneity/ inter-relatedness within the cluster/factor,* and ii) heterogeneity between clusters/factors. Naturally, correlation/ association and/or regression analysis may be used as the base of factor analysis. Since indexing, including the human resource index, involves the further combining of numerous variables/factors into one single value, ***factor/component analysis comes as a handy tool for indexing.*** This study has, therefore, chosen to use an innovative model of factoring/ clustering.

The New Model of Identification and Clustering of variables into factors is an *extension and modification of the Kendall's method of ranking rank sums and then combining these clusters with Kendall's Coefficient of Concordance with a view to further reduce the number of variables within each brad cluster to form factors, and extract weights scientifically rather than arbitrarily. .*

Rank correlation coefficient has been calculated according to Spearman's rather than Kendall's formulation (1975):

$$\rho_{sj} = 1 - [6\sum\{d^2(i)\}/\{n(n^2-1)\}] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(10)}$$

where ρ is rank correlation coefficient, d(i) is difference between two ranks allotted according to criteria s and j to i^{th} unit, n denotes units ranked, and s/j=1…..m are variables/ranking criteria.

The coefficient of concordance, W, is estimated as follows:

$$W = [12S]/[m^2(n^3-n)] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(11)}$$

Where S is the sum of squares of deviations of the sum of ranks of an individual unit from the mean of rank sums according to all m criteria used for ranking n units/entities; mean of ranks is given by

M= (1/2)(n+1) m

where M denotes mean of ranks of rank sums.

χ^2, used to test the statistical significance of W, is calculated as follows:

$\chi^2 = m(n-1)W$ ..... (21)

Factoring Model

Identification and Clustering of variables into distinct factors is the first step. Rank correlation coefficients are the base of this step. Rank correlation coefficients between all possible pairs of variables of n units/entities, ranked according to m criteria, taken two at a time, say ${}^{m}c_2$, have to be worked out for clustering variables into homogenous clusters. One pair of two most closely related variables is ***identified as the core variables of first factor***. Further scrutiny is done for identifying variables for inclusion in the core group of first factor. First factor will have only such variables as are positively correlated significantly with either variable of the core pair. Second group, having variables negatively correlated significantly with either variable of the core pair, is kept apart. Third group will have all other variables, having no correlation with variables of first two groups.

Variables, correlated with either variable of the first core pair, are added to the group one by one successively. Coefficients of Concordance are calculated for successive groups of 3,4…. variables of this group. The significance of observed value of W, Chi-square, χ^2 with (N-1) degrees of freedom will be used as the test, where N denotes total observations, $N \leq mn$. Results are then used for grouping variables into first cluster. If the coefficient of concordance indicates statistical significance the additional variable(s) is included in the first cluster. If the concordance coefficient indicates non-significant agreement, the additional variable(s) is excluded from the core pair. Change in concordance as a result of addition of a new variable to the core group is thus used to decide whether additional variable(s) is retained in the cluster. If a third variable is retained in the group, then another variable, showing high correlation with any of these three core variables, is brought into this cluster of 3 variables. The coefficient of concordance for this augmented cluster of 4 variables is again evaluated. The process is repeated till all variables, having significant positive correlation with any variable(s) of the core group, have been exhausted. In all probability, some variables initially included in the first cluster will be left out. Once the first cluster has been finalized, the process is repeated on the second most closely correlated pair of another two variables till all such variables, that show a very high degree of concordance with second group, have been identified and clubbed together in the second cluster. We end up with say, k clusters of m variables, n-m variable having been excluded from all the clusters, $1 \leq k \leq {}^{m}c_2$. *Each cluster constitutes one factor or component.* **First core set of variables *furnishes one cluster of variables, defined as first factor. Each of the subsequent sets will furnish one cluster/ factor each. The coefficients of concordance and rank correlation are thus used jointly as discriminants for clustering given variables into a limited number of factors.**

All unrelated variables are excluded from the given cluster one by one. ***The coefficient of correlation and coefficient of concordance rather than rank sums are used to eliminate subjective judgment involved in Kendall's method of arriving at appropriate components/ factors,*** each factor having 2 or more variables showing maximum degree of inter-relatedness and concordance.

