Centre for Efficiency and Productivity Analysis

Working Paper Series
No. WP05/2006

Title

Analysis of Development Performance Using a Development Index Based on Factor Analysis

(Old Title: Constructing Multideminsional Indexes of Economic Structure and Development)

Authors
K. Renuka Ganegodage
Alicia N. Rambaldi
D.S. Prasada Rao
Kam Ki Tang

Date Previous version: January 2007
Date Current version: October 2008

School of Economics
University of Queensland
St. Lucia, Qld. 4072
Australia

ISSN No. 1932 - 4398

Analysis of Development Performance Using a Development Index Based on Factor Analysis♣

K. Renuka Ganegodage, Alicia N. Rambaldi*, D.S. Prasada Rao, Kam Ki Tang

The University of Queensland
Centre for Efficiency and Productivity Analysis, School of Economics

25 October 2008

ABSTRACT

Development involves changes in social, economic and institutional structures. Quantifying development requires a large array of variables with different characteristics often highly correlated. A factor analysis approach using inferential decisions based on computed standard errors is proposed. The significant factors are used to construct sub-indexes of structural characteristics and a new development index (DI). The properties of DI are analysed and contrasted with more traditional measures, real per capita income (RIPC) and the Human Development Index (HDI). The methodology is applied to data on 45 variables for 97 countries for 1995-2004. DI is found to have stronger discriminating power.

Key words **–** development indexes, factor analysis, Jackknife

JEL Classification: O10, C19, C49

♣ This is a new title. The previous version was titled: Constructing Multideminsional Indexes of Economic Structure and Development.

* Corresponding author. The University of Queensland, School of Economics, QLD 4072, Australia. Tel.: +61733656576, Fax: +61733657299. Email: a.rambaldi@uq.edu.au

1. INTRODUCTION

In the age of information abundance, indexes have become an indispensable tool to distil and condense an otherwise incomprehensible amount of information, reducing the processing costs of the end users. In fact, over the last few decades, indexes have increasingly become a fixture in applied research and policy evaluation. For instance, the United Nations Development Program has regularly used the Human Development Index (HDI) and its variants to evaluate development performance of countries (Human Development Report, UNDP, various years); various public and private enterprises have published indexes designed to capture political risks or institutional qualities of countries, ranging from the long established Index of Economic Freedom by the Heritage Foundation, to the more recently developed Governance Indicators by the World Bank (2008). These indexes have already found their way into the wider literature. Several empirical studies have employed HDI to examine inequality (Mbaku, 1997) and sustainability issues (Moran, et al., 2008). Governance Indicators and the Index of Economic Freedom have been used in empirical investigations of issues like government expenditure efficiency (Rayp & van de Sijpe, 2007), the effectiveness of market oriented reforms on economic growth (Pitlik, 2002), and the institution-growth annex (Aralica & Budak, 2004). The proliferation of indexes is not limited to the economics discipline. For example, dynamic factor models are being used in the finance literature as a method to handle a large range of variables; for applications in stock market indexes, see Corielli & Marcellino (2006) and Rocha & Sekkel (2006). The growing use of indexes in applied research clearly points towards the need to undertake research to improve the quality of the indexes constructed.

Indexes compress the information content of a large number of indicators into one or two summary measures. The essence of the index construction process is to extract the "relevant" information content from an array of different indicators and weight them appropriately in constructing a single series. There are several ways to derive such indexes, and factor analysis has emerged a widely accepted “scientific method” to perform such tasks, mostly for its statistical foundations. However, closer inspection of many empirical applications of factor analysis shows major deficiencies in the statistical tools used in several crucial decisions required to obtain the factors. Amongst these decisions relate to the determination of the number of significant factors, identification of the variables that are significantly loaded on each factor, and to decide on the degree of rotation of the factors. The common practice is to base such decisions on arbitrary rules of thumb without any statistical reasoning. The consequence is that the resulting index may not adequately capture the information content of the underlying indicators or might even provide a distorted message. The adverse effects of these arbitrary treatments are likely to be larger when the number of indicators gets larger and the prior knowledge regarding the appropriate composition of the indexes being constructed diminishes. It is therefore important to have sound statistical principles guiding the implementation of the factor analysis procedures. Development indexes are a good example of a situation where the number of indicators can be fairly large and there is limited *a priori* information on the composition of the indexes being constructed.

Measuring development is a challenging task because the term development is often used as a broad, catch-all term which does not have a precise definition to aid its quantification. As a consequence, development often means different things to different people, and measures based on a particular point of view may not gain acceptance among advocates of alternative perspectives. In meeting the challenge of measuring development, the role of income per capita is unequivocal, but its limitations are also well recognised (Booysen, 2002; Cahill & Sánchez, 2001). There are many development indexes aiming to fill the gap, such as the General Indexes of Development (GID) (McGranahan, et al., 1985; McGranahan, et al., 1972), the Morris's Physical Quality of Life Index (PQLI) (Doessel & Gounder, 1991; Morris, 1996; Morris, 1979), the Economic and Social Development Index (ESDI) (Cahill & Sánchez, 2001), and the HDI (HDR, various issues). However, most of them incorporate only a few variables and are unable to achieve a comprehensive measurement of development (Booysen 2002). Embracing the multidimensional concept of development, the pioneering work of Adelman and Morris (1967) tried to include a range of economic, social and institutional indicators in their measure. To overcome problems associated with the multiplicity of indicators and the high correlation among them, they used factor analysis. Although they pioneered the use of this technique in development studies, their estimation procedure is open to the criticisms mentioned earlier.

Against this background, this paper aims to propose an index construction procedure that is more grounded in statistical theory than those previously employed in the literature. The paper then applies the procedure to a cross country dataset to construct a new development index.

Firstly, we adopt image analysis as the estimation method for the factor model. This choice is based on the sound statistical properties of the method and its robustness to lack of normality in the data. Secondly, the paper estimates standard errors of factor loadings to aid the selection of significant factors and variables that are *loaded* on each factor. The term "loaded" is used in factor analysis to indicate that the estimate of the weight of a given variable on a given factor is significant, although in the current literature this is not commonly based on statistical inference. To demonstrate the procedure, we conduct exploratory factor analysis using data on 45 variables for 97 countries at various income levels covering the period 1995 to 2004. We find four significant factors that capture different defining characteristics of development at different stages of development. Two of these factors are able to discriminate between countries of a wide range of development stages, and the factors are then used to construct a single development index. The index is compared to existing indexes, specifically real income per capita and the HDI, allowing us to draw some conclusions about the socio-economic characteristics of a country that appear to impose constraints on the development of countries at the low and middle stages of development.

The rest of the paper is organized as follows. The next section explains the proposed statistical methodology and the estimation procedure proposed for the construction of the sub-indexes and the development index. Section 3 describes the data used in the study. Section 4 presents the results and discussion. The final section offers some concluding remarks.

2. METHODOLOGY

Index construction becomes challenging when the number of indicators to be considered becomes very large and when the indicators are highly correlated. Factor analysis – a multivariate statistical technique – has been commonly used as a method to overcome this problem. The purpose of factor analysis is to describe the covariance/correlation relationship among many variables in terms of a few underlying but unobservable random quantities called factors. These factors have little or a small correlation with each other, and each factor captures a particular underlying aspect of the original variables (Johnson & Wichern, 2002). Our method differs from those used in previous studies in that we compute standard errors associated with factor loadings, which are then used to choose the number of significant factors, the variables that are significantly loaded on each factor and the degree of rotation of the factors. The details of the procedure are explained in the following sections.

2.1 Estimation of parameters and underlying factors

This section follows the work of Wansbeek and Meijer (2000) and Johnson and Wichern (2002). Consider a vector of p random variables, $\mathbf{X}$, which has mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Each random variable represents a socio-economic indicator. The factor model hypothesizes that **X** is linearly dependent upon m unobservable random variables, $F_1, F_2 \ldots F_m$ called common factors and p additional sources of variation, $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_p$ called the idiosyncratic components or specific factors. This can be written in matrix notation as:

$$\underset{(p\times 1)}{\mathbf{X}-\boldsymbol{\mu}} = \underset{(p\times m)}{\mathbf{L}}\,\underset{(m\times 1)}{\mathbf{F}} + \underset{(p\times 1)}{\boldsymbol{\varepsilon}}$$

$$\boldsymbol{\Sigma} = \mathrm{Cov}(\mathbf{X}) = \mathrm{E}(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})' \qquad (1)$$

where the element l_{ij} in matrix **L** is the loading of variable i on factor j. We assume that:

$$\mathrm{E}(\mathbf{F}) = \underset{(m\times 1)}{\mathbf{0}},\ \mathrm{Cov}(\mathbf{F}) = \mathrm{E}[\mathbf{F}\mathbf{F}'] = \underset{(m\times m)}{\mathbf{I}}$$

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \underset{(p\times 1)}{\mathbf{0}},\ \mathrm{Cov}(\boldsymbol{\varepsilon}) = \mathrm{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\right] = \underset{(p\times p)}{\boldsymbol{\psi}} \qquad (2)$$

where **I** is the unit matrix, $\boldsymbol{\psi}$ is a diagonal matrix, and that **F** and $\boldsymbol{\varepsilon}$ are independent, so $\mathrm{Cov}(\boldsymbol{\varepsilon}, \mathbf{F}) = \mathrm{E}(\boldsymbol{\varepsilon}\mathbf{F}') = \underset{(p\times m)}{\mathbf{0}}$ and $\mathrm{Cov}(\mathbf{X}, \mathbf{F}) = \mathbf{L}$.

Using these assumptions, the covariance matrix of **X** can be written as:

$$\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}' + \boldsymbol{\psi} \qquad (3)$$

In this decomposition of the variance covariance of **X**, the proportion of the variance of the ith variable contributed by the m common factors is known as *communality* (ie the ith communality is the sum of squares of the loadings of the ith variable on the m common factors).

It is useful to note that the factor loadings are determined only up to an orthogonal transformation. Thus, if $\mathbf{T}$ is $m \times m$ and orthogonal, so that $\mathbf{TT}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$, then:

$$\underset{(p\times1)}{\mathbf{X}-\boldsymbol{\mu}} = \underset{(p\times m)}{\mathbf{L}}\,\underset{(m\times1)}{\mathbf{F}} + \underset{(p\times1)}{\boldsymbol{\varepsilon}} = \underset{(p\times m)}{\mathbf{L}}\,\mathbf{TT}'\,\underset{(m\times1)}{\mathbf{F}} + \underset{(p\times1)}{\boldsymbol{\varepsilon}} = \underset{(p\times m)}{\mathbf{L}^*}\,\underset{(m\times1)}{\mathbf{F}^*} + \underset{(p\times1)}{\boldsymbol{\varepsilon}} \tag{4}$$

where $\mathbf{LT} = \mathbf{L}^*$ and $\mathbf{T}'\mathbf{F} = \mathbf{F}^*$.

2.1.1 Estimation of factor loadings

Given a sample of observations on **X**, the aim is to estimate elements l_{ij} and ψ_i of the matrices $\mathbf{L}$ and $\boldsymbol{\psi}$ respectively. The available estimation procedures can be divided into two groups, namely common factor analysis and principal components analysis. The differences between the two lie in the consideration of the idiosyncratic components. The principal components method is based on a decomposition of the correlation matrix that does not take into account the random noise in the data. On the other hand, there are several common factor extraction methods to choose from. The main methods are maximum likelihood, principal axis, alpha factor analysis, image factor analysis, unweighted least square and generalized least square factor analysis methods (Gorsuch, 1983). However, the choice is not obvious as the theoretical arguments and empirical investigations are limited and controversial (Costello & Osborne, 2005). Many common factor extraction methods, including the maximum likelihood method, require multivariate normality and suffer from convergence problems. We propose to use image factor analysis, a restricted form of common factor analysis, for a number of reasons. First, image analysis accounts for the effect of the idiosyncratic components and, thus, is more applicable when there are different degrees of data accuracy across countries. Second, it is scale invariant and robust to the lack of multivariate normality. Third, it is computationally more efficient as it can be estimated without an iterative procedure, avoiding the problems related to convergence and the Heywood case (Acito & Anderson, 1980; Jöreskog, 1969). In practice, maximum likelihood and generalized least squares estimation of the common factor analysis models often yield estimates of the diagonal elements of the LL´ matrix that are greater than one, and this is known as the Heywood case (Gorsuch 1983; Harman 1976; Johnson & Wichern 2002). Theoretically, communalities cannot exceed unit length because all variables are standardized. In addition, both maximum likelihood and generalized least square estimation often face problems of improper solution or convergence (Ichikawa & Konishi 1995; Wansbeek & Meijer 2000).

Image analysis uses the multiple regression approach to identify the communality (Guttman, 1953; Harman, 1976; 1969; 1965; Jöreskog, 1962; 1965; 1969). The partial image of each variable in the sample is defined as its predicted value from the remaining p–1 variables – the commonness. The limit of the partial images of x_i as p becomes infinite is the total image of x_i in the *universe of content*. The squared multiple correlation for a particular variable represents the proportion of its total variance that depends on the remaining variables, providing an estimate of the communality, and the remainder of the unit variance is the proportion of the variance that is unique to the variable or unexplained by the remaining p–1 variables (that is the variance of the noise). Defining the vector $\mathbf{Y} = \mathbf{X} - \bar{\mathbf{X}}$, ie demeaned data, we can write:

$$\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{E} \tag{5}$$

where $\hat{\mathbf{Y}}$ is a vector of predictions of $\mathbf{Y}$ and $\mathbf{E}$ is a vector of errors of predictions, and they are orthogonal to each other. Then, the partial image of a variable, y_i , is given by:

$$\hat{y}_i = \sum_{k=1, k \neq i}^{p} \beta_{ik} \mathbf{y}_k \tag{6}$$

where β_{ik} are the estimated regression coefficients from regressing y_i on the set of remaining p–1 y variables. The anti-image is the residual from the multiple regression estimation.

In order to link these ideas to the factor model we need to introduce a few more concepts. Let **R** denote the $p \times p$ matrix of pairwise correlations amongst the variables. If all variables are standardized (that is to zero mean and variance unity), then $\mathbf{R} = \mathbf{\Sigma}$. In order to keep the new notation to a minimum, we will continue with using $\mathbf{\Sigma}$ and it is understood that it is the correlation matrix. Guttman (1953) establishes a decomposition of $\mathbf{\Sigma}$ that links image theory with factor analysis. The decomposition is:

$$\mathbf{\Sigma} = \mathbf{G} - \mathbf{\Gamma} + 2\mathbf{Z}^2 \tag{7}$$

where $\mathbf{G}$ and $\mathbf{\Gamma}$ are the covariance matrices of the images and anti-images, and $\mathbf{Z}^2$ is a diagonal matrix with elements equal to the variance of the anti-images of each variables (that is regression residual variances). The partition in (7) corresponds roughly to the covariance partition in equation (3).

