Is human development multi-dimensional? *

DECLAN FRENCH†, MICHAEL MOORE† and DAVID CANNING‡

† UKCRC Centre of Excellence for Public Health, School of Management, Queens University, Belfast BT7 1NN. (email - declan.french@qub.ac.uk)

‡ Department of Global Health and Population, Harvard School of Public Health, Harvard University,

Aug 23rd, 2011

Abstract

Stiglitz's *Commission on the Measurement of Economic Performance and Social Progress* (CMEPSP) argued that well-being is multidimensional and identified eight distinct dimensions. Conventional linear techniques confirm that a large number of dimensions are needed to describe development. In contrast, a new non-linear technique which we introduce from chaos theory shows that a smaller number of dimensions are needed to span the development space. From the analysis, variables representing the Health, Education, Inequality and Individual Rights areas of life quality would provide a broad picture of development while income per capita adds little extra information.

JEL Classification*:* I32, O10, O15

Keywords*: development, well-being, dimension, measure, indicator.*

---

* **Acknowledgements:** This work was undertaken by the Centre of Excellence for Public Health (Northern Ireland), a UKCRC Public Health Research Centre of Excellence. Funding from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, Research and Development Office for the Northern Ireland Health and Social Services and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration is gratefully acknowledged

1 Introduction

Human development is a broad concept and has many components. The human development Report 1990 (United Nations Development Program 1991) identified three important components of development: longevity, knowledge, and the standard of living. Taking life expectancy, literacy and school enrolment rates, and income per capita as measures it put forward the human development index as a measure of a country's performance in these dimensions.

The World Bank has shifted from an emphasis on promoting economic growth (e.g. Commission on International Development, 1969) to include a focus on health, education, and social exclusion (see World Development Report 2000/1). The Millennium Development Goals, adopted at the Millennium Summit of the United Nations in September 2000, focused on poverty, mortality, education, health, gender equality and environment.

More recently The *Commission on the Measurement of Economic Performance and Social Progress* (2009) put forward eight distinct spheres of wellbeing in addition to consumption and wealth: health, education, personal activity, political voice and governance, social connectedness, environmental conditions, personal insecurity, and economic insecurity. In addition, like Hicks (1997) and Anand and Sen (2000), it argues for a focus on inequality in these indicators as well measures of their average levels. While it did not give concrete measures of each type of wellbeing it argued that they should be measured and that an index of these measures could be constructed. Alkire (2002) discusses the conceptual basis for thinking of human development as taking place in multiple dimensions.

Different dimensions of development could be aggregated into a single measure if we were utilitarian and had a well defined social welfare function. However, it is difficult to construct such a welfare function since it involves interpersonal comparisons of utility (Harsanyi, 1955). Indeed it is not even clear that this is what we mean by human development. Sen (1999) conceptualizes development as enlarging the set of human capabilities; the choice set people have, without any need to consider the utility of the choices that are made.

While it is difficult to construct a theoretically justified single measure of development it is noticeable that many of the development indicators we consider move together. Even though the indicators measure different things, if they move together a single index might be a good proxy for the overall level of human development. Cahill (2005) shows that the longevity, knowledge, and the standard of living components of the human development index are highly positively correlated and can be reduced to an index based on a weighted average. Using ten indicators on the same dimensions of health, education and standard of living, Alkire and Santos (2010) also construct an aggregate multi-dimensional poverty measure. When considering broader descriptions of development including many other distinct dimensions, a single index will not suffice. However linear data reduction techniques, such as principal component analysis or factor analysis, as employed by Adelman and Morris (1967) and McGillivray (2005) tend to overestimate the number of dimensions in development because they do not allow for the possibility of nonlinear relationships between indicators. We measure the dimension of development indicators using techniques developed by Grassberger et al (1983) and Takens (1985). These techniques are non-parametric and allow for non-linear relationships in the data. They were originally developed for the study of dynamic systems. in high dimensions where the outcome paths may be cycles, or strange attractors, that may appear random at first sight but have a structure that means they occupy only a low dimensional subspace. We consider a set of twenty three development indicators. We find these have dimension five. Our method identifies the dimension of the development space but does not provide a coordinate system. We investigate which variables span the five dimensions of the larger development space. We find that variables representing Health, Education, Inequality and Individual Rights are consistently represented in those subsets with the highest dimension and so should be included in any summary picture of development.