III. Determining Optimum Number of Variables and Factors

Different factors are formed according to the *principles of i) minimizing variables in a factor, ii) maximizing inter-relationship between variables within the cluster, and iii) minimizing relationship between variables constituting different factors.* Variables excluded from a factor show independence. Each consecutive factor *captures the maximum joint variability of the variables excluded from*

preceding factor. These selected factors will then constitute the base of composite index. In some cases, criterion of significance of correlation and/or concordance may lead to an ambiguity regarding inclusion of a variable(s) in one or other factor, if the variable(s) qualifies for inclusion in more than one factor. Besides, initially formed clusters on the basis of significance of correlation and concordance coefficients may not minimize the number of variables within a cluster. A supplementary condition, over and above the above criteria, is imposed. ***Some value of significant correlation coefficient is specified as the minimum for inclusion of the variables in a factor***. The minimum critical value, taken as the cut off, will thus *represent cases where at least the specified percentage of total change in one variable is explained by the other. Explained proportion of variation of lesser magnitude than the cut-off will leave too wide a margin for residuals to be more dominant than the chosen determinants.* ***The process of factoring should stop at a point where only little relatedness/ joint variability remains for exploration****. This is a point beyond which random variation can not be reduced and systematic variation can not be raised.*

Maximum Number of Factors

There arises the question of number of factors to be included in an index. As more and more factors are formed, less and less joint variability remains to be explored. *Extraction of more and more systematically related variables through factoring also tends to lower random variance*. But the factors formed should neither be far too many nor far too few. There exist criteria to guide the determination of the optimal number of factors to be extracted from the given variables. Kaiser (1960) criterion, widely used in conventional principal component analysis, suggests the retention of factors having an eigen value larger than one. It is based on the assumption that the factor should *explain variance at least equal to that of one original constituent variable,* variances of all variables having been standardized to unity. Cattel suggests the point *where the decrease of eigen values tapers off.* To the right of this point, one finds only *what is defined as Factorial Scree.* Scree, in geology, refers to debris at the lower part of a rocky slope. Kaiser criterion may at times suggest the retention of too many factors, while Cattel criterion may lead to the retention of too few factors. This may suggest the combined use of two criteria.

Can we formulate similar guidelines for determining optimum number of factors for the new method? For the given number of observations, there is a minimum critical value of the coefficient of concordance among the variables of a factor, which varies according to the number of variables included in the factor. The number of variables included in a factor may lie between the minimum 2 and maximum m. *Maximum of these minimum values/ variables may be taken to determine the point below which the debris of the scree will lie.* Besides, average of rank correlations among the variables of *first and last clusters may be taken to set up the average degree of variability* as acceptable for the formation of a factor. These two parameters may be supplemented by the interpretability of the results furnished by differing the number of factors to be finally included in the index.

Generally, 4-5 factors exhaust 95 to 100 per cent of total variation. Besides, first 2-3 factors normally exhaust 65-85 per cent of total variation. Peripheral factors are generally excluded and a rule of thumb and nail is employed to pick up only such factors as explain at least 65-75 per cent of total variation. This generally requires the inclusion of 2-3 factors/components (Prakash and Mohapatra, 1981, Kendall, 1975). This technique may be adopted for the formation of composite index as efficiently as Principal Component/ Factor Analysis or Kenall's method of ranking successive rank sums.

A common *facet of multi criteria of ranking is that larger the number of variables, lower tends to be the degree of inter-relation, and hence, concordance among them. It implies that any factor should not have*

too many variables to jeopardize the homogeneity and coherence, and it should not be too small to make factoring redundant.