Guttman (1956) shows that as $p \to \infty$ with the ratio $m / p \to 0$, $\mathbf{Z}^2 \to \boldsymbol{\psi}$, and thus, the estimation of the common factor model could be achieved through factorizing an approximation of the theoretical relationship, $\mathbf{\Sigma} - \boldsymbol{\psi}$ **:**

$$\mathbf{G} \to \mathbf{\Sigma} - \mathbf{Z}^2 \text{ as } p \to \infty \text{ if } \mathbf{Z}^2 = (diag\,\mathbf{\Sigma}^{-1})^{-1}$$

The corresponding image factor model is then:

$$\mathbf{\Sigma} = \mathbf{LL}' + \theta(diag(\mathbf{\Sigma}^{-1}))^{-1} \tag{8}$$

where θ is a positive scalar’. The early work of Guttman (1956) and Jöreskog (1969) shows that the corresponding elements of $\boldsymbol{\psi}$ and $(diag\,\mathbf{\Sigma}^{-1})^{-1}$ tend to the same limit, which gives support to the use of $(diag\,\mathbf{\Sigma}^{-1})^{-1}$ for $\boldsymbol{\psi}$. However, when p is finite, $\mathbf{\Sigma} - (diag\,\mathbf{\Sigma}^{-1})^{-1}$ is not necessarily positive definitive, thus image factor analysis equates $\boldsymbol{\psi}$ to a proportion of the corresponding elements of $(diag\,\boldsymbol{\Sigma}^{-1})^{-1}$:

$$\boldsymbol{\psi} = \theta(diag(\boldsymbol{\Sigma}^{-1}))^{-1} \tag{9}$$

Hayashi and Bentler (2000) study the relationship between the model in (3) and a restricted form of it,

$$\boldsymbol{\Sigma} = \mathbf{L}^{*}\mathbf{L}^{*\prime} + kI_{p} \tag{10}$$

where k is a positive scalar and I_p is an identity matrix. They derive conditions under which $\mathbf{L}$ and $\mathbf{L}^{*}$ are "close" (defined by a measure of closeness which is equal to 1 when they are identical). They show that the image factor model (8) and the restricted form in (10) are nested in (3). Thus, their work links to the original results by Guttman (1956) and Jöreskog (1969) and establish the conditions that link the image factor model in (8) to the factor model in (3).

2.1.2 Rotation of factors

In practice the element l_{ij} in $\mathbf{L}$ may not be readily interpretable, as they may not show a clear indication of how a given variable loads onto a particular factor. It is customary in empirical applications to rotate factors until a "simple structure" is achieved (details in Section 2.1.3). The available rotational methods can be divided into orthogonal and oblique rotational methods (Gorsuch, 1983). We define by $\mathbf{L}^{ot}$ and $\mathbf{L}^{ob}$ the rotated loading matrices obtained through an orthogonal and oblique transformation, respectively. Then, the orthogonally rotated loadings matrix is given by:

$$\mathbf{L}^{ot} = \mathbf{LT} \tag{11}$$

where $\mathbf{T}$ is an orthogonal transformation matrix such that $\mathbf{TT}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$. We note that the estimated covariance matrix remains unchanged, since $\mathbf{L}^{ot}\mathbf{L}^{ot\prime} + \hat{\boldsymbol{\psi}} = \mathbf{LTT}'\mathbf{L}' + \hat{\boldsymbol{\psi}} = \mathbf{LL}' + \hat{\boldsymbol{\psi}}$. These orthogonal rotated loadings are commonly obtained through a method known as *varimax*, which is used in this study. In contrast, an oblique rotation allows for some degree of correlation between factors. We compute the oblique loadings using the *promax* method developed by Hendrickson and White (1964) which is recognized as a more efficient method than other available alternatives. The oblique loadings of the *promax* method are a function of the orthogonal loadings as follows:

$$\mathbf{L}^{ob} = (\mathbf{L}^{ot\prime}\mathbf{L}^{ot})^{-1}\mathbf{L}^{ot\prime}\mathbf{P} \qquad (12)$$

where $\mathbf{P}$ is known as the pattern matrix and it has elements $\frac{\left|l_{ij}^{\kappa+1}\right|}{l_{ij}}$ (that is each element of $\mathbf{P}$ is the power of the corresponding element in the raw column normalized orthogonal matrix), $\kappa > 1$ and its value determines the degree of correlation amongst the factors (further discussion in Section 2.1.5). The term "pattern" initially appeared in the literature to refer to factor patterns, both orthogonal and oblique patterns. In an oblique rotation there are two matrixes: one is called the structure matrix which contains the correlations between the variables and the factors (these correlations may be inflated because some of the variance in a factor may not be unique to it); and the other is called the pattern matrix and it contains the unique correlations between variables and factors.

2.1.3 Computing standard errors for factor loadings and the selection of significant variables

Factor analysis meets the criteria defined by Thurstone (1947) to deliver a simple structure (Gorsuch, 1983; Reyment & Jöreskog, 1993) in that, first, it is a data reduction technique;

second, it identifies "salient loadings" of individual variables on factors; and third, through rotation, the factors are more interpretable. However, it is necessary to make a number of statistical decisions during the analysis. The common practice in the literature is to use rules of thumb or ad-hoc criteria. Moreover, there is substantial evidence that both under- and over-estimation of the number of factors can result in poor factor loading patterns and interpretation (Fabrigar, et al., 1999; Fava, & Velicer 1992; Velcier & Jackson, 1990). Next, the term "salient loading" in the literature is similar to significant loading, and some criteria are required to decide whether a variable's loading on each factor is significant. In most studies, a variety of rules of thumb, such as a loading above 0.3 or 0.5 is used to decide whether a loading is salient. From a statistical perspective, though a loading around 0.7 is high in magnitude, if it is not statistically significant, it cannot be considered to be a salient loading. Finally, the rotation method also plays a critical role as a simple structure is commonly obtained through the application of a rotation technique until the resulting factors are interpretable. There are numerous studies on various types of rotations and their properties; however, none of them provide criteria to choose the method of analytical rotation, or the degree of correlation to be allowed in a rotational method (Cudeck & O'Dell, 1994).

In this study we compute standard errors for the estimated loadings in order to decide on the significance of the factor loadings, as well as the number of the factors to be retained and the type of rotational method. There have been attempts in the literature to compute the standard errors of factor loadings in order to determine their level of significance (Chatterjee, 1984; Cudeck & O'Dell, 1994; 1974; Jennrich, 1973; Jennrich &Clarkson, 1980); however their use has not been widespread. Since most of the available methods to estimate standard errors are not suitable for variables that are not normally distributed, we implement a Jackknife procedure which requires re-estimation of the factor model *N* times (*N* is the number of observations in the data set, ie countries in our case), each time dropping one of the observations (in our case one of the countries). The square root of the sum of the variability of the estimates provides the standard errors (*SE*) of each estimated factor loading l_{ij} . that is;

$$SE = \sqrt{\left[(N-1)/N\right] \sum_{t=1}^{N} \left[l_{ijn} - (1/N)\sum_{s=1}^{N} l_{ijs} \right]^2} \qquad (13)$$

The corresponding significance level is determined using Bonferroni critical points ($\alpha^* = \alpha / d_u$, $\alpha = 0.05$, and d_u is the degrees of freedom; for the orthogonal rotation d_u = *pm*-*m*(*m*-1)/2 where *p* is the number of variables, and *m* is the number of factors; for oblique rotation d_u = *pm*-*m*(*m*-1). Bonferroni critical points are used because we wish to test the significance of the loadings of a particular variable across *m* factors. Therefore, all loadings are to be examined. To consider all loadings, the problem is handled as a special case of the joint probability that all the *pm* parameters are significant in a simultaneous system. The reader is referred to Cudeck and O'Dell (1994) for details.

The correlation across factors can be changed through an appropriate choice of the kappa value, that is κ in equation (12). However, there are no specific guidelines to select the extent of correlation. We use the standard error estimates to decide the appropriate kappa value by evaluating the significance of the variables over values of kappa. We have found (empirically) that within an interval of values of kappa the significance of the variables does not change although the correlation among the factors does (in our case within the low to moderate level of correlation). We label this the "stable region" and chose the value of kappa within this region so that the correlation among factors is the lowest. When the kappa value is chosen to be one, the final sets of factors are orthogonal.

2.1.4 Summary of the proposed procedure

The proposed procedure is summarized schematically in Figure 1. An initial number, m^*, of factors must be chosen as a first approximation to the factor model. The choice of m^* could be arbitrary; however, it is important to choose a relatively large number of factors so that the likelihood of omitting a significant factor is minimized. Methods such as choosing the number of factors to be those with eigenvalues greater than one or using the Scree Plot can be used to find m^*. Standard errors for loadings are then computed for the initial model with m^* factors. A factor is eliminated from the initial set if no variable loading on that factor is statistically significant. Once the number of significant factors, $\hat{m}$, has been determined the significance of individual variables is evaluated. By construction these factors are orthogonal. A variable is dropped from the initial set of p variables if it is found not to be significantly loaded on any of the $\hat{m}$ significant factors. The *promax* procedure is then used to obtain non-orthogonal (or oblique) rotations, where the correlation across factors changes by changing the value of kappa.

Estimated standard errors are used to decide the appropriate kappa value. As stated in the previous section, the significance of the variables is found to be robust within the stable region.

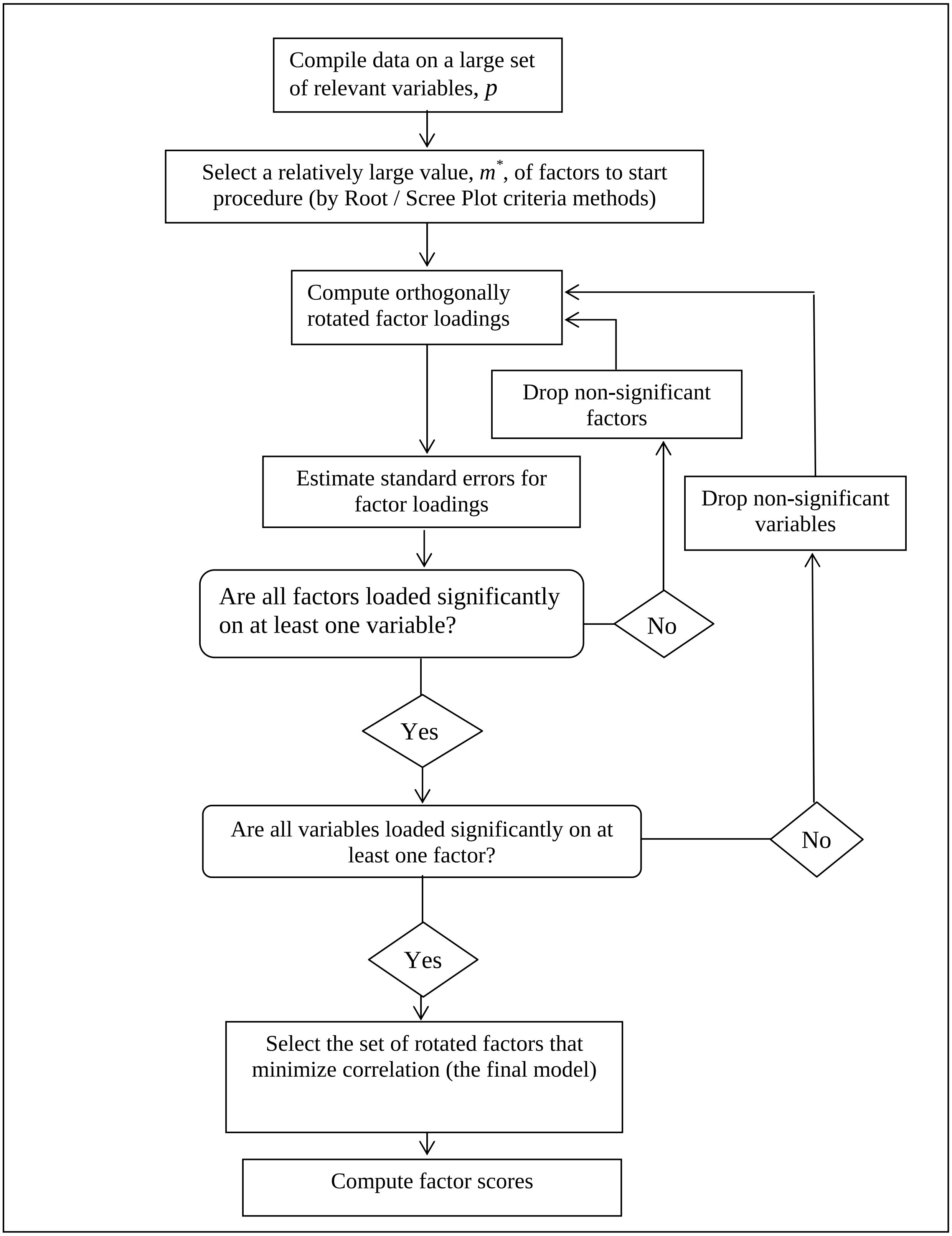

Figure 1: Flow chart of the estimation procedure

2.2 Construction of sub-indexes and a development index

In this section we present how the identified factors can be used to construct sub-indices of defining characteristics for each country in the sample and describe how these can be combined to form a development index.

2.2.1 Computation of factor scores and construction of structural characteristics indexes

Once significant factors have been extracted, factor scores for each factor can be estimated for individual countries. A weighted least square procedure is used to generate factor scores:

$$\hat{\mathbf{f}}_g^t = (\hat{\mathbf{L}}^{t\prime}\hat{\boldsymbol{\psi}}^{-1}\hat{\mathbf{L}}^t)^{-1}\hat{\mathbf{L}}^{t\prime}\hat{\boldsymbol{\psi}}^{-1}(\mathbf{x}_g - \overline{\mathbf{x}}),\ t = ot, ob \qquad (14)$$

where $\hat{\mathbf{f}}^t$ is the vector of estimated factor scores with elements $\hat{f}_{jg}$ for factor *j* of country *g*, and $\mathbf{x}_g$ is the vector of observable variables for the country.

The significant factors capture specific aspects of the socio-economic structure of the countries. The factor scores can be negative or positive in value. For the purpose of presentation of the structural characteristic indexes in Section 4, the raw factor scores are rescaled using the following formula to create *m* sub-indexes that allow comparison across countries:

$$S_{jg} = \left[\frac{f_{jg} - f_{j\min}}{f_{j\max} - f_{j\min}} \right] \qquad (15)$$

where $f_{j\min}$, $f_{j\max}$ and f_{jg} are respectively the minimum value, maximum value and individual scores g $(g = 1, 2, ...n)$ of each factor $(j = 1, 2, ...m)$. These rescaled factor scores are in the range of 0 to 1 and their empirical interpretation is presented in Sections 4.1-4.4.