Our results suggest that that the process of development is not arbitrary; there are "laws" that seem to limit possible outcomes. For example, Preston (1975) argues that at each point in time there is a (non-linear) relationship between a country's level of income per capita and its life expectancy, a relationship that continues to hold (Bloom and Canning, 2007). Kuznets (1955) suggested an inverted–U relationship between income and inequality with the distribution of income initially worsening with economic growth and improving only at higher levels of income. There may also be links between income levels and political rights (Dasgupta and Weale, 1992; Barro 1996), political stability (Fedderke and Klitgaard, 1998) and education (Fedderke and Klitgaard, 1998). Ranis et al. (2000) find two way linkages between economic growth and other measures of human development, and these linkages will limit the range of possible outcomes.

Section 2 describes the data we use. In section 3 we describe the techniques we employ to estimate the dimension of the space. Results are presented in section 4 with conclusions in section 5.

2 Data

We are interested in broader indicators of development. To this end we use variables from a dataset constructed by Easterly (1999) which is the most comprehensive collection of development indicators in number, time span and country availability known to the authors. In addition to income the Easterly dataset covers seven further areas of life quality: Individual rights and democracy, political instability, education, health, transport and communications, inequalities and negative factors ("bads") associated with development such as pollution. These data are more fully described in the original paper. Data were assembled from a number of sources including United Nations, World Bank and International Labor Organization as well as academic research such as Barro-Lee for education measures and Penn World tables for data on income.

Our methods do not allow us to deal with discrete data, or missing observations. We therefore removed discrete variables and then the variables/ countries with the highest proportion of missing values until a dataset of continuous variables with no missing data was formed. [1] As none of the Individual Rights and Democracy variables were represented in the reduced dataset, the first principal component accounting for 70% of the total variance of the variables in this section was extracted and this constructed variable was then matched to the countries in the reduced dataset. [2] This produced a dataset of 23 variables across 62 countries with a

---

[1] A closer examination of the Under-5 mortality rate shows it is unreliable and it was therefore excluded. For example, according to Easterly there were 8 deaths per 1000 in the under 5s in Bangladesh in 1990 and 139 deaths in this group in the US in the same year.

[2] Two variables were excluded as they were only available for a small subset of countries: *Human Rights Rating* and *% of Children (age 10-14) working (-)*. Principal component analysis was performed

fairly even spread across the World Bank classifications by Income Group. [3]

Descriptive statistics are provided in Table 1.[4]

TABLE 1: *Descriptive statistics*

Variable	*Symbol*	*Mean*	*Max*	*Min*	*Std dev*
CO2 Emissions, per capita	EMSCO2PC	4.07	19.75	0.04	4.72
Cellular phones	CELLUL	5.37	53.88	0.00	11.76
Carbon Dioxide Emissions	EMSCBDX	1.09	5.24	0.01	1.25
Armed Forces	AMRFORC	6.08	44.00	0.80	6.65
Cars per capita	CARS	0.17	0.74	0.00	0.20
F/M gross enrollment ratio for primary ed	ENRPRFM	0.94	1.04	0.56	0.12
F/M gross enrollment ratio for secondary ed	ENRSCFM	0.90	2.39	0.24	0.33
F/M gross enrollment ratio for higher ed	ENRHGFM	0.74	1.39	0.11	0.32
Female to male average schooling years (age 26+)	AVGSCHFM	0.75	1.02	0.30	0.23
War Deaths	WARDEATH	164	6056	0	815
Daily calorie intake	NSDCI	2730	3987	1803	589
daily protein intake (grams)	NSDPI	74	120	30	24
Radios per capita	NSRADPC	0.43	2.13	0.05	0.39
TV sets per capita	NSTVPCB	0.18	0.69	0.00	0.17
Mortality - Infant	NSMORT	51	169	5	44
Percentage of "no schooling" in total population	NSNOSCH	26	88	0	25
Average schooling years in the total population age 25+	NSSCHL	5.37	12.04	0.82	2.97
Total gross enrollment ratio for primary ed	NSPENRL	0.92	1	0.25	0.17
Total gross enrollment ratio for secondary ed	NSSENRL	0.54	1	0.04	0.30
Total gross enrollment ratio for higher ed	NSHENRL	0.17	0.58	0.00	0.14
Individual Rights & Democracy	Ind Rts & Dem	0.24	1.88	--1.61	1
Life Expectancy	LE	65	77	40	10
GDP per capita	GDP	5653	18054	519	5286

Our variables cover material living standards, health, education, political voice and governance, physical and economic insecurity, and the environment. We do not cover two areas that CMEPSP recommended: (a) personal activities including work e.g. leisure, commuting, paid/unpaid work, housing and (b) social connections and

on the remaining eight : *Freedom from Expropriation*, *Government does not break contracts*, *Bureaucratic quality*, *Rule of law*, *Freedom from Corruption*, *Civil Liberties, Political Rights* and *Index of independence of politics from military*. All variables were highly correlated with the Principal Component. The *Political Rights* variable had the lowest communality ($h^2 = 0.582$).