IV. Determination of Weights of Factors

The composite index may be estimated from the chosen factors. But the factors have to be weighted. Eigen values of principal components furnish the weights. But Kendall's method does not use weighting. This method derives weights to be attached to each factor as follows: Sum of ranks of each spatial unit, according to variables of each factor is divided by the number of variables constituting a factor for each ranked unit separately. Average of all units for all criteria, taken together, is also determined. Let $\overline{X}(1i)$ be the average rank sum of ith unit, *based on variables of first factor,* and let $\overline{X}$ *be the average of all ranks according to all criteria of all units taken together.* The product of the proportion of divergence of $\overline{X}(1i)$ from this overall average, $\overline{X}$, and the ratio of concordance coefficients of the factor and its difference from unity is defined as w(1i) as the weight of first factor in indexing:

$$w(1i) = \left[\frac{\overline{X}(1i) - \overline{X}}{\overline{X}}\right] . \left[\frac{W(1)}{1 - W(1)}\right] \qquad \text{(15)}$$

where w(1i) is the weight assigned to first factor for ith unit for estimating composite index and W(1) is the coefficient of concordance of first factor. Sum of all weighted values of all factors for each unit will determine its composite index of human resource development.

Data Massaging

First, 40 of the 115 sectors of the economy have been identified as the base of human resource development. These sectors pertain to i) health, health care and health service, ii) education and training, and iii) production of consumer goods, specially the goods satisfying basic needs, and commodities/ services satisfying professional needs. Remaining 75 sectors constitute the group of other commodity production. Both groups of sectors have been aggregated into 45 sectors. Data difficulties, pertaining to employment and its education-occupation composition, led us to aggregate 115 into 45 broad sectors of the economy. Nature and level of aggregation are data rather than concept or method related.

In all, 10 out of 45 aggregated sectors pertain to the provision of goods needed as investment in physical strength and stamina of human resources. Number of sectors in these groups is listed below:

Table - I

Sectors Producing Provision For	**Number of Sectors**
Food	7
Personal Effects/ Necessities	3
Education	5
Health	3
Total	**18**

Sectors, considered as the base of HRD, pertain to food and food processing, clothing, other necessities such as gas, electricity and water supply, transport, education, training and health. Sectors needed to

sustain education and health such as printing and publishing, health care services and pharmaceutical industry also constitute a part of this group.

Employment and Productivity

Productivity is the manifestation of economic contribution of human capital to output and its growth. For its estimation, employment involved in the production of given output has to be determined first. In order to relate productivity to human resource capability, labour coefficients a^1_{0i} and a^2_{0j} have been decomposed into sub-coefficients according to educational qualifications. But the level and type of education required in relation to actual education possessed by employees are considered in relation to i) sector, ii) occupation, and iii) position in the hierarchy: i) *workers having below average general education and training, including illiterates; ii) workers having below average professional/ technical/ vocational education and training; and iii) workers having professional/technical, vocational education required for occupations of their employment* (Cf. Parnes, 1959, Tinbergen and Correa, 1963). Estimates have been derived on the basis of Education-Occupation-Industry matrix, based on information culled from Census data, 1991 and Annual DGET reports.

Sectoral profits have been used to decompose sectoral value added to yield estimates of wage and salary component. These two components have facilitated the estimation of contribution of human and physical capital to value added. Whereas wages and salaries reflect the productivity of human capital, profits manifest the productivity of physical capital. Manpower is the carrier of human capital; while machinery, equipment, and tools are the carriers of physical capital and technology. As a first approximation, *it is assumed that expenses on food, clothing, hygiene, sanitation, potable water, medicines and health care reflect the physical component of human capital,* whereas *investment in education and training reflect intellectual component of human capital.* Three fold classification of sectoral employment according to education is used to estimate the contribution of human capital to output in terms of productivity differentials.

Thus, we have 3 distinct groups of variables that relate to i) a) consumption designed for the building of physical strength and stamina of manpower, b) health and health care services that service the needs of preservation of health, and cure and prevention of disease, and ii) education and training. Since a certain level of both direct and indirect inter-relatedness has already been exhausted through the use of matrix multiplier, these variables are kept as clusters. But within the group, interrelatedness and variability are detected by the use of clustering model in order to extract distinct factors within each group. All above parameters and variables have been used to rank in ascending or descending order of values according to the nature of the variable. For example, lowest death and birth rates command first rank, whereas the highest expected age at birth is ranked first.