2.2.2 Construction of a development index

The (raw) factor scores obtained from the procedure presented in Section 2.1 are used to construct a new development index. In order to avoid difficulties in handling negative factor scores, all the factor scores are transformed into positive values by using an exponential transformation. Since factors have different contributions to the total variance, the computed contribution to the total variance is employed to weight the individual factor to construct the index. We construct the index employing the geometric mean of the weighted values of the exponential value of the factor scores. The geometric mean is chosen as the scores are exponentiated which creates a left skewed distribution. Under these circumstances, the geometric mean is preferred over other methods of mean calculation as it avoids the bias towards large numbers (Feinstein 2001; Neuspiel 2004; Palaniswamy U. R. & Palaniswamy K. M. 2005):

$$DI_g = \prod_{j=1}^{m} [\exp(f_{jg})]^{w_j} \quad (16)$$

where DI_g is the index of development developed for each country g; w_j is the weight based on the proportion of the total variance due to the j^{th} factor, with $0 \le w_j \le 1$ and $\sum w_j = 1$.

This formula ensures that the constructed index is bounded below at zero. However, there is no upper bound. In order to make the resulting DI scores comparable across countries, we re-scale the DI scores into a scale of 0 to 1 as:

$$DI_g^* = DI_g / DI_{\max} \quad (17)$$

This index is employed to explore the level of development of countries and their differences in Section 4.5.

3. DATA

What indicators should be used to measure development depends on how we conceptualize development, that is, what we think development is about. In the 1950s there emerged a broad consensus that viewed development as a process of structural change (Chenery 1960; Kuznets, 1951; 1961; Rostow, 1960). The modern view of development has its roots in this lineage. It sees development as a dynamic process that involves major transformations in economic, institutional and social structures (Chenery, 1979; Syrquin, 1988; Todaro & Smith, 2003). The present study adopts this modern concept of "multidimensional structural change" as the working definition of development. In this perspective, structure shapes the characteristics of a country and, therefore, change in structure is an inseparable part of the development process.

The adopted working definition can avoid the debate over whether the means/causes/inputs or the ends/effects/outputs of development should be the focus of the measurement. Economic variables such as investment, technological, and institutional variables can be considered as "means" of development, while social indicators are commonly considered as "ends" (Adelman & Morris, 1972; Booysen, 2002). Some researchers argue that development should be based on ends, as they are the ultimate goal of development, while others argue that both means and ends are important components in setting the direction of policies (Morris, 1979). Yet in cases like education, ends and means cannot be separated. Since our working definition of development is not based on this restrictive view of development, we incorporate variables that capture both the means and ends of development. Our dataset includes 97 low, medium and high income countries, with 45 variables covering their economic, technological, institutional, and social characteristics. The list of countries included in the study along with a description of the variables is provided in the Appendix.

It should be mentioned at the outset that income is not included as a variable in the factor analysis for two reasons. From a conceptual point of view, development is defined as encompassing the various aspects of social, economic and institutional structures; therefore, the attention is on economic structure rather than income. From an analytical point of view, income can be used as an alternative indicator to validate the usefulness of the new index. This validation process is important because income is still viewed as a measure of development due to its high correlation to other economic, institutional and social variables.

Variables are selected to reflect various viewpoints and theories of development. To account for the structure of the economy, the shares of various sectors in the total GDP are included. Though technology plays a vital role in the growth process, direct measures of technology are hard to obtain. As a result, technology is indirectly assessed through measures of knowledge output and technological infrastructure. Knowledge output is measured using the volume of scientific and technical journal articles published. Technological infrastructure is considered to be important for knowledge diffusion and is measured by the availability of computers, internet usage, telecommunication facilities, and the use of mass media such as newspapers, radio and television. In addition, data on the number of vehicles and cars are also collected. Though these variables do not measure the level of technology directly, they provide an overview of transport infrastructure available for citizens. Since technology contributes to economic growth, we expect these measures to have a positive relationship with the level of development.

Physical and human capital play a significant role in standard growth theory. The accumulation of physical capital can be examined through the rate of saving, domestic and foreign investments as well as capital formation as a percentage of GDP. Human capital is a broader concept and is more difficult to define and measure. To represent human capital, education, health, and migration measures are included. School enrolments at primary, secondary as well as at tertiary educational levels are included as education measures. The ratio of pupils to teachers, which is considered as a quality measure, is also included. It is expected that all these measures to be positively associated with the level of development, with the exception that the ratio of pupils to teachers in primary education is expected to have a negative sign. In addition, the number of physicians, life expectancy of individuals, as well as life expectancy of females are included as health measures in the analysis. Fertility rate is also taken into account because some theories suggest a decrease in fertility rates during the development process (Chenery, 1979; Galor & Weil, 2000). Moreover, the age dependency ratio, population between age 14-64 years and labor force between ages 10-14 years are considered because they provide useful demographic information. Migrant stock as a percentage of total population, net migration and the percentage of urban population are also included.

Institutions have received a lot of attention in the development debate in recent years (Rodrik, 2003). This study includes several governance indicators published by the World Bank. The indicators cover six dimensions of institutional development, including voice and accountability, political stability, government effectiveness, regulatory burden, rule of law, and control of corruption (Kaufmann, et al., 2003). These institutional measures are expected to have a positive relationship with the level of development.

In addition, Gini coefficient is included since income distribution and inequality are expected to undergo changes during the development process. To represent cultural diversity, a religious diversity index and an ethnic diversity index are constructed by the authors using the Hirschman diversity index (Massell, 1970) (Details of the construction of this variable are provided in the Appendix, Table A.2.). Lastly, in recent years the AIDS epidemic has become a major concern in developing countries, therefore the incidence of AIDS is also included in our analysis.

We constructed averages for all the variables over the 10 year period of 1995 to 2004. This helps smooth the annual fluctuations of flow variables and minimize the gaps in data for several countries. If the data are not available for this time period, we collect data from the year nearest to the period as alternatives (for example 1990-1995). The factor analysis is conducted on standardised variables and, therefore, the correlation matrix (see Section 2) is used in the computations.

4. RESULTS

Seven factors are extracted initially to start the procedure. These factors are those that have eigenvalues greater than one. Standard errors are computed through a Jackknife re-sampling. Re-sampling can potentially create a few problems. Factors can be identified up to an orthogonal transformation (see equation (4)). Factor loadings of variables on some factors may change in sign during the re-sampling as specific observations (countries) might be more influential. This issue was raised in the literature as a problem for techniques that involve re-sampling (Clarkson, 1979; Pennell, 1972). Additionally, the ordering of factors can also change during the re-sampling process. This is known as the permutation problem (this typically occurs when the computed eigenvalues of two factors are very close and one is not consistently larger than the other across the re-sampling). In our case, the seven initially selected factors showed instability and the estimated loadings were not significant for the factors with smaller eigenvalues. As indicated in Figure 1, we proceed by lowering the number of factors by one each time and re-computing the factor loadings and standard errors until we arrive at a set of factors that have significant loadings and do not permute during re-sampling. For our data this is obtained when the number of factors drops to four. The four factors account for over 70 per cent of the overall variation. The first three factors explain respectively 45, 39, and 10 per cent of the total variance accounted for by the four factors, and the remaining factor explains only six per cent of the variation. The percentage of variation explained by each factor is expressed relative to the number of eigenvalues to convert into a percentage of variation:

$$V_i = \frac{\text{\% of variation of factor } i \text{ of oblique rotation}}{\text{number of eigenvalues}}$$

$$\text{Cumulative variation of factor } i = CV_i = \frac{V_i}{\sum V_i}$$

$$\text{and } 0 < CV_i < 1$$

Therefore, it is clear that more than half of the variation is accounted for by the first three factors.

Forty two out of the original 45 variables have a statistically significant loading on *at least* one factor (the Bonferroni critical points for m= 4, p=45 and varimax orthogonal rotation at the 5 and 10 percent significance level are given by: $\alpha^*_{0.05} = 0.00029$ is 3.70, $\alpha^*_{0.10} = 0.00057$ is 3.51 as the degrees of freedom are 174. For p=42, the Bonferroni critical points for $\alpha^*_{0.05} = 0.00031 =$ 3.69, $\alpha^*_{0.10} = 0.000617 =$ 3.49 as the degrees of freedom reduces to 162). Three variables do not significantly load on any of the factors, namely, foreign direct investment, gross primary enrolment ratio and industrial value added and therefore are dropped from the analysis. The loading of variables on the four factors are shown in Table 1 and significant loadings are in boldface. Inspection of the loadings in Table 1 makes it clear that if a typical rule of thumb with a cutoff point of 0.3 or 0.5 were used to determine "salient loadings", a number of variables would not have been considered as significant loadings on some of the factors. For instance, the loadings of birth rate, final consumption expenditures, life expectancy at birth and pupil-teacher ratio (primary education) have statistically significant loadings for Factor 1 that are below 0.3. This shows that using inferential decisions can make a significant difference to the results of factor analysis and on the index construction compared to the decisions on significant loading made using rules of thumb. In the next section we discussed in detail each factor and those variables that are significantly loaded on each of them.

The value of kappa, which determines the non-oblique degree of rotation and therefore the degree of correlation among factors, is found to be 1.6. This results in a correlation between the first two factors of 0.3 which is statistically significant. Factors III and IV are not significantly correlated. However, Factors II and IV are significantly correlated.

We now turn to a brief discussion of each of the four factors constructed through oblique rotation.

4.1 Factor I: Index of the level of technology and institutional development

The indicators that load significantly on Factor I, which explains the largest share of the variation of the variables, are mostly related to technology and institutions. The variables that are found to load significantly are: internet users, the availability of personal computers, daily newspapers, radios and television sets, fixed line and mobile phones, gross secondary and tertiary school enrolment rates, migration measures including share of urbanization, the share of high technology exports to the total output, and all the six governance measures (see Appendix A for a description of these measures). All these indicators are positively related to the factor. The Gini coefficient is loaded negatively on the factor, reflecting the importance of income distribution in characterizing development. In addition, the share of value added of the service sector is loaded positively while the ratio of industrial value added to service value added is loaded negatively. Overall, a higher score on Factor I indicates advancement in technological capabilities and the presence of quality institutions. Accordingly, Factor I is named as “the level of technological and institutional development”.

Figure 2 is a scatter plot of the rescaled scores for Factor I and income per capita which is a widely used, but crude, measure of development. The income data in this paper are real per capita income in 1995 PPP international dollars. A positive trend in the factor score for all countries with income per capita above $5000 is observed. However, the relationship becomes ambiguous when incomes are below that level. This implies that once a threshold level of development has been reached, technological know-how and institutional quality become crucial for reaching higher levels of development. The finding is also in line with the argument of Rodrik (2003) that, during the early stages of the transition countries do not need the best technology and institutions, instead, gradual changes in institutional structure should be introduced with the stage of development. We take the view that Factor I can be used as an indicator of the capabilities of countries to advance at the upper end of the development ladder. The loading of other variables in Factor I also indicate that at this upper end, higher development is associated with urbanization, rising income equality, expansion of the service sector at the expense of industrial sector, and higher migration levels. These findings are also broadly consistent with expectations (Chenery, 1979; Echevarria, 1997; Kuznets, 1955; Syrquin, 1988).

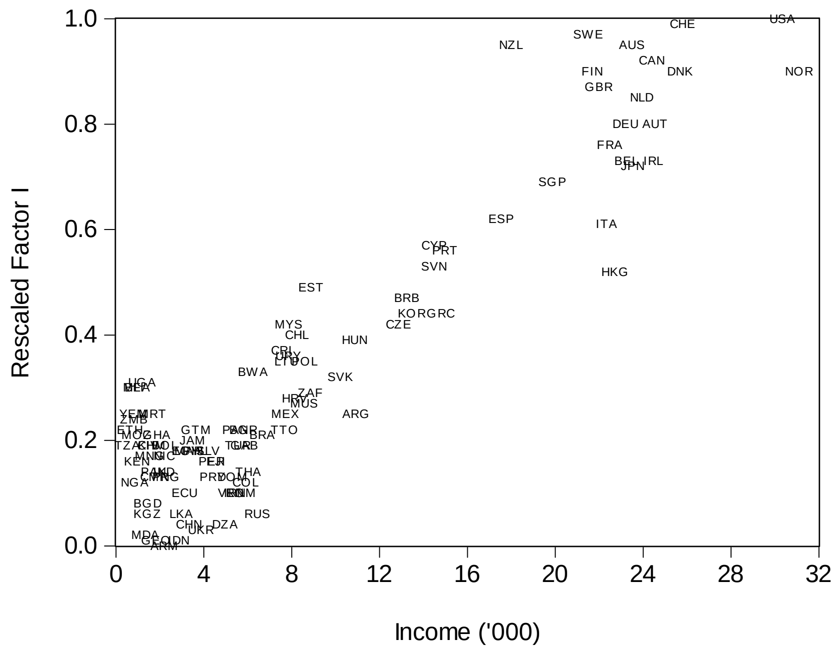

Figure 2: Income per capita and rescaled scores of Factor I

Table 2 reports the rescaled scores of each of the four factors and income per capita of the 97 countries, and countries are ranked accordingly. For Factor I, the United States has the highest score, due to its strong technological capability, followed by Switzerland which performs best in terms of institutional quality. Not surprisingly, countries that are ranked at the top are all OECD countries. However, Japan, a technologically advanced country, is only ranked the 17th with a rescaled score of 0.72. This is because the country scores low on government efficiency, regulatory quality, and control of corruption. Those in the bottom tier are expectedly low or lower middle income countries. Also, some countries with relatively high income such as South Africa (rescaled score = 0.29), Mauritius (0.27) and Argentina (0.25) also score fairly low, suggesting that they have low institutional and technological development compared to countries with similar income levels like Hungary (0.39) and Slovenia (0.53).

4.2 Factor II: Index of the level of basic development

Factor II contains both negatively and positively loaded indicators (see Table 1). Negatively loaded indicators include the ratio of agriculture to industry valued added, the age dependency ratio, birth rate, fertility rate, child labour force and the ratio of pupil to teachers in primary education. The values of these indicators, except the age dependency ratio, are typically high in developing countries and, therefore, load negatively on this factor. On the other hand, the following indicators are positively loaded to the factor: the share of urban population, the percentage of population aged 15-64, life expectancy of females at birth, life expectancy of the total population, the ethnic diversity index and the number of physicians. From Table 1 it can be seen that some variables, such as birth rate, daily newspapers, and secondary and tertiary school enrolment rates, load on both Factors I and II but typically with different magnitudes. Many of the indicators that loaded heavily on Factor II like life expectancy, the number of physicians, and the size of child labour force are commonly seen as a barometer of whether a society has the very basic level of resources and capability for survival. In fact, some of these indicators, including life expectancy and educational enrolments, have a central role in the construction of prominent development indicators such as PQLI and the HDI (Morris, 1979; 1996; UNDP, various years). Therefore, Factor II is labelled as "the level of basic development".