[3] Low (9 countries), Middle (27), High (21). Source for classifications : web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20421402~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html

[4] All variables used were then transformed by scaling relative to the data range

$$\text{scaling index} = \frac{(\text{actual value} - \text{min value})}{(\text{max value} - \text{min value})}.$$

relationships. There are no robust international data sources in these areas at present. Future responses by national statistical agencies to the CMEPSP report's recommendations may allow us to test whether these extra areas add additional dimensions.

3 Methods

Assuming Linear Relationships

There are a number of statistical techniques to determine the number of underlying factors affecting a set of variables. One of the most commonly used is *principal component analysis* (PCA). This technique finds orthogonal components and ranks them by the amount of variation in the data they can explain. The determination of dimensionality using PCA is complicated by the fact that a criterion must be set to limit the components used; components below this limit explain little of the variation in the data and are considered unnecessary dimensions. Peres-Neto et al. (2005) compare 20 stopping rules used commonly in the literature often involving Monte Carlo simulation to generate critical values. In this paper, we examine a simple rule which is based on extracting components until 95% of the total variance is explained; this rule is advocated by Joliffe (1986).

Assuming Non-Linear Relationships

As we expect the existence of non-linearities in associations between quality of life variables, approaches based on linearly relating variables to a number of factors will inevitably overestimate the number of dimensions. Non-linear adaptation of factor analysis has been developed but relies on *a priori* knowledge of the non-linear model (McDonald, 1967a, 1967b, 1979 ; Yalcin and Amemiya, 2001). Non-linear developments of Principal Component Analysis using neural networks avoid assumptions of functional form but have a number of limitations including large approximation error and inability to cope with self-intersections and discontinuities (Malthouse, 1998).

We use the approach of estimating the correlation dimension of the data (Theiler, 1990) which produces results closely related to other standard measures of dimension (Grassberger et al, 1983). The idea of this measure is that we take an arbitrary data point, draw a ball of radius r around it and count the number of data points that lie within the radius r. We then double the radius r and count again. The volume of space in the ball will increase by a factor 2^{υ} when the radius doubles in size where the data has dimension υ. For example if all the data points are uniformly distributed on a one dimensional line, doubling the radius of the ball will approximately double the number of points captured. Similarly if the data is uniform on a two dimensional surface doubling the radius of the ball will quadruple the number of points captured. This intuition then forms the basis of the correlation dimension measure. The mathematics of the calculation are explained in more detail in the Appendix. In simple terms, the average number of points C(r) contained within a ball of radius r is found for each r. The relationship between C(r) and r is then used to determine the dimension υ.

A simple way to estimate υ is as the slope of the graph of log C(r) against log r for r small (Grassberger and Procaccia, 1983). The choices of r and s on horizontal axis of the graph used to estimate the slope will obviously not use information at all points. An alternative would be to use weighted least squares to estimate the slope (Theiler, 1990), however, C(r) values are not independent for increasing r since balls of increasing radius necessarily include the smaller balls so this method may not have desirable properties.

Takens (1985) provides an alternative approach based on a maximum likelihood estimator based on the distribution of distances between points in the dataset. An estimate, $\hat{\upsilon}$, and a confidence interval can be then determined for the dataset for a given choice of boundary radius R_0.

In practice, the relationship between C(r) and r does not hold consistently over a large range of distances, r. When r is large, C(r) is affected by the 'edge' of the dataset since points X_i near the edge have less nearby X_j to form pairs than points in the interior. C(r) at large r is therefore biased downwards and the scaling relationship only holds consistently over shorter r. Estimating over smaller distances also raises problems. If the radius is small then the number of distances between pairs being consideredwill be small and consequently the standard error will be large. A degree of compromise is therefore necessary between bias at large radius and inaccuracy at small radius.To address this we follow the procedure suggested by Smith (1997) which involves using the largest radius for which the estimate lies inside the confidence interval of estimates for any smaller radius .

4 Results

A principal component analysis of the data set of 23 development indicators. was carried out and results are given in Table 2. It can be seen that in order to account for 95% of the variation of the data it is necessary to extract 10 factors. Even setting the criterion at the more moderate level of 90% of the total variance we would have to extract 7 factors. There are therefore a large number of dimensions to this data set according to this rule.

