Trends in Human Development Index of European Union

Serpil Aktaş, Derya Karagöz

Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey

spxl@hacettepe.edu.tr, deryacal@hacettepe.edu.tr

Abstract: The Human Development Index is a measure of development index calculated from life expectancy, literacy, education and standards of living. In this paper, Human Development Index data of European Union are collected for periods 1980-2013, and analyzed using Generalized Estimating Equations to investigate whether there is a trend through the years. Generalized Estimating Equations method is often employed to analyze longitudinal and other correlated response data does not require any multivariate distribution assumption.

Key Words: Human Development Index, Generalized Estimating Equations, European Union

Introduction

Longitudinal studies involve repeated observations of the same items over long periods of time and, often arise in psychology, sosiology, education, medical sciences to analyze developmental trends across time (Diggle et. al. 1994 ;Diggle et. al. 2002; Twisk, 2002). Variable for each subject observed repeatedly over time cause dependency structure between variables. Correlated data are particularly very common in educational and more generally in social science researches. Longitudinal studies also allow researches to reveal the short from long-term phenomena, such as poverty, infant mortality rate, economic development ect. Ignoring the dependency of the observations will overestimate the standard errors of the the time-dependent predictors. This means that we also ignore the between-subject variability. Repeated measure ANOVA is used for longitudinal studies because of simplicity, but it has some limitations. For instance, it assumes categorical predictors; does not take the time-dependent covariates into account; assumes that subjects are measured at the same and equally spaced time intervals and it requires restrictive assumptions about the correlation structure. Hence, Generalized Linear Model approach to longitudinal studies has been growing in recent years. Generalized Estimation Equatios (GEE) methodology were developed by Liang and Zeger (1986); Zeger and Liang (1986) as the extension of the Generalized Linear Models (GLM) (McCullagh and Nelder,1989) for the data in longitudinal form.

In this paper, it is aimed to model the Human Development Index (HDI) data of the European Union (EU) countries via GEE. HDI has become an important alternative measure of development. The HDI data are collected between the 2005-2010 periods, and analyzed using GEE to investigate whether there is the trend through the years.

The simple model can be written as

$$HDI{=}\beta_0 + \beta_1 TIME + CORR + ERROR. \qquad (1)$$

Where, the intercept β_0 and the slope β_1 are unknown parameters. Time is treated as a continuous variable and measured in years. The aim is test the trend over years.

Generalized Estimating Equations

The idea of GEE was first introduced by Liang and Zeger (1986); Zeger and Liang (1986). GEE methodology fits a model to repeated categorical responses, that could be correlated and clustered responses. The advantages of GEE can be also summarized as: It does not require a multivariate distribution; estimates of model parameters are valid even if misspecification of the covariance structure; it is preferred to Maximum Likelihood (ML) because of its computational simplicity. In recent years, GEE has been a popular alternative to maximum likelihood.

Let Y_{ij} be the jth outcome for the ith subject, where we assume that observations on different subjects are independent; the association between outcomes is observed on the same subject. Y_i denotes a response vector for each subject j and Y is the vector of measurement off all units.

$$Y_i = (Y_{i1}, Y_{i2}, \ldots, Y_{in}), \text{ i = 1,..., N; j = 1,..., } n_i$$

Marginal response is defined as $\mu_i = E(Y_i)$. Linear combination of the covariates are $g(\mu_i)$=$X_i\beta$. Where , X_i is a $n_i \times p$ matrix of covariates ; β is a p × 1 vector of unknown regression coefficients and g(·) is the link function.

For unknown parameter vector β , Equation (2) is given as

$$U(\beta) = \sum_{i=1}^{N} \frac{\partial \mu_i}{\partial \beta} V_i^{-1}(Y_i - \mu_i) = 0 \qquad (2)$$

where, V_i is the $n_i x n_i$ variance covariance matrix, $V_i = A_i^{\frac{1}{2}} R_i(\alpha) A_i^{\frac{1}{2}} / \phi$. A_i is a diagonal matrix with elements Var (Y_{ij}) and $R_i(\alpha)$ is referred as working correlation matrix (Liang and Zeger, 1986; Zeger and Liang, 1986). ϕ is the over-dispersion parameter. Working correlation matrix choices are: Independent, Exchangeable, Autoregressive , M-dependent and Unstructured. But, the advantage of GEE is that it is fairly robust against a misspecification of correlation matrix (Hin and Wang, 2009).