Figure 3 is a scatter plot of the rescaled scores of Factor II against income per capita. The two variables exhibit an approximately log linear relationship. Between the level of income per capita of $2500 and $10 000 a scattered relationship is observed, while for income per capita above $10 000 there exists a stagnant relationship. A major portion of the variation in this factor is attributable to countries at the low income level since middle and high income countries have reached or surpassed the minimum levels of the indicators that determine this factor, suggesting that the measure provided by Factor II is more useful in differentiating countries at the lower end of the development ladder. This finding is also reflected in Table 2 in that nearly 80 percent of the sample countries have a score of Factor II equal to or above 0.5, indicating that most of the countries have reached certain basic levels of development. Countries that ranked at the top end of the table are mostly emerging economies. However, the scores of OECD countries are not far behind. This means that the ranking of countries only becomes truly indicative of lack of development at the bottom end of the table. Countries that ranked the lowest are all sub-Sahara African countries, such as Burkina Faso, Mali, and Uganda**.** These countries are not only the poorest but also with some of the lowest levels of life expectancy.

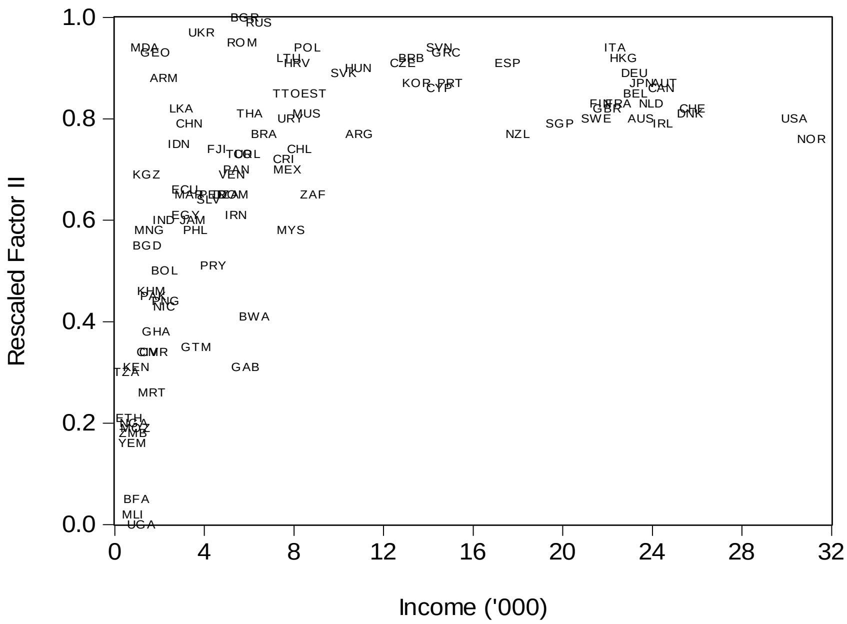

Figure 3: Income per capita and rescaled scores of Factor II

4.3 The relationship between Factor I and Factor II

The factor scores in Table 2 show that only 24 out of the 97 countries have a rescaled score of Factor I equal to or higher than 0.5, compared to 76 countries for Factor II. That is, while over 80 percent of the countries in the sample have already been elevated above the bottom end of the development ladder, only 25 percent have reached the top end. This means that over half of the countries have the “middle status”. Furthermore, the scatter plots of Factors I and II against income per capita are in great contrast. Figure 2 shows that Factor I is more useful in differentiating countries at the higher end of the development ladder, whereas Figure 3 shows that Factor II is more useful in differentiating countries at the lower end of the development ladder.

An important implication of this contrast is that Factors I and II could jointly provide information about a certain threshold level, below which development is more about survival, and above which development is associated with advancement of capabilities. If this is the case, the relative scores of Factor II and Factor I might provide an indication of the transformation process from the basic level to the advanced level of development during the development process. To see this possibility, we plot the raw scores of Factor II against those of Factor I in Figure 4. The raw rather rescaled scores are used because as the raw scores can be positive or negative, the scatter plot is naturally divided into four quadrants according to the signs of Factors I and II respectively.

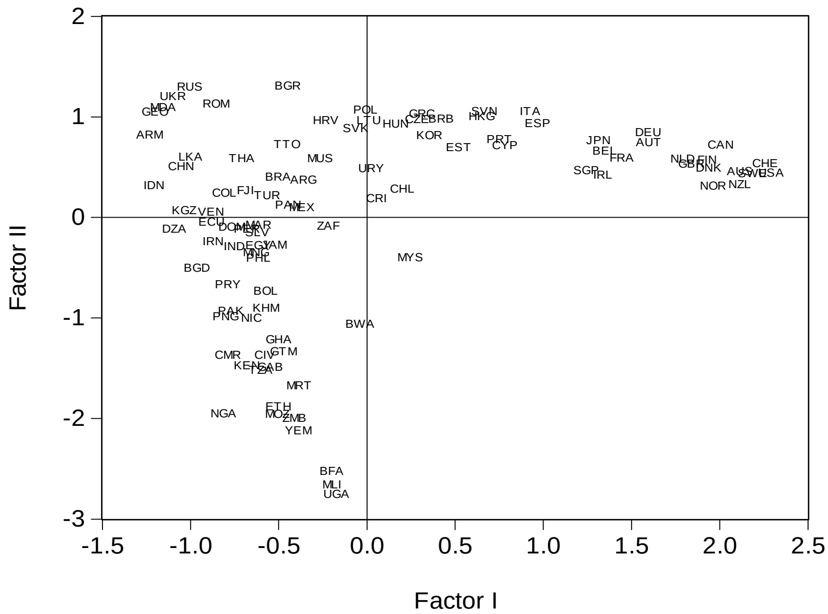

Figure 4: The Relationship between Factor I and II (Factor scores are not rescaled)

Since Factor I represents the advancement of capabilities while Factor II represents the survival capabilities, it is natural to expect that the development path for most countries will be from the lower-left quadrant to the upper-left quadrant and eventually to the upper-right one. Therefore, it is not surprising to see that only one country occupying the lower-right quadrant.

Countries in the lower-left quadrant have negative scores for both Factors I and II. The majority of these countries are low income, sub-Sahara African countries. But a few African countries with relatively high income per capita such as Botswana and South Africa, and some Latin American countries such as Nicaragua and Peru are also in this quadrant. Real per capita incomes of countries in this group range from \$476 to \$8772, with an average equal to \$2708. A puzzling observation is that within this quadrant there appears to be a negative relationship between the two factor scores. Whether this reflects some sort of imbalance in resource allocation between different social and economic needs in this group of countries deserves further investigation.

Countries in the upper-left quadrant have negative scores for Factor I but positive scores for Factor II. These countries are mostly the transitional or emerging economies like Armenia, Indonesia, Moldova, Ukraine, China, Kyrgyzstan, Russia, and Sri Lanka. Countries in this quadrant have achieved the basic level of development but are yet to achieve the technological and institutional capabilities for advancing to a higher stage of development. Their incomes range from $1269 to $10 873, with an average equal to $5391.

Countries in the upper-right quadrant have positive scores for both factors, and they are mostly OECD countries. These countries have long passed the threshold of basic development. Therefore, development for these countries is more about enhancement of life and living, and expanding the choices available to individuals. At this level of development, further advancement is closely associated with technological progress and a good institutional environment (Aghion 2003; Rodrik, 2003).

Malaysia is the only country in the lower-right quadrant, indicating its unique characteristics. It has a positive score for Factor I and a negative score for Factor II, implying that it has a relatively low level of basic development compared to other countries with a similar income level. Although the country registers a positive score for Factor I, the score value is not much larger than zero.

4.4 Factor III: Index of capital accumulation

Factor III loads significantly on four variables, namely, final consumption, gross domestic savings, gross capital formation and gross fixed capital formation. All the four variables except consumption are positively loaded on Factor III. Based on the findings, Factor III is named as “capital accumulation”. The top two countries, China and Singapore, have a rescaled factor score substantially higher than the rest due to their impressive saving and capital formation rates.

Factor III does not show any distinct relationship with the level of income (Figure 5). There is considerable development literature that highlights the importance of saving and capital accumulation for growth (Caballero, et al., 2005; Rogers, 2003). However, Denison (1964; 1980) argues that capital has sometimes contributed to growth differences in places but falls short of fully explaining international differences in growth rates. Furthermore, even if saving and capital accumulation is important to growth, Factor III is still only a leading indicator of future development levels, but not an indicator of the current stage of development (this is our interpretation, we are not aware of any theoretical literature to support this observation).

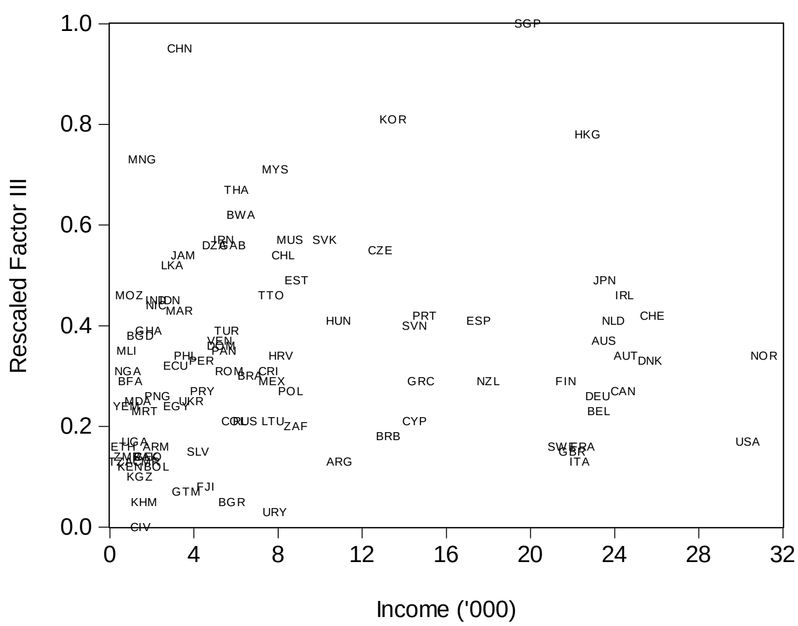

Figure 5: Income per capita and rescaled scores of Factor III

4.5 Factor IV: Index of HIV incidence and its impact on life quality

Factor IV loads negatively on HIV incidence and positively on total life expectancy and female life expectancy. The opposite signs of the loadings are consistent with the observed negative relationship between these variables. In addition, the religious diversity index is loaded positively onto this factor. Based on its loading composition, Factor IV is labelled "the incidence of HIV and its impact on life quality".

The two countries that have the lowest rescaled scores for Factor IV are Botswana and South Africa, both having the highest HIV incidence amongst the sample countries. Other countries with low scores also register a high level of HIV incidence and a low level of life expectancy, such as Tanzania, Cameroon, Cote d'Ivoire, Nigeria, Ethiopia, and Zambia. At first glance, no clear relationship appears to emerge between Factor IV and the level of income of countries (Figure 6); but the variation in the scores of Factor IV changes with income. When income exceeds $10 000, the volatility in the scores appears to have reach and stayed at its minimum level. However, most recently the United Nations Programme on HIV/AIDS and World Health Organization (WHO) have substantially revised the HIV/AIDS prevalence figures worldwide. The new estimate of the number of person living with HIV in 2007 was 16 percent lower than the 2006 estimate, and 70 percent of the reduction are due to changes in five African countries plus India (UNAIDS & WHO, 2007). Given that the dataset covers pre 2007 data, we need to be cautious about the findings of Factor IV.

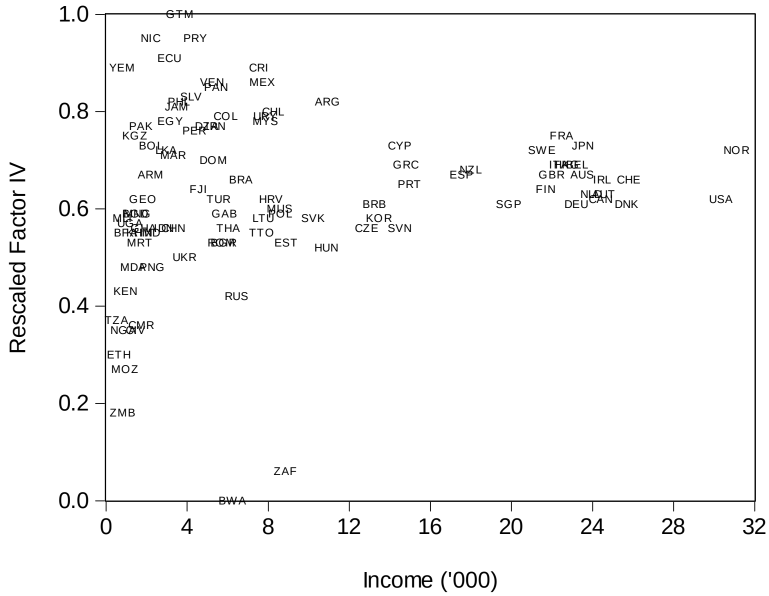

Figure 6: Income per capita and rescaled scores of Factor IV

4.6 A new development index

The extracted four factors describe and measure different aspects of development. However, not every one of them is useful in constructing a development index. In particular, there is no *a priori* indication of whether a higher score for Factor III implies a higher or lower stage of development and, in fact, the factor shows no pattern against income per capita. Next, though Factor IV captures some important aspects of development in relation to HIV incidence, it fails to displace an expected positive relationship with income per capita. The data issue mentioned above also weights against the inclusion of Factor IV in the final development index construction. For these reasons, we compute the development index using only scores for Factor I and II. Further, the first two factors account for 84 percent of the total variation explained by the four factors and nearly 60 percent of the total variation of the 42 variables in the dataset.

The numerical values of the development index (DI) are reported in Table 2. The top five countries by DI ranking for the period 1990-2000 were Switzerland, USA, Canada, Sweden, Australia, and Finland, while the bottom five countries by DI ranking were Yemen, Burkina Faso, Mali, Nigeria, and Uganda. To examine how useful the DI is as a development measure, we use the most popular development measure – income per capita – as an external validator. Figure 7 shows a scatter plot of the DI against income per capita. Despite the overall linear relationship, there is substantial variation in the DI for a given income level, especially at the top and bottom ends of the income scale. This is due to the fact that Factor I has particularly good discriminating power amongst the high income countries while Factor II is the same amongst low income countries, so jointly the two indices can successfully separate countries that cluster at the two ends of the income spectrum.