Assuming the relationships between these variables and the underlying factors may be non-linear we calculate the correlation dimension of this data. Using the Grassberger-Procaccia approach, we plot the relationship between log C(r), the average proportion of all points within radius r, and log r in Figure 1. The slope of this line then gives us the dimension of this data. The uneven part to the left is attributed to noise at small measurement scales and the remainder can be seen to be approximately linear. The slope is then calculated over a suitable section (-0.80 to -0.20 in this case) giving $\hat{\upsilon} = \frac{(5.71 - 2.71)}{(-0.20 - (-0.80))} = 5.0$ for the points indicated.

The Takens approach gives us a maximum likelihood estimate and allows us to calculate confidence intervals. Also when used in conjunction with the Smith procedure, determination of the dimension estimate becomes a more objective process. In figure 2 we plot the estimate $\hat{\upsilon}$ against N* the number of data pairs used to calculate distances with 95% confidence intervals. It can be seen that the width of the confidence interval is dependent on the number of distances (N*) over which the estimate $\hat{\upsilon}$ is taken. On the left of the graph the confidence interval is very large as only a small number of the shortest distances between points are considered in the estimate. Increasing N* increases the accuracy of the estimate but also causes a downwards bias in $\hat{\upsilon}$ due to edge effects which become gradually more severe. For example when N*=500, the dimension estimate $\hat{\upsilon}$ falls to 3.2. Following Smith (1999), we then find a compromise between bias and inaccuracy by proceeding from small N* on the left to larger N* on the right and checking if each estimate fits within the confidence intervals of estimates at all smaller N*. The estimate indicated was selected as this was the largest possible radius such that $\hat{\upsilon}(R_0)$ was within the 95% confidence intervals at all $R < R_0$ For this value, $\hat{\upsilon} = 5.21 \pm 0.56$ or a dimension of 5 to the nearest integer.

TABLE 2 : *Principal Component Analysis (Easterly data)*

Component	*Eigenvalue*	*% of Variance*	*Cumulative %*
1	14.838	64.5	64.5
2	2.432	10.6	75.1
3	1.096	4.8	79.9
4	1.002	4.4	84.2
5	0.708	3.1	87.3
6	0.572	2.5	89.8
7	0.480	2.1	91.9
8	0.326	1.4	93.3
9	0.247	1.1	94.4
10	0.237	1.0	95.4
11	0.202	0.9	96.3
12	0.162	0.7	97.0
13	0.141	0.6	97.6
14	0.128	0.6	98.1
15	0.113	0.5	98.6
16	0.083	0.4	99.0
17	0.069	0.3	99.3
18	0.055	0.2	99.5
19	0.041	0.2	99.7
20	0.030	0.1	99.8
21	0.024	0.1	99.9
22	0.009	0.0	100.0
23	0.004	0.0	100.0