On the basis of sectoral employment and average private per person consumption of specified goods, output for final consumption for the sustenance of human capital deployed on productive activities is estimated. Estimated output of such goods is then allocated among the states according to their SDP. *It furnishes estimates of output produced for the sustenance of human capital in the commodity production sectors.* Same procedure is applied to estimate the component of output of education and health to estimate their contribution to the physical and intellectual capabilities of human capital. Population is not included in human capital.

Empirical Results

Experiment shows that the statistical significance of either rank correlation or concordance coefficients is a necessary but not sufficient condition of high degree of inter-relatedness and/or high degree of concordance. Inclusion of too many variables in a cluster leads to a decline in the i) average value of rank correlation coefficients, and ii) degree of concordance.

Therefore, a critical minimum value for significant correlation has been determined. The variables within a cluster, having inter-correlation lower than the critical minimum, are dropped to maximize inter-relatedness and degree of concordance. This facilitates the formation of factors having optimum degree of homogeneity, as determined by inter-correlations and concordance. This explains the exclusion of as many as 25 educational variables from the finally chosen four rather than five factors.

Concordance tends to increase with a decline in the number of variables in a cluster upto a point.

The acceptable degree of inter relatedness and agreement in our empirical exercises is .388 for concordance, and average of 0.56 for rank correlation coefficients. These values have been chosen as the minimum of the cut off point for inclusion of variables in factors. On the whole, average of rank correlations and corresponding value of Chi-square for concordance increase with the increasing degree of concordance.

As many as 37 variables relating to education have been analysed. The following table shows the factors formed, variables they comprise of, average of rank correlation among the variables of each factor, and their corresponding coefficients of concordance:

Table - II

Sector	**Factor**	**Factor Variables**	**Average rank**	**Coefficient of Concordance**
Health	Factor – I	7	49	.388
	Factor – II	2	27	.693
Education	Factor – I	3	49.5	.893
	Factor – II	2	.33	.821
	Factor – III	4	66	.579
	Factor – IV	3	49.5	.67
Consumption	Factor - I	9	38	1

Initial analysis leads to the identification of 15, 8, 8, 3 and 2 variables constituting 5 factors.

After the identification of final clusters and their constituent variables, the variate values have been normalized by the national average corresponding to a given variables so as to make them scale free. The normalized value have then been used to estimate each factor and the associated sub index separately. These have then been used to estimate three indexes, one each corresponding to health, education and consumption. These indexes, in turn, have been converted into a single composite index of human resource development. Results reported in table 2 facilitate the formation of three crude indexes of three factors of human development. The crude indexes have been merged to furnish an estimate of one single composite index of human development, which is expected to differ significantly from UNDP's index in i) magnitude, ii) range of variation among spatial units, and iii) number of variables considered, and hence, the coverage.

National index serves as the reference point for the relative position of individual states. But this index does not have any base year for reference. Results are reported in Table I and Table II.

The results show that i) average human development index has a value of 232.39; ii) the human development index varies sharply among the states; iii) the coefficient of variation has as high a value as 190 per cent; and iv) the range of variation is from the minimum 161.43 for Rajasthan to a maximum of 307.38 for Andhra Pradesh.

A perusal of the three constituent factors of human development index reveals that i) education, the dominant determinant of intellectual profile of human resource, *is the pivot of human development index: on an average, contribution of EDI to composite HD index ranges from approximately two thirds to nine-tenths of the total,* though, like the composite HDI, *EDI and its contribution to HDI varies sharply among the variables;* ii) education is followed by health, though its relative contribution to HDI is much lower than that of education; iii) surprisingly, investment in physical component of HD occupies the last position, its relative and absolute contribution being very low; iv) *these results, however, may act as the sign posts and delineate the road map for future research.* Since the study is highly experimental and illustrative of concepts and methodology, the author expects to come forward with more precise empirical estimates with comparable values of UNDP's HDI in near future.