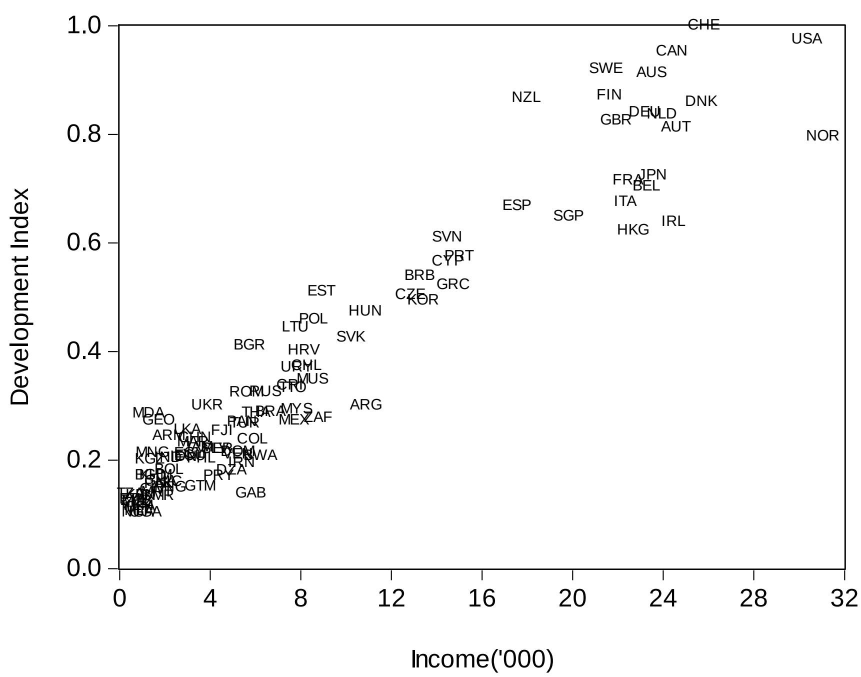

Figure 7: Income per capita and Development Index

To further examine the usefulness of the DI as a development measure, we compare it with another popular development indicator, the HDI in Figure 8. The figure displays an exponential relationship between the two indexes. In particular, it seems that the HDI has a higher differentiating power for the least developed countries while the DI has more power for the more developed countries. This is because the HDI concentrates only on a limited number of development indicators, namely the literacy rate, school enrolment rates, life expectancy, and income. Since high income countries mostly have reached the “satiation levels” of these indicators except for income, the differentiating power of the HDI diminishes quickly at the upper end of the development ladder. Although the DI also includes school enrolment rates and life expectancy, its inclusion of many other indicators, especially those related to technological capability and institution quality, means that it has a strong differentiating power even amongst the high income countries.

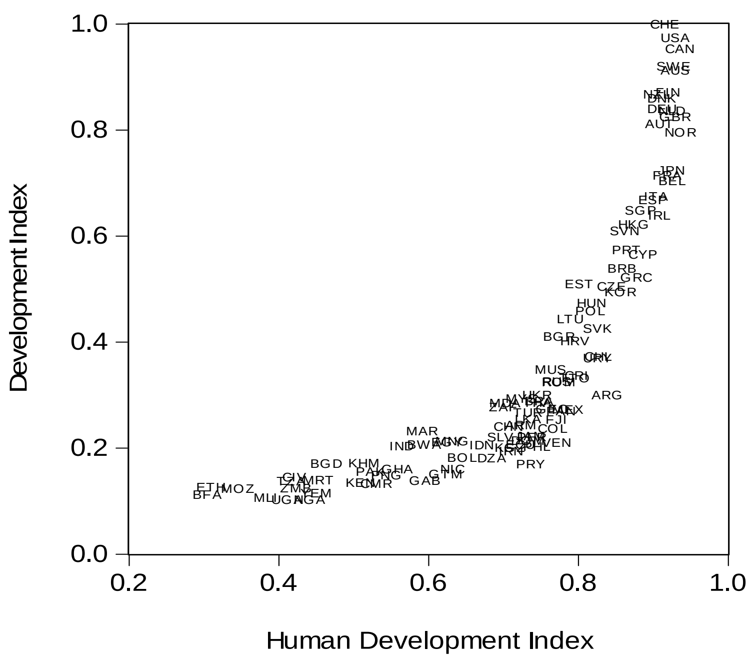

Figure 8: Human Development Index and Development Index

In terms of ranking, although the three measures are broadly consistent with others, there are also remarkable differences for individual countries (Table 2). For instance, Bulgaria and Gabon have very similar income per capita at about $5700 and therefore have adjacent income rankings (43 and 41, respectively), but their DI rankings are more than 50 apart (32 and 79, respectively). Gabon's low ranking score is due to its relatively low level of basic standards of living captured by Factor II. Compared to income per capita, the HDI is more capable of identifying this difference between Bulgaria and Gabon. Nevertheless, the HDI ranking gap between the two countries is still far smaller than that of the DI.

Income and the HDI noticeably neglect institutional quality. Argentina is a case in point. Having been one of the wealthiest countries in the world 100 years ago because of its rich natural resources and highly literate population, Argentina currently is still an upper-middle income country. Moreover, due to continuous effort on welfare provision, its human development is relatively high and that is captured by the HDI. As a result Argentina has a high rank of the 29th on either per capita income or the HDI, slightly ahead of Hungary. Despite this, Argentina's institutional quality measured by government efficiency, political stability, rule of law and corruption has stayed low, posing a major stumbling block on its efforts to achieve a higher development standard. As a result, the DI ranks Argentina at the 44th, 14 places behind Hungary.

Lastly, China and India have been undergoing a tremendous economic transformation in recent decades. Numerous studies envisage India and China to be prominent world powers by the end of this century (Drezner, 2007). However, China only shows a moderate difference between its rankings on the DI (57th) and income per capita (66th). This is because China's extraordinary economic transformation is facilitated by a high level of capital accumulation which is mostly captured by Factor III. On the other hand, services are playing an increasingly important role in India's economy and the DI does capture the service sector's development impacts. However, the country's strong performance in the economic sphere is counterbalanced by its poor performance in the social sphere. For instance, India has a life expectancy of 60 years, compared to China's

70 years. The country also has low values for telephone availability, a low secondary school enrolment rates, coupled with a high birth rate and a large presence of child labour. As a result, India is ranked below China by the DI at a position of 70th.

5. CONCLUSION

The construction and use of indices is increasingly popular across social sciences, business and economics disciplines. Factor analysis is a data reduction technique and it remains one of the most popular techniques used in the construction of indexes. The key methodological contribution of the paper is to incorporate standard errors, computed using the Jackknife technique, into the decision making process of factor analysis, which hitherto is mostly based on rules of thumb, that is arbitrary criteria. These standard errors aid the identification of the number of factors to be retained as well as variables with statistically significant loadings.

To illustrate the improved procedure, we apply the technique on a data set covering a total of 45 development indicators over a ten year period of 1995-2004 for 97 countries. Three out of the 45 original variables had insignificant loadings on all the factors and, therefore, were dropped from the analysis. The remaining indicators loaded on four factors at different levels of significance and magnitude. We found that the four selected significant factors explain over 70 per cent of inter-country variation exhibited in the data on the 42 variables. A re-scaled form of these factors can be used as a sub-index as each of these factors related to different dimensions of development and thus can provide useful insights on the structure of particular countries.

Out of the four factors extracted, the first two factors account for more than 80 per cent of the total variation explained by the four factors together and also are the most important in terms of explaining the inter-country development patterns. These two factors depict two crucial, yet different aspects of development. The first factor captures the role of technology and institutional quality while the second factor reflects the information related to the basic level of development. The contrast of Factors I and II signifies that the notion of development is not only multidimensional, but also changing with the stage of development. As these two factors emphasize different aspects of development, they were subsequently combined to form a development index (DI) to study the development ranking of countries. A comparison of the new DI to income per capita and the HDI strongly established that that the DI is a valid and relevant development index, and that it has better discriminating power across different parts of the development ladder.

Our empirical investigation has also revealed an important third factor which largely provides a measure of the general saving and capital formation efforts of countries. Factor scores for the countries may be used as a leading indicator of future development performance. Not surprisingly, China has the highest score for this factor which may be the main reason for the extraordinary growth performance of this country over the last ten years.

Overall, the paper has developed a procedure to incorporate statistical inference into factor analysis, eliminating some of the arbitrariness of its implementation in many previous studies. It also successfully demonstrated how the new procedure can be applied to construct a more powerful development index from a large set of diverse indicators.

Acknowledgment

The authors would like to acknowledge valuable discussions with Derek Headey and comments from the participants of the Development Workshop at Monash University, July, 2006.

References

Acito, Franklin, and Ronald D. Anderson (1980), "A Monte Carlo comparison of factor analytic method," *Journal of Marketing Research,* 17(2): 228-236.

Adelman, Irma, and Cynthia T. Morris (1972), "The measurement of institutional characteristics of nations: methodological considerations," *Journal of Development Studies,* 8(3): 111-135.

Adelman, Irma, and Cynthia T. Morris (1967), *Society politics & economic development: A quantitative approach,* Baltimore: The Johns Hopkins Press.

Aghion, Philippe (2006), On Institutions and Growth Institutions, ed. Theo S. Eicher & Cecilia García-Peñalosa Cambridge, Mass: MIT Press.

Aralica, Zoran, and Jelena, Budak (2004), "Institutional Development and Good Governance Assessments in Croatia: An Extended Focus on Corruption," *Zagreb International Review of Economics and Business*, 7 (Special Conference Issue): 17-34.

Booysen, Frederik, le R. (2002), "An overview and evaluation of composite indices of development," *Social Indicators Research,* 59(2): 115-151.

Caballero, Ricardo J. & Kevin, Cowan and Jonathan Kearns (2005)," Fear of Sudden Stops: Lessons from Australia and Chile," *Journal of Policy Reforms,* 8 (4): 313-354.

Cahill, Miles. B., and Nicolas, Sánchez (2001), "Using principal components to produce an economic and social development index: An application to Latin America and the US," *Atlantic Economic Journal* 29 (3): 311-329.

Chatterjee, Sangit (1984), "Variance estimation in Factor Analysis: An application of the Bootstrap," *British Journal of Mathematical and Statistical Psychology*, 37(2): 252-262.

Chenery, Hollis B. (1960), "Pattern of industrial Growth," *American Economic Review*, 50(4): 624-654.

Chenery, Hollis B. (1979), *Structural Change and Development Policy,* New York: Oxford University Press.

Clarkson, Douglas B. (1979), "Estimating the standard errors of rotated factor loadings by jackknifing," *Psychometrika* 44(3): 297-314.

Corielli, Francesco, and Massimiliano Marcellino (2006), "Factor based index tracking," *Journal of Banking & Finance,* 30 (8): 2215-2233.

Costello, Anna B. & Jason Osborne (2005), "Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis," *Practical Assessment Research and Evaluation,* 10 (7): 1-9.

Cudeck, Robert, and O'Dell, Lisa L. (1994), "Application of standard error estimates in unrestricted Factor Analysis: Significant tests for factor loadings and correlations," *Psychological Bulletin,* 115 (3): 475-487.

Denison, Edward F. (1964), "Capital theory and the rate of return," *American Economic Review,* 54 (5): 721-725.

Denison, Edward F. (1980), "The contribution of capital to economic Growth," *American Economic Review,* Papers and proceedings of the ninety-second annual meeting of the American economic association, 70 (2): 220-224.

Doessel, Darrel, P., and Rukmani Gounder, (1991),"International comparisons of levels of living and the Human Development Index: some empirical results," Department of Economics, The University of Queensland.

Drezner, Daniel, W. (2007), "The New World Order," *Foreign Affairs,* 86 (2): 34-46.

Echevarria, Cristina (1997), “Changes in sectoral composition associated with economic growth,” *International Economic Review* 38 (2): 431-452.
Fabrigar, Leandre R., Wegener, Duane T., MacCallum, Robert C., and Erin J. Strahan, (1999), "Evaluating the use of exploratory factor analysis in psychological research," *Psychological Methods*, 4 (3): 272-299.
Fava, Joseph L. and Velicer, Wayne F. (1992), "The effects of overextraction on factor and component analysis," *Multivariate Behavioral Research,* 27 (3): 387-415.
Galor, Oded, and David N. Weil. (2000), "Population, technology and growth: from Malthusian stagnation to the demographic transition and beyond," *American Economic Review* 90(4): 806-828.
Gorsuch, Richard L. (1983), *Factor Analysis,* 2nd ed. New Jersey: Lawrence Erlbaum Associates, Publishers.
Guttman, Louis (1956), "Best Possible" systematic estimates of communalities," *Psychometrika,* 21(3): 273-285.
Guttman, Louis (1953), "Image theory for the structure of quantitative variates," *Psychometrika* 18(4): 277-296.
Harman, Harry, H. (1976), *Modern factor analysis,* 3rd (revised) ed. Chicago and London: The University of Chicago Press.
Hayashi, Kentaro, and Peter M. Bentler (2000), "On the relations among regular, equal unique variances, and image factor analysis models," *Psychometrika,* 65(1): 59-72.
Hendrickson, Alan E., and Paul O. White (1964), "Promax: A quick method for rotation to Oblique simple structure," *The British Journal of Statistical Psychology* 17 (1): 65-70.
Jennrich, Robert I. (1974), "Simplified formulae for standard errors in Maximum Likelihood Factor Analysis," *British Journal of Mathematical and Statistical Psychology* 27(1): 122-131.
Jennrich, Robert I., (1973), "Standard errors for obliquely rotated factor loadings," *Psychometrika* 38(4): 593-604.
Jennrich, Robert I., and Douglas B. Clarkson (1980), "A feasible method for standard errors of estimates in Maximum Likelihood Factor Analysis," *Psychometrika,* 38(2): 571-580.
Johnson, Richard A., and Dean W. Wichern. (2002), *Applied multivariate statistical analysis,* New Jersey: Prentice Hall.
Jöreskog, Karl G. (1969), "Efficient estimation in image factor analysis," *Psychometrika,* 34 (1): 51-75.
Jöreskog, Karl G. (1965), *Image analysis,* Vol. 65-5. Research Bulletin, Princeton, N. J.: Educational Testing Service.
Jöreskog, Karl G. (1962), "On the statistical treatment of residuals in factor analysis," *Psychometrika* 27(4): 335-345.
Kaufmann, Daniel, Massimo Mastruzzi and Diego Zavaleta (2003), Sustained macroeconomic reforms, tepid Growth: A governance puzzle in Bolivia? ed. D. Rodrik. Princeton and Oxford. Princeton University Press.
Kuznets, Simon (1961), "Quantitative aspects of the economic growth of nations: Long term trends in capital formation proportions," *Economic Development and Structural Change,* 9(4): 1-124.
Kuznets, Simon (1951), "The state as a unit in the study of economic growth," *Journal of Economic History,* 11(1): 25-41.