Figure 1. Grassberger-Procaccia estimate of dimension for Easterly data

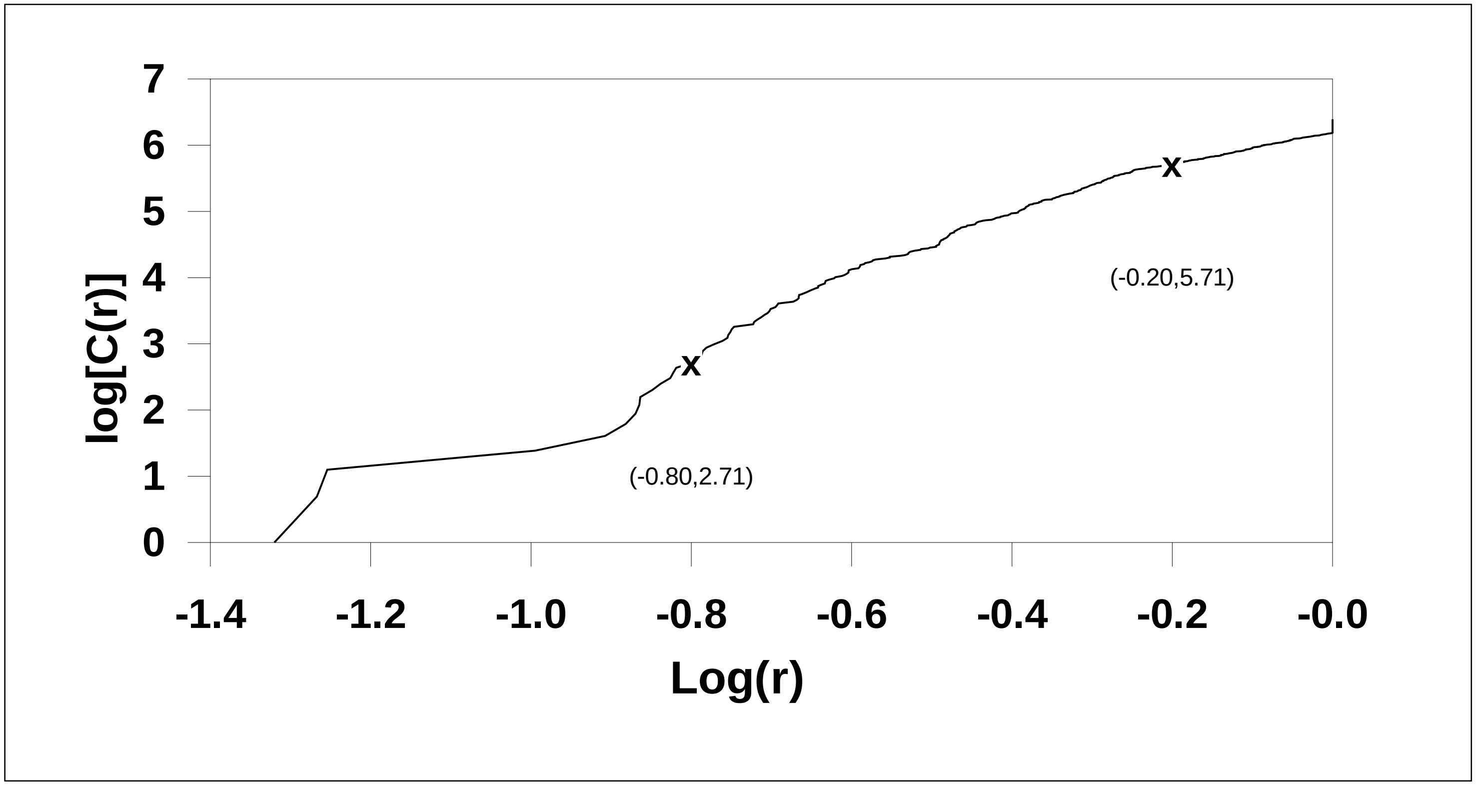
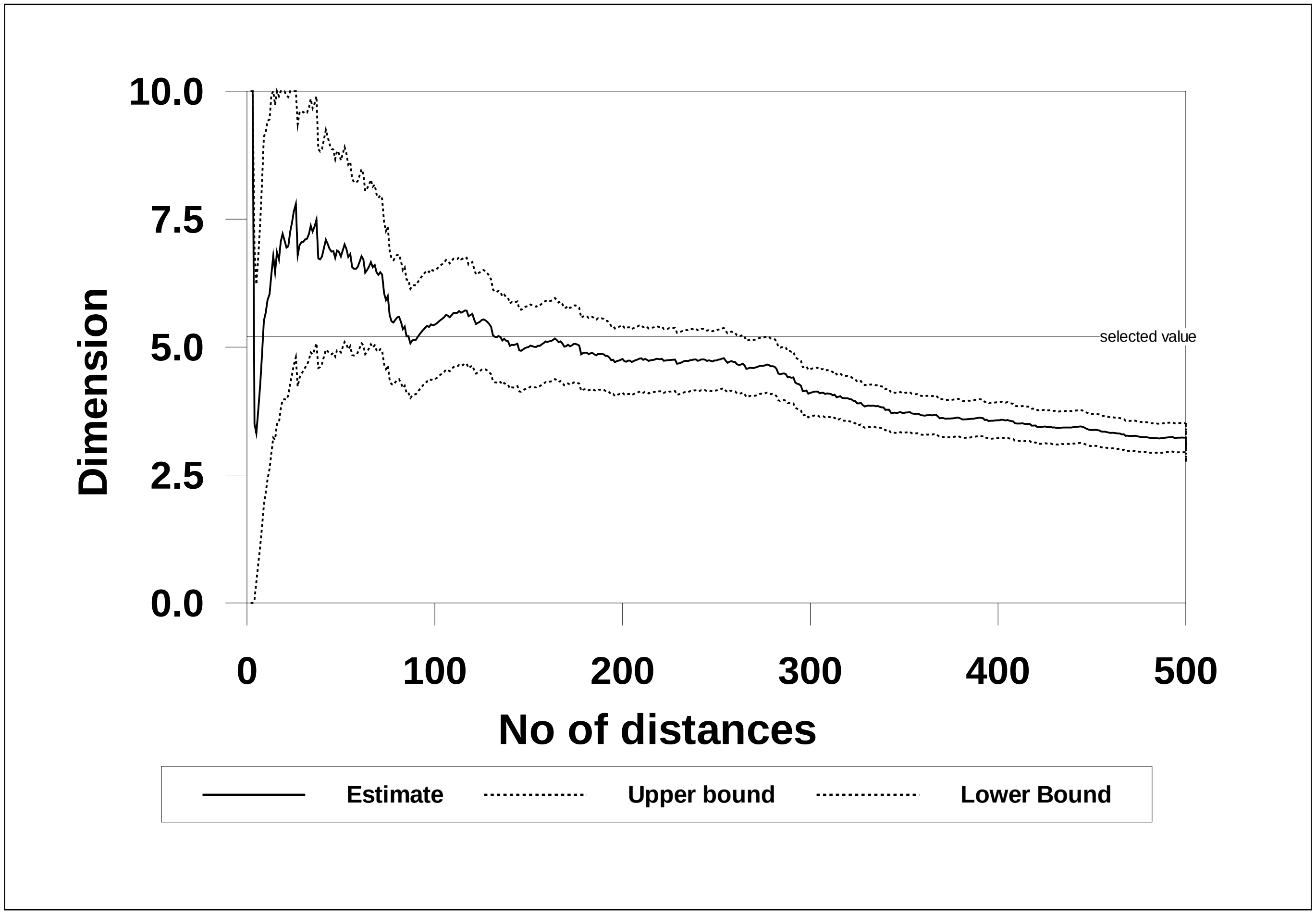