Solution the Equation (2) gives the parameter estimates. In the GEE procedure, ordinary linear regression analysis is firstly performed, assuming the observations within subjects are independent. Then, residuals are calculated from the ordinary model and a working correlation matrix is estimated from these residuals. Then the regression coefficients are estimated, correcting for the correlation.

Human Development Index

HDI is an aggregated measure of development index calculated from life expectancy, literacy, education and standards of living (UNDP, 2011). Until 2010, the HDI had been defined as a simple arithmetic average of normalized indices in the dimensions of health, education and income:

$$HDI = \frac{1}{3}\left(H_{health} + H_{education} + H_{living\ standards}\right). \qquad (3)$$

Each of these indices are normalized indicators of achievements for each dimensions and based on life expectancy (LE) , GDP per capita (GDP), literacy (LIT) and the gross enrolment ratio (GER).

Where, the subindices:

$$H_{health} = \frac{(LE - LE_{min})}{(LE_{max} - LE_{min})} \qquad (4)$$

$$H_{education} = \frac{1}{3}\left(\frac{GER-GER_{min}}{GER_{max}-GER_{min}}\right) + \frac{2}{3}\left(\frac{LIT-LIT_{min}}{LIT_{max}-LIT_{min}}\right) \quad (5)$$

$$H_{living\ standards} = \frac{(ln(GDP)-ln(GDP_{min}))}{(ln(GDP_{max})-ln(GDP_{min}))} \quad (6)$$

Hence, the indices are normalized using given upper and lower bounds which were defined in the 2009 report. The 2010 Human Development Report presented some changes in the HDI as

$$HDI = \sqrt[3]{H_{health} \,.\, H_{education} .\, H_{living\ standards}} \quad (7)$$

Life expectancy still represents the health dimension, while Gross National Income (GNI) replaces GDP as the measure for living standards. Mean years of schooling (MYS) and expected years of schooling (EYS) now are the new indicators of the education dimension.

$$H_{health} = \frac{(LE-LE_{min})}{(LE_{max}-LE_{min})} \quad (8)$$

$$H_{education} = \left(\frac{MYS-MYS_{min}}{MYS_{max}-MYS_{min}}\right)\left(\frac{EYS-EYS_{min}}{EYS_{max}-EYS_{min}}\right) \quad (9)$$

$$H_{living\ standards} = \frac{(ln(GNI)-ln(GNI_{min}))}{(ln(GNI_{max})-ln(GNI_{min}))} \quad (10)$$

The HDI enables to researchers to detect the changes in development levels over time and to compare development levels in other countries. The value of HDI vary between 0 and 1. The interpretation of HDI can be made as:

HDI $\geq$ 0.800 is high development,

HDI 0.500—0.799 is medium development,

HDI $<$ 0.500 is low development

(UNDP, 2011). High HDI means more prosperity and achievement on the developmental factors.

Analysis of HDI data for the Member Counties of European Union

United Nations Development Program has been calculating HDI for the member countries. This paper's goal is to asses the changes HDI for the member countries of EU over nine years . The human development indices of the countries were obtained from a Human Development Report (Table 1). Data set was downloaded from the United Nations Development Program web page (http://hdr.undp.org/en/data).

Recall the member states of the European Union: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and United Kingdom. These countries are included in the analysis.

Table 1: Human Development Index trends, 1980-2013.