Kuznets, Simon (1955), “Economic growth and income inequality” *American Economic Review* 45(1): 1-28.
Massell, Benton F. (1970), "Export instability and economic structure," *American Economic Review, Papers and Proceedings,* 60 (4): 618-630.
Mbaku, John M. (1997), "Inequality in income distribution and economic development: Evidence using alternative measures of development," *Journal of Economic Development,* 22 (2): 57-67.
McGranahan, Donald V., Eduardo, Pizarro and Claude, Richard (1985), *Measurement and analysis of socio-economic development,* Geneva: Research Institute for Social Development, United Nations.
McGranahan, , Donald . V., et al. (1972), *Contents and measurement of socio-economic development,* Staff Study. New York: United Nations Research Institute for Social Development.
Moran, Daniel D., Mathis, Wackernagel, Justin A. Kitzes, Steven, H. Goldfinger, and Aurélien, Boutaud (2008), "Measuring Sustainable Development-Nation by Nation," *Ecological Economics,* 64 (3): 470-74.
Morris, Morris D. (1996), *Measuring the changing condition of the world's poor: the Physical Quality of Life Index*, *1960-1990*, Working Paper no 23/24. Brown University.
Morris, Morris D., (1979), "Measuring the condition of the world's poor: the physical quality of life index," Pergamon Press, New York.
Pennell, Roger (1972), "Routinely computable confidence intervals for factor loadings using the 'jack-knife'," *British Journal of Mathematical and Statistical Psychology,* 25(1): 107-114.
Pitlik, Hans (2002),"The path of liberalization and economic growth," *Kyklos,* 55:1: 57-79.
Rayp, Glenn, and Nicolas Van De Sijpe (2007),"Measuring and Explaining Government Efficiency in Developing Countries," *Journal of Development Studies,* 43 (2): 360-81.
Reyment, Richard A, and Karl G. Jöreskog (1993), *Applied factor analysis in the natural science,* New York: Cambridge University Press.
Rocha, Bruno de P., and R. M. Sekkel (2006), "The Dynamic Factor Model: An Application to International Stock Market Integration," *Revista de Economia,* 13 (1): 91-116.
Rodrik, Dani (2003), *In search of prosperity: analytic narratives on economic growth*, New Jersey, Princeton University press.
Rogers, Mark (2003), "A survey of economic growth," *The Economic Record,* 79 (244): 112-135.
Rostow, Walt W. (1960), *The stages of economic growth: a non-communist manifesto*, Cambridge: University Press, London.
Syrquin, Moises, (1988) Patterns of structural change, ed. H. Chenery, and T. N. Srinivasan, vol. I. Amsterdam, Elsevier Science Publishers B.V.
Thurstone, Louis L. (1947), *Multiple factor analysis*. Chicago: University of Chicago Press.
Todaro, Michael P. and Stephen C. Smith (2003), *Economic development,* 8 ed. New York: Addison-Wesley.
UNAIDS and WHO (2007), *AIDs epidemic updates*, Geneva, World Health Organization.
UNDP (various years). *Human development report,* United Nations Development Programme.
Velicer, Wayne F. and Douglas N. Jackson(1990), "Component analysis versus common factor analysis: Some further observations," *Multivariate Behavioral Research,* 25 (1): 97-114.
Wansbeek Tom, and Erik, Meijer (2000), *Measurement error and latent variables in econometrics*. Amsterdam: North-Holland.

World Bank (2008), *Governance and Anti-corruption: Worldwide governance indicators,* 1996-2006.
UNDP (various years), *Human development report,* United Nations Development Programme.

Table 1: Factor loadings and test statistics for obliquely rotated factors (kappa value = 1.6)[a]

Variables	Factor				Test Statistics			
	1	2	3	4	1	2	3	4
Age dependency ratio (ratio to the working age pop.)	-0.11	**-0.92**	-0.18	0.06	-3.6	-32.1	-2.3	1.1
Birth rate, crude (per 1000 people)	**-0.17**	**-0.91**	-0.07	0.01	-4.5	-36.9	-2.0	0.2
Daily newspapers (per 1000 people)	**0.60**	**0.31**	0.10	0.01	9.3	4.8	1.5	0.2
Fertility rate (no. of births per women)	-0.09	**-0.93**	-0.12	-0.01	-2.8	-38.6	-2.7	-0.3
Final consumption expenditure, etc. (% GDP)	**-0.24**	-0.11	**-0.65**	0.08	-3.6	-1.1	-7.6	0.9
Fixed line & mobile phone subscribers (per 1000 people)	**0.79**	**0.34**	0.04	0.02	24.2	9.2	0.8	0.5
Gross capital formulation (% GDP)	-0.09	0.06	**0.89**	0.02	-1.8	1.1	11.0	0.4
Gross domestic savings (% GDP)	0.13	0.15	**0.74**	0.00	2.1	1.7	9.3	0.0
High-technology exports (% of manufactured exports)	**0.39**	0.01	0.29	0.13	3.7	0.2	2.0	1.4
Migration stock (per 1000 people)	**0.45**	0.15	0.13	-0.09	4.5	1.8	0.7	-1.3
Internet uses (per 1000 people)	**0.84**	**0.20**	0.13	0.02	18.0	5.8	1.9	0.7
Life expectancy at birth, female (years)	**0.23**	**0.67**	0.07	**0.46**	5.3	8.2	1.92	6.5
Life expectancy at birth, total (years)	**0.26**	**0.63**	0.10	**0.48**	5.7	7.1	2.4	6.9
Passenger cars (per 1000 people)	**0.77**	**0.36**	-0.16	-0.04	16.3	5.5	-1.6	-0.8
Personal computers (per 1000 people)	**0.89**	**0.12**	0.08	0.02	34.7	3.7	1.1	0.6
Physicians (per 1000 people)	0.21	**0.71**	-0.23	0.02	2.5	9.9	-2.6	0.4
Population ages 15-64 (% of total)	**0.14**	**0.91**	0.19	-0.08	3.9	27.9	2.4	-1.5
HIV incident (per 1000 people)	0.03	-0.28	0.04	**-0.64**	0.6	-1.9	0.2	-7.3
Pupil-teacher ratio (primary education)	**-0.22**	**-0.74**	0.01	-0.15	-4.7	-14.1	0.3	-2.6
Gross secondary school enrolment rate	**0.42**	**0.63**	-0.05	0.09	7.0	11.3	-0.9	1.8
Gross tertiary school enrolment rate	**0.51**	**0.52**	-0.09	0.07	7.7	8.2	-1.2	1.2
Secure Internet servers (per 1 million people)	**0.82**	0.02	-0.01	0.01	23.5	0.6	-0.2	0.3
Service sector value added (% GDP)	**0.53**	**0.38**	-0.11	0.16	7.8	4.8	-0.8	1.9
Urban population (% of total population)	**0.37**	**0.42**	0.05	0.18	4.9	4.7	0.5	2.8
Vehicles per (per 1000 people)	**0.77**	**0.36**	-0.14	-0.03	16.3	5.7	-1.4	-0.7
Television sets (per 1000 people)	**0.63**	**0.49**	-0.09	0.06	13.0	8.2	-1.7	1.1
Radio (per 1000 people)	**0.67**	**0.33**	-0.10	-0.01	11.5	4.9	-1.7	-0.1
Ratio of value added of Indus. to that of services	**-0.33**	-0.10	0.37	-0.12	-4.2	-0.8	2.4	-1.0
Ratio of value added of agric. to that of industry	-0.21	**-0.49**	-0.23	-0.17	-2.6	-6.3	-3.1	-1.6
Scientific & technical journal articles (per 1000 people)	**0.88**	0.11	-0.05	-0.01	41.0	3.1	-1.2	-0.4
Fixed capital formulation (% GDP)	0.03	0.00	**0.92**	0.01	0.7	0.1	13.1	0.1
Net migration (per 1000 people)	**0.55**	-0.17	0.22	-0.13	10.2	-1.5	1.6	-1.3
Labour force, children 10-14 (% age group)	-0.06	**-0.79**	-0.08	-0.17	-1.4	-14.7	-1.4	-2.2
Gini coefficient	**-0.31**	-0.35	0.08	0.09	-3.7	-3.3	0.7	0.5
Voice and accountability	**0.71**	**0.33**	-0.04	-0.04	14.5	4.7	-0.5	-0.6
Political stability	**0.71**	0.26	0.14	-0.07	15.0	3.4	1.9	-1.1
Government effectiveness	**0.86**	**0.21**	0.15	-0.03	28.4	4.6	2.1	-0.8
Regulatory quality	**0.78**	0.20	0.11	0.05	18.2	2.7	1.1	0.8
Rule of law	**0.84**	**0.24**	0.16	-0.03	24.9	5.1	2.5	-0.9
Control of corruption	**0.88**	**0.18**	0.11	-0.02	31.0	4.2	1.5	-0.4
Ethnic diversity index	0.13	**0.43**	0.07	0.04	1.6	4.3	0.8	0.4
Religious diversity index	-0.28	0.11	-0.08	**0.39**	-1.8	0.8	-0.8	5.5

Significant loadings at the 5% level are in bold face.

Table 2: Scores of factors[a] , income per capita[b] , HDI, DI, and ranking differences between income and DI, and HDI and DI

Factor I			Factor II			Factor III			Factor IV			Income ('000)		HDI		DI		Income rank -DI rank	HDI rank -DI rank
USA	1	(2.29)	BGR	1	(1.31)	SGP	1	(3.27)	GTM	1	(2.06)	NOR	31.04	CAN	0.935	CHE	1.000	2	11
CHE	0.99	(2.26)	RUS	0.99	(1.30)	CHN	0.95	(3.11)	NIC	0.95	(1.79)	USA	30.33	NOR	0.934	USA	0.974	0	1
SWE	0.97	(2.19)	UKR	0.97	(1.20)	KOR	0.81	(2.37)	PRY	0.95	(1.76)	CHE	25.78	USA	0.929	CAN	0.952	4	-2
AUS	0.95	(2.11)	ROM	0.95	(1.13)	HKG	0.78	(2.10)	ECU	0.91	(1.53)	DNK	25.67	AUS	0.929	SWE	0.920	14	2
NZL	0.95	(2.11)	MDA	0.94	(1.10)	MYS	0.71	(1.94)	CRI	0.89	(1.49)	AUT	24.54	GBR	0.927	AUS	0.912	5	-1
CAN	0.92	(2.00)	POL	0.94	(1.07)	MNG	0.73	(1.93)	YEM	0.89	(1.44)	IRL	24.44	SWE	0.926	FIN	0.871	11	5
NOR	0.9	(1.96)	SVN	0.94	(1.06)	THA	0.67	(1.71)	VEN	0.86	(1.29)	CAN	24.38	NLD	0.925	NZL	0.867	13	11
DNK	0.9	(1.94)	ITA	0.94	(1.06)	BWA	0.62	(1.48)	MEX	0.86	(1.28)	NLD	23.94	BEL	0.925	DNK	0.860	-4	5
FIN	0.9	(1.93)	GEO	0.93	(1.05)	GAB	0.56	(1.25)	PAN	0.85	(1.22)	JPN	23.53	JPN	0.924	DEU	0.839	3	5
GBR	0.87	(1.84)	GRC	0.93	(1.03)	DZA	0.56	(1.22)	SLV	0.83	(1.10)	AUS	23.48	FRA	0.917	NLD	0.836	-2	-3
NLD	0.85	(1.79)	HKG	0.92	(1.01)	IRN	0.57	(1.20)	PHL	0.82	(1.08)	BEL	23.24	FIN	0.917	GBR	0.825	5	-6
AUT	0.8	(1.60)	BRB	0.92	(0.98)	MUS	0.57	(1.14)	ARG	0.82	(1.04)	DEU	23.18	CHE	0.915	AUT	0.812	-7	3
DEU	0.8	(1.59)	CZE	0.91	(0.98)	SVK	0.57	(1.11)	JAM	0.81	(1.01)	HKG	22.67	DNK	0.911	NOR	0.796	-12	-11
FRA	0.76	(1.44)	LTU	0.92	(0.97)	CZE	0.55	(1.09)	CHL	0.8	(0.97)	FRA	22.44	DEU	0.911	JPN	0.723	-5	-5
BEL	0.73	(1.35)	HRV	0.91	(0.97)	CHL	0.54	(1.03)	URY	0.79	(0.92)	ITA	22.32	AUT	0.908	FRA	0.714	-1	-5
IRL	0.73	(1.34)	ESP	0.91	(0.94)	JAM	0.54	(0.98)	COL	0.79	(0.89)	GBR	21.92	IRL	0.907	BEL	0.703	-5	-8
JPN	0.72	(1.31)	HUN	0.9	(0.93)	LKA	0.52	(0.87)	MYS	0.78	(0.85)	FIN	21.62	ITA	0.903	ITA	0.675	-2	0
SGP	0.69	(1.24)	SVK	0.89	(0.89)	JPN	0.49	(0.74)	EGY	0.78	(0.82)	SWE	21.47	NZL	0.903	ESP	0.667	3	1
ESP	0.62	(0.97)	DEU	0.89	(0.85)	EST	0.49	(0.72)	IRN	0.77	(0.77)	SGP	19.82	ESP	0.899	SGP	0.648	0	2
ITA	0.61	(0.92)	ARM	0.88	(0.82)	TTO	0.46	(0.67)	DZA	0.77	(0.77)	NZL	17.95	CYP	0.886	IRL	0.639	-14	-4
CYP	0.57	(0.78)	KOR	0.87	(0.82)	IDN	0.45	(0.66)	PAK	0.77	(0.76)	ESP	17.53	SGP	0.881	HKG	0.622	-8	2
PRT	0.56	(0.75)	PRT	0.87	(0.78)	IRL	0.46	(0.63)	PER	0.76	(0.74)	PRT	14.97	GRC	0.875	SVN	0.610	3	3
SVN	0.53	(0.66)	JPN	0.87	(0.77)	MOZ	0.46	(0.60)	FRA	0.75	(0.73)	GRC	14.73	HKG	0.872	PRT	0.574	-1	1
HKG	0.52	(0.65)	AUT	0.87	(0.75)	IND	0.45	(0.55)	KGZ	0.75	(0.67)	CYP	14.49	PRT	0.864	CYP	0.565	0	-4
EST	0.49	(0.52)	TTO	0.85	(0.73)	NIC	0.44	(0.47)	JPN	0.73	(0.63)	SVN	14.46	SVN	0.861	BRB	0.539	2	1
BRB	0.47	(0.42)	CAN	0.86	(0.72)	MAR	0.43	(0.46)	SWE	0.72	(0.60)	KOR	13.40	BRB	0.858	GRC	0.522	-3	-4
KOR	0.44	(0.35)	CYP	0.86	(0.72)	CHE	0.42	(0.39)	CYP	0.73	(0.58)	BRB	13.23	KOR	0.854	EST	0.510	5	8
GRC	0.44	(0.31)	EST	0.85	(0.70)	HUN	0.41	(0.37)	NOR	0.72	(0.56)	CZE	12.82	CZE	0.843	CZE	0.504	0	0
CZE	0.42	(0.28)	BEL	0.85	(0.66)	PRT	0.42	(0.37)	BOL	0.73	(0.54)	ARG	10.87	ARG	0.837	KOR	0.494	-3	-2
MYS	0.42	(0.25)	LKA	0.82	(0.60)	NLD	0.41	(0.33)	LKA	0.72	(0.50)	HUN	10.85	CHL	0.826	HUN	0.473	0	3
CHL	0.4	(0.20)	FRA	0.83	(0.59)	ESP	0.41	(0.32)	MAR	0.71	(0.41)	SVK	10.21	SVK	0.825	POL	0.459	3	3
HUN	0.39	(0.16)	THA	0.81	(0.59)	SVN	0.4	(0.32)	DOM	0.7	(0.39)	EST	8.89	URY	0.825	LTU	0.444	8	7
CRI	0.37	(0.05)	MUS	0.81	(0.59)	GHA	0.39	(0.26)	ITA	0.69	(0.39)	ZAF	8.77	HUN	0.817	SVK	0.425	-2	-2
URY	0.36	(0.02)	NLD	0.83	(0.58)	TUR	0.39	(0.25)	BEL	0.69	(0.38)	POL	8.54	POL	0.814	BGR	0.411	16	8
LTU	0.35	(0.01)	FIN	0.83	(0.57)	BGD	0.38	(0.23)	NZL	0.68	(0.37)	MUS	8.51	EST	0.801	HRV	0.401	2	2
POL	0.35	(-0.01)	CHE	0.82	(0.54)	VEN	0.37	(0.23)	GRC	0.69	(0.37)	CHL	8.24	CRI	0.797	CHL	0.373	0	-6
BWA	0.33	(-0.04)	GBR	0.82	(0.54)	AUS	0.37	(0.13)	HKG	0.69	(0.35)	HRV	8.13	HRV	0.795	URY	0.370	1	-5
SVK	0.32	(-0.06)	CHN	0.79	(0.51)	DOM	0.36	(0.09)	AUS	0.67	(0.29)	URY	7.84	TTO	0.793	MUS	0.348	-3	11
UGA	0.31	(-0.17)	DNK	0.81	(0.50)	MLI	0.35	(0.06)	ESP	0.67	(0.28)	MYS	7.80	LTU	0.789	CRI	0.337	4	-3
MLI	0.3	(-0.20)	URY	0.8	(0.49)	NOR	0.34	(0.05)	GBR	0.67	(0.28)	LTU	7.76	MEX	0.784	TTO	0.332	2	-2
BFA	0.3	(-0.20)	SGP	0.79	(0.47)	AUT	0.34	(0.02)	ARM	0.67	(0.24)	MEX	7.65	PAN	0.776	RUS	0.325	4	2
ZAF	0.29	(-0.22)	AUS	0.8	(0.46)	HRV	0.34	(0.01)	CHE	0.66	(0.22)	TTO	7.63	BGR	0.772	ROM	0.324	9	2
HRV	0.28	(-0.24)	USA	0.8	(0.45)	NGA	0.31	(0.01)	IRL	0.66	(0.21)	CRI	7.53	RUS	0.771	UKR	0.300	19	9
MUS	0.27	(-0.27)	SWE	0.8	(0.44)	PAN	0.35	(0.01)	BRA	0.66	(0.15)	BRA	6.65	ROM	0.77	ARG	0.300	-15	-15
ARG	0.25	(-0.36)	IRL	0.79	(0.43)	PHL	0.34	(0.00)	FIN	0.64	(0.13)	RUS	6.43	VEN	0.77	MYS	0.293	-6	15
MEX	0.25	(-0.37)	BRA	0.77	(0.40)	DNK	0.33	(-0.05)	PRT	0.65	(0.12)	BWA	6.20	FJI	0.769	BRA	0.288	-2	4
MRT	0.25	(-0.38)	ARG	0.77	(0.37)	PER	0.33	(-0.05)	NLD	0.63	(0.06)	THA	6.01	COL	0.764	THA	0.286	0	4
YEM	0.25	(-0.39)	NZL	0.77	(0.33)	ECU	0.32	(-0.12)	AUT	0.63	(0.06)	COL	5.85	GEO	0.762	MDA	0.285	39	18
ZMB	0.24	(-0.41)	IDN	0.75	(0.32)	ROM	0.31	(-0.13)	USA	0.62	(0.02)	GAB	5.78	MUS	0.761	ZAF	0.277	-16	18
PAN	0.22	(-0.45)	NOR	0.76	(0.32)	CRI	0.31	(-0.15)	FJI	0.64	(0.01)	BGR	5.73	BRA	0.747	GEO	0.273	28	-2
BGR	0.22	(-0.45)	CHL	0.74	(0.28)	BRA	0.3	(-0.19)	CAN	0.62	(0.00)	ROM	5.58	THA	0.745	MEX	0.273	-10	-11