Figure 2. Takens estimates for Easterly data (estimates for first 500 out of total 1891 distances are shown) with 95% confidence interval



The number of dimensions needed to describe the data then using this non-linear approach is much less than the number indicated by the principal component analysis approach. Unfortunately the theory does not allow us to construct the underlying variables or determine the relationship of each variable in our dataset to these dimensions. Instead we search for combinations of five indicators from the twenty-three available which have the largest dimension as we aim to describe the entire development data set with as few variables as possible avoiding redundancy and multicollinearities.[6] The combination that gives the highest dimension using the Takens approach above and a 95 % confidence interval is *Armed Forces* (Easterly area of life quality: POLITICAL INSTABILITY AND WAR), *Cars per capita* (TRANSPORT AND COMMUNICATIONS), *Female to male average schooling years (age 26+)* (INEQUALITY ACROSS CLASS AND GENDER), *Total gross enrolment ratio for secondary education* (EDUCATION) and *Total gross enrolment ratio for higher education* (EDUCATION). However the confidence interval associated with this estimate is very large so it is more instructive to look more generally at the highest results. A summary of results is given in column 1 of Table 4 for the indicators which combine to give the 100 highest dimension estimates broken down by Easterly area of life quality. All of these estimates include a dimension of 5 within their 95% confidence interval.

[6] Calculations were performed in MATLAB and took approximately three hours to perform dimension calculations for all 33649 combinations.

TABLE 4: *Breakdown by Easterly area of life quality indicators which combine to give the 100 highest dimension estimates*

	Easterly area	95% CI		99% CI	
		appearances	**rescaled**	**appearances**	**rescaled**
		(1)	**(2)**	**(3)**	**(4)**
1	Individual Rights and Democracy	21	21	27	27
2	Political instability and war	31	16	20	10
3	Education	93	19	103	34
4	Health	103	26	111	28
5	Transport and communications	67	17	65	13
6	Inequality across class and gender	135	34	126	21
7	"Bads"	34	17	31	4
8	Income	16	16	17	2
	Total	**500**		**500**	

Note : Columns (2) and (4) are formed by dividing the number of appearances in column (1) and (3) by the number of indicators used for each area of life quality (e.g. 4 for health : Life expectancy at birth, Daily calorie intake, daily protein intake (grams), Mortality – Infant)

It can be see that Health, Education and Inequality measures feature most prominently even when rescaling for the number of indicators used in the analysis for each area (column 2). On rescaling the Individual Rights and Democracy area becomes the third most frequent area of life quality. It would appear that the addition of an income per capita variable does not give us much any additional information not already included in other development measures. A measure from any of the remaining areas of life quality could then form a basket of five indicators spanning the development space. Using 99% confidence intervals in the Takens estimates makes these conclusions even more apparent (columns 3 and 4) with Health, Education, Inequality and Individual Rights and Democracy being the four most frequent areas and Transport and Communications becoming the fifth most frequent area when rescaled.