Main Findings

The results highlight that

i) Contribution of both human and physical capital vary sharply between sectors;
ii) Technology and physical capital dominates sectors where processing of materials are pivotal, whereas knowledge based sectors are dominated by human capital;
iii) Intellectual component of human capital dominates such sectors where high proportion of professional and technical workers are employed;
iv) Effect of technical and professional education and skills thereof is reflected in the contribution of human capital component of productivity;
v) Practioners constitute non-negligible proportion of professional and technical workers (approximately 13%) across the sectors, though the proportion varies greatly between sectors. Naturally, contribution of such constituents of human capital, overlooked by conventional indexes is substantive. So long as on the job learning and non-formal training remain important sources of human capital accumulation and means of filling up the supply-demand gaps, it is not rational to leave it out of reckoning in HDI as is done by UNDP and Planning Commission. A priori reasoning and thesis proposed in this study are thus supported by empirical evidence. UNDP has to have a look at such results; and
vi) For sectors having low capital and technology base such as primary sectors, physical component of human capital emerges as dominant.

Notes:

1. Entire Output is attributed to labour, implying as if it were the only factor used in production. Besides primary factor labour, capital is an important factor of production. The grossness arises from two sources (i) Total output is attributed to labour; and (ii) gross rather than net output is used as denominator.

Net productivity may be estimated by working out the ratio of factors used in production and net output. Net output may itself be conceptualized in two alternative ways: i) Output of a given good net of its inter-industry use/demand for sectoral output, which is defined as surplus output; and ii) value added, which is defined as excess of output over and above the value of intermediate inputs used up in production.

2. Marshall, in a different context, describes the experience of mining industry in 1870's in similar terms. See *Principles of Economics, Macmillan.*

References

Cattel Quoted from Principal Components and Factor Analysis, File:///c:/ Documents %20Settings/ Administrator/ My %20Document

Correa, H. and Tinbergen, J. (1962) Quantitative Adaptation of Education to Accelerated Growth, *Kyklos*, Vol. 15.

Parnes, Herbert, H. (1956) Manpower Planning

Jain, Manish (2001) *Managing Deskilling Effect of IT* Enabled Services (ITES) In Electronics Engineering Industry- A Study of Competence Profiles of Design Engineers, Synopsis of Ph. D. Thesis, BIMTECH.

Kaiser (1960) Quoted from Principal Components and Factor Analysis, File:///c:/ Documents %20Settings/ Administrator/ My %20Document

Kendall, M. G. (1975) *Multi-Variate Analysis,* Charles Griffin and Co., London.

Kendall, M. G. (1962) *Rank Correlation Methods,* Hazuer Publishing Co., New York.

Mohanty, N. K. (2003) A Comparative Study of Education Indices of UNDP's Human Development Report and National Human Development Report of India, *Business Perspectives*, Vol. 5, No. 2.

Oscar de Juan & Eladio Febrero (2000) Measuring Productivity from Vertically Integrated Sectors, *Economic Systems Research*, Journal of the International Input- Output Research Association, Vol 12, No1.

Prakash, S. and Chowdhury, S. (1994) *Expenditure on Education- Theory Models and Growth*, NIEPA, New Delhi.

Parnes, Herbert, S., Manpower Analysis in Educational Planning, in Parnes, *Ibid*.

Prakash, Shri (2000) Utilization of Human Resources, *Business Perspectives*, Vol. 2, No. 2.

Prakash, Shri (1995) *Cost of Education,* Anamika, Delhi.

Prakash, S., Mehta, A. C. and Zaidi, S.M.A. (1995) *Indicators of Quality of Education, National Status Paper*, UNESCO Regional Office, Bangkok and NIEPA, N. Delhi.

Prakash, S. and Mohapatra, A. C. (1980) Development of the State of M.P.-A Study of Intra State and Inter – Temporal Variations, *Third World Planning Review*, Vol. 2, No. 2, Liverpool, U. K.

Prakash, Shri, Srivastava, R., Mehta, A. C. and Zaidi, S. M. I. A. (1994) *Use of Sample Survey Techniques for Educational Statistics*, UNESCO, Paris.

Principal Components and Factor Analysis, File:///c:/ Documents %20Settings/ Administrator/ My %20Document

Thurstone Quoted from Principal Components and Factor Analysis, File:///c:/ Documents %20Settings/ Administrator/ My %20Document