Table 2 continued

Factor I			Factor II			Factor III			Factor IV			Income ('000)		HDI		DI		HDI rank -DI rank	Income rank -DI rank
TTO	0.22	(-0.46)	FJI	0.74	(0.27)	FIN	0.29	(-0.21)	DNK	0.61	(-0.05)	TUR	5.53	PHL	0.744	PAN	0.270	1	-11
GTM	0.22	(-0.48)	COL	0.73	(0.24)	MEX	0.29	(-0.24)	HRV	0.62	(-0.06)	PAN	5.41	UKR	0.744	TUR	0.266	-1	5
GHA	0.21	(-0.50)	TUR	0.73	(0.22)	BFA	0.29	(-0.24)	GEO	0.62	(-0.07)	IRN	5.38	PER	0.737	LKA	0.255	16	3
ETH	0.22	(-0.50)	CRI	0.72	(0.19)	GRC	0.29	(-0.25)	SGP	0.61	(-0.07)	VEN	5.21	PRY	0.736	FJI	0.253	3	-9
BRA	0.21	(-0.51)	PAN	0.7	(0.13)	PNG	0.26	(-0.26)	DEU	0.61	(-0.07)	DOM	5.20	JAM	0.735	ARM	0.243	20	6
MOZ	0.21	(-0.51)	MEX	0.7	(0.10)	NZL	0.29	(-0.26)	BRB	0.61	(-0.09)	DZA	4.92	LKA	0.733	CHN	0.240	9	7
JAM	0.2	(-0.53)	KGZ	0.69	(0.07)	CAN	0.27	(-0.32)	TUR	0.62	(-0.10)	FJI	4.52	TUR	0.732	COL	0.237	-10	-11
GAB	0.19	(-0.55)	VEN	0.69	(0.06)	POL	0.27	(-0.34)	MUS	0.6	(-0.17)	PRY	4.39	DOM	0.729	MAR	0.231	8	19
TUR	0.19	(-0.57)	ECU	0.66	(-0.04)	PRY	0.27	(-0.36)	POL	0.59	(-0.19)	PER	4.35	ECU	0.722	JAM	0.223	5	-4
BOL	0.19	(-0.58)	MAR	0.65	(-0.07)	DEU	0.26	(-0.41)	KOR	0.58	(-0.24)	SLV	4.20	MYS	0.722	SLV	0.221	0	7
KHM	0.19	(-0.58)	ZAF	0.65	(-0.08)	UKR	0.25	(-0.42)	MNG	0.59	(-0.26)	UKR	3.86	ARM	0.721	PER	0.221	-2	-8
CIV	0.19	(-0.58)	DOM	0.65	(-0.09)	YEM	0.24	(-0.43)	GAB	0.59	(-0.27)	PHL	3.59	IRN	0.709	DOM	0.214	-7	-4
TZA	0.19	(-0.60)	DZA	0.65	(-0.11)	EGY	0.24	(-0.46)	BGD	0.59	(-0.28)	GTM	3.53	CHN	0.706	MNG	0.213	19	9
EGY	0.18	(-0.62)	PER	0.65	(-0.11)	MDA	0.25	(-0.48)	SVK	0.58	(-0.28)	JAM	3.44	KGZ	0.706	EGY	0.211	3	9
MAR	0.18	(-0.62)	SLV	0.64	(-0.15)	MRT	0.23	(-0.50)	LTU	0.58	(-0.29)	CHN	3.31	MDA	0.7	VEN	0.210	-11	-21
PHL	0.18	(-0.62)	IRN	0.61	(-0.24)	BEL	0.23	(-0.55)	MLI	0.58	(-0.31)	MAR	3.24	ZAF	0.697	ECU	0.207	2	-6
SLV	0.18	(-0.62)	JAM	0.6	(-0.27)	RUS	0.21	(-0.61)	SVN	0.56	(-0.34)	EGY	3.11	SLV	0.696	BWA	0.207	-22	8
MNG	0.17	(-0.63)	EGY	0.61	(-0.28)	COL	0.21	(-0.63)	CZE	0.56	(-0.36)	ECU	3.11	DZA	0.683	IDN	0.205	2	1
NIC	0.17	(-0.66)	IND	0.6	(-0.29)	CYP	0.21	(-0.65)	UGA	0.57	(-0.38)	LKA	2.98	IDN	0.67	IND	0.203	5	9
KEN	0.16	(-0.68)	MNG	0.58	(-0.34)	LTU	0.21	(-0.67)	THA	0.56	(-0.40)	IDN	2.83	BOL	0.643	PHL	0.202	-8	-18
PER	0.16	(-0.68)	MYS	0.58	(-0.40)	ZAF	0.2	(-0.71)	CHN	0.56	(-0.42)	PNG	2.23	NIC	0.631	KGZ	0.201	14	-7
FJI	0.16	(-0.69)	PHL	0.58	(-0.40)	BRB	0.18	(-0.82)	GHA	0.56	(-0.42)	NIC	2.18	MNG	0.628	IRN	0.194	-19	-10
THA	0.14	(-0.71)	BGD	0.55	(-0.50)	USA	0.17	(-0.85)	IDN	0.56	(-0.44)	BOL	2.16	EGY	0.623	BOL	0.182	0	-3
IND	0.14	(-0.75)	PRY	0.51	(-0.66)	ARM	0.16	(-0.86)	TTO	0.55	(-0.45)	IND	2.15	GTM	0.619	DZA	0.181	-18	-6
DOM	0.13	(-0.77)	BOL	0.5	(-0.73)	UGA	0.17	(-0.86)	BFA	0.55	(-0.50)	ARM	2.15	BWA	0.593	KHM	0.172	6	8
PAK	0.14	(-0.77)	KHM	0.46	(-0.90)	SWE	0.16	(-0.88)	KHM	0.55	(-0.50)	GHA	1.79	GAB	0.593	BGD	0.171	8	9
PRY	0.13	(-0.79)	PAK	0.45	(-0.93)	FRA	0.16	(-0.89)	IND	0.55	(-0.51)	GEO	1.72	MAR	0.589	PRY	0.170	-19	-23
CMR	0.13	(-0.79)	PNG	0.44	(-0.98)	ETH	0.16	(-0.92)	ROM	0.53	(-0.55)	PAK	1.71	IND	0.563	GHA	0.160	-2	1
PNG	0.13	(-0.80)	NIC	0.43	(-1.00)	GBR	0.15	(-0.93)	BGR	0.53	(-0.57)	CMR	1.68	GHA	0.556	NIC	0.160	-7	-8
COL	0.12	(-0.81)	BWA	0.41	(-1.06)	SLV	0.15	(-0.96)	EST	0.53	(-0.57)	MRT	1.61	PNG	0.542	PAK	0.156	-2	2
NGA	0.12	(-0.81)	GHA	0.38	(-1.21)	PAK	0.14	(-0.97)	MRT	0.53	(-0.57)	KHM	1.57	CMR	0.528	GTM	0.151	-18	-7
ROM	0.1	(-0.86)	GTM	0.35	(-1.33)	ZMB	0.14	(-1.01)	HUN	0.52	(-0.59)	MNG	1.44	PAK	0.522	PNG	0.149	-11	-2
IRN	0.1	(-0.87)	CIV	0.34	(-1.37)	CMR	0.13	(-1.01)	UKR	0.5	(-0.76)	CIV	1.44	KHM	0.512	CIV	0.146	0	6
ECU	0.1	(-0.88)	CMR	0.34	(-1.37)	ARG	0.13	(-1.03)	MDA	0.48	(-0.84)	BGD	1.38	KEN	0.508	MRT	0.140	-4	2
VEN	0.1	(-0.89)	KEN	0.31	(-1.47)	GEO	0.14	(-1.03)	PNG	0.48	(-0.91)	KGZ	1.34	BGD	0.461	GAB	0.138	-37	-9
BGD	0.08	(-0.97)	GAB	0.31	(-1.49)	ITA	0.13	(-1.03)	RUS	0.42	(-1.19)	MDA	1.27	MRT	0.451	TZA	0.137	10	5
LKA	0.06	(-1.00)	TZA	0.3	(-1.52)	TZA	0.13	(-1.05)	KEN	0.43	(-1.21)	UGA	1.13	YEM	0.448	KEN	0.135	1	-3
RUS	0.06	(-1.01)	MRT	0.26	(-1.67)	BOL	0.12	(-1.06)	TZA	0.37	(-1.53)	KEN	0.94	NGA	0.439	CMR	0.133	-9	-7
KGZ	0.06	(-1.04)	ETH	0.21	(-1.88)	KEN	0.12	(-1.11)	CMR	0.36	(-1.59)	BFA	0.93	CIV	0.42	ETH	0.127	6	6
CHN	0.04	(-1.06)	NGA	0.2	(-1.95)	KGZ	0.1	(-1.17)	NGA	0.35	(-1.63)	NGA	0.80	ZMB	0.42	ZMB	0.124	4	0
DZA	0.04	(-1.09)	MOZ	0.19	(-1.95)	FJI	0.08	(-1.27)	CIV	0.35	(-1.64)	MOZ	0.79	TZA	0.415	MOZ	0.123	0	3
UKR	0.03	(-1.10)	ZMB	0.18	(-2.00)	GTM	0.07	(-1.33)	ETH	0.3	(-1.92)	MLI	0.74	UGA	0.409	YEM	0.115	1	-5
MDA	0.02	(-1.16)	YEM	0.16	(-2.12)	KHM	0.05	(-1.39)	MOZ	0.27	(-2.08)	YEM	0.73	MLI	0.38	BFA	0.112	-4	3
GEO	0.01	(-1.20)	BFA	0.05	(-2.52)	BGR	0.05	(-1.44)	ZMB	0.18	(-2.62)	ZMB	0.73	MOZ	0.341	MLI	0.106	-2	-1
IDN	0.01	(-1.21)	MLI	0.02	(-2.66)	URY	0.03	(-1.54)	ZAF	0.06	(-3.24)	ETH	0.61	ETH	0.309	NGA	0.103	-5	-7
ARM	0	(-1.24)	UGA	0	(-2.75)	CIV	0	(-1.67)	BWA	0	(-3.58)	TZA	0.48	BFA	0.303	UGA	0.103	-9	-4

a. Columns are in the descending order of the factor scores and figures outside and inside parenthesis are respectively rescaled and raw factor scores. Raw factor scores re-scaled by equation (15).

b. Income per capita is the 10 year average of 1995 to 2004, measured in PPP$ 1995.