An appropriate multidimensional measure of wellbeing could then be formed with at least representative indicators from each of the domains: Health, Education, Inequality and Individual Rights and Democracy. These are seen to be describing different aspects of development and it is therefore not correct to combine indicators of distinct dimensions into a single measure. Instead assessment of the level of development in a country should consider each dimension of development in turn.

The Easterly (1999) data used in this study could also be updated for more recent years. It would also then be necessary to expand the dataset to include areas of development not originally covered. Identification of supplementary sources of data as well as superior data sources for areas already identified by Easterly has been left to further work. Carrying out a number of analyses at different points in time would also allow us to examine the consistency of results under different geopolitical circumstances.

5 Conclusion

This paper introduces a non-linear technique to estimate the dimension of development. The CMEPSP argued that well-being is multidimensional and identified eight distinct dimensions. This paper has explored the dimensionality of international development data to decide empirically what dimension the development space takes. It would appear that the 23 development variables assembled which do not cover all the categories identified by the CMEPSP but cover many of the generally accepted categories of indicators pertinent to development, span only five dimensions. Potential candidates were sought that would span the empirical development space. From the analysis, variables representing the Health, Education, Inequality and Individual Rights would provide a broad picture of development while income per capita adds little extra information. As the CMEPSP highlighted, however, there are other areas such as social wellbeing, crime and housing for which we do not yet have internationally comparable data sources. Adding measures from these areas may add extra dimensions to the development space.

Appendix

Formally a ball of radius r is drawn on each data point, X_j , and the proportion of all data points within that ball is calculated ($B_{X_j}(r)$). If the points are evenly dispersed on a manifold of dimension υ , in general

$$B_{X_j}(r) \sim r^{\upsilon} \tag{1}$$

As there may be local anomalies in the distribution of data points, we can then find the average dimension over the manifold. We first find the average of $B_{X_j}(r)$ over all all points and label it C(r):

$$C(r) = \frac{1}{N}\sum_{j=1}^{N} B_{X_j}(r) = \frac{1}{N}\sum_{j=1}^{N}\frac{\#\left\{X_i \mid X_i \neq X_j \text{ and } \left\|X_i - X_j\right\| \leq r\right\}}{N-1} \tag{2}$$

Since C(r) scales with r^{υ} , we have log C(r) = a + υ log r. The correlation dimension is then defined as

$$\upsilon = \lim_{r,s \to 0^{+}} \frac{\log C(r) - \log C(s)}{\log r - \log s} \tag{3}$$

Takens (1985) provides an alternative approach based on a maximum likelihood estimator. Choosing a boundary radius R_0 and taking the distance between any two points, $R_i = \left\|X_k - X_j\right\|$, he shows that the random variable $S_i = -\log(R_i / R_0)$ for $R_i < R_0$ has a probability density given by the exponential function with parameter υ . [7] From the properties of this probability distribution, it then follows that the maximum likelihood estimator of υ is given by the reciprocal of the mean of S_i ,

---

[7] In this derivation, the norm is defined in such a way that the furthest distance between two points is 1.

$\hat{\upsilon} = \frac{1}{S_i}$ and the standard error is then estimated by $\hat{\sigma} = \hat{\upsilon} / \sqrt{N^*}$ where N* is the number of distances for which $R_i < R_0$. An estimate, $\hat{\upsilon}$, and a confidence interval can be then determined for the dataset for a given choice of boundary radius R_0.

Bibliography

Adelman, I. and Morris, C.T. (1967), *Society, politics & economic development; a quantitative approach*, Johns Hopkins Press, Baltimore.

Amemiya, Y. and Anderson, T.W. (1990), 'Asymptotic Chi-square tests for a large class of factor analysis models', *Annals of Statistics*, Vol. 18, No.3, pp. 1453-1463.

Anand S. and A. Sen (2000), 'Human Development and Economic Sustainability', World Development, Vol. 28, No. 12, pp. 2029-2049.

Barro, R. (1996), 'Health and Economic Growth', paper prepared for Program For Public Policy and Health, Pan American Health Organisation

Bartlett, M.S. (1950), 'Tests of significance in factor analysis', *British Journal of Psychology*, Vol. 3, Part 2, pp. 77-85.

Bloom, D. E., and Canning, D. (2007) 'The Preston Curve 30 Years On: Still Sparking Fires', *International Journal of Epidemiology*, Vol. 36, No. 3, pp. 498-499.

Camastra, F. and Vinciarelli, A. (2001), 'Intrinsic dimension estimation of data: An approach based on Grassberger-Procaccia's algorithm', *Neural Processing Letters*, Vol.14, No.1, pp 27-34.