Appendix

Table A.1: Countries included in the sample

Low income[a]		Lower middle income		Upper middle income		High income	
Country code	**Country name**	**Country code**	**Country name**	**Country code**	**Country name**	**Country code**	**Country name**
BFA	Burkina Faso	ARM	Armenia	ARG	Argentina	AUS	Australia
BGD	Bangladesh	BOL	Bolivia	BGR	Bulgaria	AUT	Austria
CIV	Cote d'Ivoire	CHN	China	BRA	Brazil	BEL	Belgium
ETH	Ethiopia	CMR	Cameroon	BWA	Botswana	BRB	Barbados
GHA	Ghana	COL	Colombia	CHL	Chile	CAN	Canada
IND	India	DOM	Dominican Rep.	CRI	Costa Rica	CHE	Switzerland
KEN	Kenya	DZA	Algeria	EST	Estonia	CYP	Cyprus
KGZ	Kyrgyzstan	ECU	Ecuador	GAB	Gabon	CZE	Czech Rep.
KHM	Cambodia	EGY	Egypt	HRV	Croatia	DEU	Germany
MLI	Mali	FJI	Fiji	HUN	Hungary	DNK	Denmark
MNG	Mongolia	GEO	Georgia	LTU	Lithuania	ESP	Spain
MOZ	Mozambique	GTM	Guatemala	MEX	Mexico	FIN	Finland
MRT	Mauritania	IDN	Indonesia	MUS	Mauritius	FRA	France
NGA	Nigeria	IRN	Iran	MYS	Malaysia	GBR	UK
PAK	Pakistan	JAM	Jamaica	PAN	Panama	GRC	Greece
PNG	Papua N. Guinea	LKA	Sri Lanka	POL	Poland	HKG	Hong Kong
TZA	Tanzania	MAR	Morocco	ROM	Romania	IRL	Ireland
UGA	Uganda	MDA	Moldova	RUS	Russia	ITA	Italy
YEM	Yemen	NIC	Nicaragua	SVK	Slovakia	JPN	Japan
ZMB	Zambia	PER	Peru	TUR	Turkey	KOR	South Korea
		PHL	Philippines	URY	Uruguay	NLD	Netherlands
		PRY	Paraguay	VEN	Venezuela	NOR	Norway
		SLV	El Salvador	ZAF	South Africa	NZL	New Zealand
		THA	Thailand			PRT	Portugal
		UKR	Ukraine			SGP	Singapore
						SVN	Slovenia
						SWE	Sweden
						TTO	Trinidad & Tobago
						USA	US

a. The categorization of countries by income groups is based on the World Bank (2008)

Table A.2: Variables included in the sample

Variable	Description of construction	Expected relationship with the level of development	Source
Economic variables			
Age dependency ratio (ratio to the working age pop.)	The ratio of population of age dependence (percentage of combine population of age below 15 years and age above 64) to the working age population (those ages 15-64).	Positive	WDI
Final consumption expenditure, etc. (% GDP)	Sum of household final consumption expenditure (private consumption) and general government final consumption expenditure (general government consumption).	Not clear	WDI
Gross capital formation (% GDP)	Gross capital formation consists of expenditure on the fixed assets of the economy and net changes in the level of inventories. Fixed assets consists of land improvements, plant, machinery, and equipment purchases, and the construction of roads, railways, schools, offices, hospitals, private residential dwellings, and commercial and industrial buildings.	Positive	WDI
Gross domestic savings (%GDP)	Gross domestic savings which are calculated as GDP less final consumption expenditure.	Positive	WDI
Gross fixed capital formulation (%GDP)	Gross fixed capital formation includes land improvements, plant, machinery, and equipment purchases; and the construction of roads, railways, schools, offices, hospitals, private residential dwellings, and commercial and industrial buildings.	Positive	WDI
Gross secondary school enrolment rate	The ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the level of secondary education.	Positive	WDI
Gross tertiary school enrolment rate	The ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the level of tertiary education.	Positive	WDI
Pupil-teacher ratio (primary education)	The number of pupils enrolled in primary school divided by the number of primary school teachers.	Negative	WDI
Migration stock (% of total population)	Number of people born in a country which is different to that where they live as a percentage of total population. It also includes refugees.	Positive	WDI
Net foreign direct investment flows (% GDP)	Net foreign direct investment flows is the sum of equity capital, reinvestment of earnings, other long-term capital, and short-term capital as shown in the balance of payments. This series shows net inflows in the reporting economy.	Positive	WDI
Net migration (per 1000 people)	The net total of migrants during the period: total number of immigrants minus the annual number of emigrants, including both citizens and non citizens.	Positive	WDI

Table A.2 continued			
Variable	**Description of construction**	**Expected relationship with the level of development**	**Source**
Industry value added (% GDP)	Industry corresponds to ISIC divisions 10-45 and includes manufacturing (ISIC divisions 15-37). It consists of value added in mining, manufacturing (also reported as a separate subgroup), construction, electricity, water, and gas.	Not clear	WDI
Services value added (% GDP)	Services correspond to ISIC divisions 50-99. They include value added in wholesale and retail trade (including hotels and restaurants), transport, and government, financial, professional, and personal services such as education, health care, and real estate services. It is included imputed bank service charges, import duties.	Positive	WDI
Population ages 15-64 (% of total)	Population ages 15 to 64 as a percentage of the total population.	Positive	WDI
Ratio of agric. To industry value added	Ratio of agricultural value added to industrial value added. Agriculture corresponds to ISIC divisions 1-5 and includes forestry, hunting, and fishing, as well as cultivation of crops and livestock production.	Negative	WDI
Ratio of industries to services sector value added	Ratio of industrial value added to services value added.	Not clear	WDI
Urban population (% of total population)	The midyear population of areas defined as urban in each country. This series expressed as a percentage of total midyear population.	Positive	WDI
Proxies for technology			
Daily news papers (per 1000 people)	Number of newspapers those published at least four times a week.	Positive	WDI
Fixed line & mobile phone subscribers (per 1000 people)	Fixed lines are telephone mainlines connecting a customer's equipment to the public telephone network. Mobile phone subscribers use portable telephones subscribing to an automatic public mobile telephone service using cellular technology that provides access to the public telephone network.	Positive	WDI
High technology exports (%manufactured exports)	Products with high R&D intensity, such as in aerospace, computers, pharmaceuticals, scientific instruments, and electrical machinery expressed as a percentage of manufactured exports. Manufactured export comprises exports of commodities in SITC sections 5 (chemicals), 6 (basic manufactures), 7 (machinery and transport equipment), and 8 (miscellaneous manufactured goods), excluding division 68 (non-ferrous metals).	Positive	WDI
Internet uses (per 1000 people)	Persons who access to the worldwide network.	Positive	WDI
Passenger cars (per 1000 people)	Number of reported for personal cars.	Positive	WDI
Personal computers (per 1000 people)	Persons with self-contained computers designed to be used by a single individual.	Positive	WDI
Radio (per 1000 people)	Persons who posses radio.	Positive	WDI

Table A.2 continued			
Variable	**Description of construction**	**Expected relationship with the level of development**	**Source**
Proxies for technology			
Scientific and technical journal articles (per 1000 people)	Scientific and technical journal articles which refer to the number of scientific and engineering articles published in physics, biology, chemistry, mathematics, clinical medicine, biomedical research, engineering and technology, and earth and space sciences.	Positive	WDI
Secure Internet servers (per 1 million people)	Servers that use encryption technology in Internet transactions.	Positive	WDI
Television sets (per 1000 people)	The share of households with a television set.	Positive	WDI
Vehicles per (per 1000 people)	Number of reported vehicles.	Positive	WDI
Social, cultural and Income Distribution Variables			
Birth rate, crude (per 1000 people)	The number of live births occurring during the year, estimated at midyear.	Negative	WDI
Ethnic diversity index	Use Hirschman Index: $D = \left[\sum y_i^2\right]^{1/2}$, where D stands for the ethnic diversity index and y_i is the i^{th} ethnic group. Categorization ethnic groups are based on country specific characteristics.	Not clear	CIA
Fertility rate (no. of births per women)	The number of children that would be born to a woman if she were to live to the end of her childbearing years and bear children with current age-specific fertility rates.	Negative	WDI
Gini coefficient	Area under the Lorenz curve. Computed using household income.	Depend on the stage of development	WDI
Gross primary school enrolment ratio	The ratio of total enrolment, regardless of age, to the population of the age group that officially corresponds to the level of education shown. Primary education includes basic reading, writing, and mathematics skills along with an elementary understanding of such subjects as history, geography, natural science, social science, art, and music.	Positive	WDI
HIV incident (per 14-45 aged group)	The percentage of people ages 15-49 who are infected with HIV.	Negative	WDI
Labour force, children 10-14 (% age group)	Children involved in economic activity for at least one hour in the reference week of the survey.	Negative	WDI
Life expectancy at birth, female (years)	The number of years a newborn female infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.	Positive	WDI
Life expectancy at birth, total (years)	The number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.	Positive	WDI
Physicians (per 1000 people)	Graduates of any facility or school of medicine who are working in the country in any medical field (practice, teaching, research).	Positive	WDI

Table A.2 continued			
Variable	**Description of construction**	**Expected relationship with the level of development**	**Source**
Religious diversity index	Use Hirschman Index: $D = \left[\sum y_i^2\right]^{1/2}$, where D stands for religious diversity index and y_i is the i^{th} religious group. There are twenty-two main types of major religions. We consider Catholic, Protestants, Christians, Jews, Muslims, Hindus, and Buddhist as separate groups to construct the index. The categorization is sometimes based on the background of a particular country.	Not clear	CIA
Institutional variables			
Control of corruption	The exercise of public power over private gain. This consists of both petty and grand scale corruption.	Positive	World Bank
Government effectiveness	The ability of a government to formulate and implement sound policies. This is a proxy for the quality of public service provisions, the quality of bureaucracy, the competence of civil servants, the independence of civil service from political pressure and the credibility of the government's commitment to policies.	Positive	World Bank
Political stability	The likelihood of the power of government to be destabilized or be overthrown by unconstitutional and violence acts including terrorism.	Positive	World Bank
Regulatory quality	Focused on market friendly policies. Incidence of price control, inadequate bank supervision, perception of burden imposed by excessive regulation on foreign trade and business development are taken into account.	Positive	World Bank
Rule of law	The Rule of Law measures perceptions of the incidence of both violent and non-violent crimes. The effectiveness and predictability of the judiciary and the enforceability of contracts are also included.	Positive	World Bank
Voice and accountability	Consider many aspects of political process, civil liberties and political rights. This indicator consists of the extent to which citizens of a country are able to participate in the selection of their government. Media independence is also included.	Positive	World Bank
Validation measure			
Income per capita	Purchasing power parity adjusted GDP per capita (1995 International $).	Positive	WDI

Table A.3: Descriptive statistics of variables from 1995 to 2004

Variable	Minimum	Maximum	Mean	Std. Deviation
Age dependency ratio (ratio to the working age pop.)	0.38	1.12	0.62	0.17
Birth rate, crude (per 1000 people)	8.26	50.1	21.07	11.22
Daily news papers (per 1000 people)	0.38	589.39	121.35	129.32
Fertility rate (No. of births per women)	1.03	7.07	2.74	1.54
Final consumption expenditure, etc. (% GDP)	51.19	106.67	80.05	10.25
Fixed line & mobile phone subscribers (per 1,000 people)	3.49	1320.25	390.82	382.47
Net Foreign direct investment (%GDP)	-1.4	23.69	3.49	3.47
Gross capital formulation (% GDP)	12.15	36.09	22.15	4.44
Gross domestic savings (% GDP)	4.79	50.62	20.41	8.08
High-technology exports (% of manufactured exports)	0.01	66.33	11.42	12.73
Value added of industries (% GDP)	9.89	52.84	30.47	8.33
Migration stock (per 1000 people)	0.04	40.01	5.63	6.87
Internet uses (per 1000 people)	0.8	577.06	148.08	171.28
Life expectancy at birth, female (years)	38.27	84.64	70.82	11.71
Life expectancy at birth, total (years)	38.39	81.1	68.16	10.88
Passenger cars (per 1000 people)	0.44	611.43	172.68	185.79
Personal computers (per 1000 people)	0.47	514.49	106.4	141.15
Physicians (per 1000 people)	0.02	4.47	1.68	1.27
Population ages 15-64 (% of total)	47.23	72.53	62.34	6.19
HIV incident (per 1000 people)	0.01	37.65	1.83	4.86
Pupil-teacher ratio (primary education)	10.03	64.78	26.12	12.63
Gross primary school enrolment rate	46.97	149.5	101.8	14.35
Gross secondary school enrolment rate	5.87	176.77	77.65	33.93
Gross tertiary school enrolment rate	0.83	84.25	31.32	22.52
Secure Internet servers (per 1 million people)	0.01	481.82	53.48	107.01
Service sector valued added (% GDP)	24.41	86.82	55.64	11.72
Urban population (% of total population)	11.99	100	58.52	20.91
Vehicles per (per 1000 people)	1.4	776.28	205.6	207.1
Television sets (per 1000 people)	5.35	849.07	286.5	210.03
Radio (per 1000 people)	42.9	2117.55	492.92	388.06
Ratio of value added of indus. to that of services	0.17	1.89	0.59	0.28
Ratio of value added of agric. To that of industry	0.004	4.98	0.58	0.75
Scientific & technical journal articles (per 1000 people)	0.0002	1.12	0.17	0.28
Fixed capital formulation (% GDP)	12.22	34.04	21.05	4.33
Net migration (per 1000 people)	-113.97	81.25	-1.79	24.25
Labour force, children 10-14 (% age group)	0.00	51.33	7.5	11.88
Gini coefficient	24.44	63.01	39.47	9.6
Voice and accountability	-1.42	1.61	0.29	0.84
Political stability	-2.04	1.62	0.17	0.87
Government effectiveness	-1.14	2.44	0.32	0.97
Regulatory quality	-1.41	2.01	0.4	0.78
Rule of law	-1.27	2.14	0.25	0.98
Control of corruption	-1.14	2.46	0.24	1.05
Ethnic diversity index	0.09	1	0.78	0.2
Religious diversity index	0.19	1	0.77	0.17
Income per capita (PPP $ 1995)	476	31,037	9120	8445