CMEPSP (2009), *Report of the Commission on the Measurement of Economic Performance and Social Progress*, www.stiglitz-sen-fitoussi.fr/en/index.htm.

Commission on International Development (1969), *Partners in Development : report*, Praeger.

Comrey, A. L. and Lee, H. B. (1992), *A First Course in Factor Analysis*, Erlbaum, New Jersey.

Dasgupta, P. and Weale, M. (1992), 'On measuring the quality of life', *World Development*, Vol. 20, No. 1, pp. 119-31.

Easterly, W. (1999), 'Life during Growth', *Journal of Economic Growth, Vol.* 4, No. 3, pp. 239-76.

Fedderke, J. and Klitgaard, R. (1998), 'Economic growth and social indicators: An exploratory analysis', *Economic Development and Social Change, Vol.* 46, pp. 455-89.

Gorsuch, R.L. (1974), *Factor Analysis*, Saunders, Philadelphia.

Grassberger, P. and Procaccia, I., (1983), ‘Measuring the strangeness of strange attractors’, *Physica D*, Vol. 9, pp. 189-208.

Harman, H.H. (1976), *Modern Factor Analysis*, 3rd ed., University of Chicago Press, Chicago.

Hicks, D. A. (1997), ‘The inequality-adjusted human development index: A constructive proposal’, *World Development*, Vol. 25, No. 8, pp 1283-1298

Jolliffe, I. T. (1986), *Principal components analysis*. Springer-Verlag, New York, New York, USA.

Kuznets, S. (1955),’Economic Growth and Income Inequality’, *American Economic Review,* Vol. 45, pp.1-28.

Lawley, D.N. (1940), ‘The estimation of factor loadings by the method of maximum likelihood’, *Proceedings of the Royal Society of Edinburgh*, Vol. 60, pp. 64-82.

Malthouse, E.C. (1998), ‘Limitations of Nonlinear PCA as performed with Generic Neural Networks’, *IEEE Trans Neural Networks*, Vol. 9, No. 1, pp. 165-173.

McDonald, R.P. (1967a), ‘Factor interaction in nonlinear factor analysis’, *British Journal of Mathematical and Statistical Psycholog*y, Vol. 20, pp. 205-215.

McDonald, R.P. (1967b), ‘Numerical methods for polynomial models in nonlinear factor analysis’, *Psychometrika*, Vol. 32, pp. 77-112.

McDonald, R.P. (1979), ‘The simultaneous estimation of factor loadings and scores’, *British Journal of Mathematical and Statistical Psychology*, Vol. 32, pp. 212-228.

Peres-Neto, P.R., D.A. Jackson, and K.M. Somers (2005), ‘How many principal components? Stopping rules for determining the number of non-trivial axes revisited’, *Computational Statistics and Data Analysis* Vol. 49 pp. 974-997.

Preston, S. (1975), ‘The Changing Relation between Mortality and Level of Economic Development’, *Population Studies*, Vol. 29, No. 2, pp. 231-248.

Ranis G., Stewart F., and A. Ramirez (2000), ‘Economic Growth and Human Development’, *World Development*, Vol. 28, No 2, pp 197-219

Sen, A. (1999), *Development as Freedom*, Oxford University Press, Oxford and New York..

Smith, L. A. (1997), ‘The maintenance of uncertainty’, in G. Cini Castagnoli and A. Provenzale, eds., *Past and Present Variability in the Solar-Terrestial System: Measurement, Data Analysis and Theoretical Models*, volume CXXXIII of International School of Physics\ Enrico Fermi, pp. 177-246, Bologna, 1997. Il Nuovo Cimento.

Takens, F. (1985), ‘On the numerical determination of the dimension of an attractor,’ in *Dynamical Systems and Bifurcations*, Groningen, 1985, B. L. J. Braaksma, H. W. Broer, and F. Takens, eds., Vol. 1125 of Lecture Notes in Mathematics, Springer-Verlag, Berlin.

Theiler. J. (1986), ‘Spurious dimension from correlation algorithms applied to limited time-series data’, *Physical Review A*, Vol. 34, No. 3, pp. 2427-32.

Theiler. J. (1990), ‘Estimating fractal dimension’, *Journal of the Optical Society of America*, Vol. 7, No. 6, pp. 1055-73.

World Bank (2001), *World Development Report* 2000/01- *Attacking Poverty*, New York.

Yalcin, I. and Amemiya, Y. (2001), ‘Nonlinear Factor Analysis as a Statistical Method’, *Statistical Science*, Vol. 16, No. 3, 275–94.