		Human Development Index (HDI)								
	Country	**1980**	**1990**	**2000**	**2005**	**2008**	**2010**	**2011**	**2012**	**2013**
1	Netherlands	0,783	0.826	0.874	0.888	0.901	0.904	0.914	0.915	0.915
2	Germany	0.739	0.782	0.854	0.887	0.902	0.904	0.908	0.911	0.911
3	Denmark	0.781	0.806	0.859	0.891	0.896	0.898	0.899	0.900	0.900
4	Ireland	0.734	0.775	0.862	0.890	0.902	0.899	0.900	0.901	0.899
5	Sweden	0.776	0.807	0.889	0.887	0.891	0.895	0.896	0.897	0.898
6	United Kingdom	0.735	0.768	0.863	0.888	0.890	0.895	0.891	0.890	0.892
7	France	0.722	0.779	0.848	0.867	0.875	0.879	0.882	0.884	0.884
8	Austria	0.736	0.786	0.835	0.851	0.868	0.877	0.879	0.880	0.881
9	Belgium	0.753	0.805	0.873	0.865	0.873	0.877	0.880	0.880	0.881
10	Luxembourg	0.729	0.786	0.866	0.876	0.882	0.881	0.881	0.880	0.881
11	Finland	0.752	0.792	0.841	0.869	0.878	0.877	0.879	0.879	0.879
12	Slovenia	.	0.769	0.821	0.855	0.871	0.873	0.874	0.874	0.874
13	Italy	0.718	0.763	0.825	0.858	0.868	0.869	0.872	0.872	0.872
14	Spain	0.702	0.755	0.826	0.844	0.857	0.864	0.868	0.869	0.869
15	Czech Republic	.	0.762	0.806	0.845	0.856	0.858	0.861	0.861	0.861
16	Greece	0.713	0.749	0.798	0.853	0.858	0.856	0.854	0.854	0.853
17	Cyprus	0.661	0.726	0.800	0.828	0.844	0.848	0.850	0.848	0.845
18	Estonia	.	0.730	0.776	0.821	0.832	0.830	0.836	0.839	0.840
19	Lithuania	.	0.737	0.757	0.806	0.827	0.829	0.828	0.831	0.834
20	Poland	0.687	0.714	0.784	0.803	0.817	0.826	0.830	0.833	0.834
21	Slovakia	.	0.747	0.776	0.803	0.824	0.826	0.827	0.829	0.830
22	Malta	0.704	0.730	0.770	0.801	0.809	0.821	0.823	0.827	0.829
23	Portugal	0.643	0.708	0.780	0.790	0.805	0.816	0.819	0.822	0.822
24	Hungary	0.696	0.701	0.774	0.805	0.814	0.817	0.817	0.817	0.818
25	Croatia	.	0.689	0.748	0.781	0.801	0.806	0.812	0.812	0.812
26	Latvia	.	0.710	0.729	0.786	0.813	0.809	0.804	0.808	0.810
27	Romania	0.685	0.703	0.706	0.750	0.781	0.779	0.782	0.782	0.785
28	Bulgaria	0.658	0.696	0.714	0.749	0.766	0.773	0.774	0.776	0.777

Descriptive statistics for HDIs are given in Table 2. From Table 2, it can be seen that the very high human development group over nine years corresponds to Netherland, Germany, Denmark, Ireland, Sweden, United Kingdom Belgium and Luxembourg.

Table 2: Descriptive statistics by country

Country	N	Minimum	Maximum	Mean	Std. Deviation
Austria	9	0.7360	0.8810	0.8437	0.0509
Belgium	9	0.7530	0.8810	0.8541	0.0448
Bulgaria	9	0.6580	0.7770	0.7426	0.0432
Croatia	8	0.6890	0.8120	0.7826	0.0438
Cyprus	9	0.6610	0.8500	0.8056	0.0674
Czech Republic	8	0.7620	0.8610	0.8388	0.0362
Denmark	9	0.7810	0.9000	0.8700	0.0457
Estonia	8	0.7300	0.8400	0.8130	0.0395
Finland	9	0.7520	0.8790	0.8496	0.0467
France	9	0.7220	0.8840	0.8467	0.0575
Germany	9	0.7390	0.9110	0.8664	0.0636
Greece	9	0.7130	0.8580	0.8209	0.0550
Hungary	9	0.6960	0.8180	0.7843	0.0506
Ireland	9	0.7340	0.9020	0.8624	0.0633
Italy	9	0.7180	0.8720	0.8352	0.0568
Latvia	8	0.7100	0.8130	0.7836	0.0407
Lithuania	8	0.7370	0.8340	0.8061	0.0378
Luxembourg	9	0.7290	0.8820	0.8513	0.0553
Malta	9	0.7040	0.8290	0.7904	0.0458
Netherlands	9	0.7830	0.9150	0.8800	0.0462
Poland	9	0.6870	0.8340	0.7920	0.0547
Portugal	9	0.6430	0.8220	0.7783	0.0622
Romania	9	0.6850	0.7850	0.7503	0.0410
Slovakia	9	0.7470	0.8300	0.7180	0.2708
Slovenia	8	0.7690	0.8740	0.8514	0.0380
Spain	9	0.7020	0.8690	0.8282	0.0598
Sweden	9	0.7760	0.8980	0.8707	0.0457
United Kingdom	9	0.7350	0.8950	0.8569	0.0610

Figures (1-28) below show trends in HDI values of EU countries separately, during the period 1990 to 2013. It can be clearly seen that the HDIs increased considerably for the years from 1990 to 2013 for all countries. Some countries 2004 there has been a steady increase such as Netherland, Germany, France and Austria.

Cyprus is the only country where it was observed a downward trend slightly in recent years. The highest level of progression in HDI is observed in for instance, for Portugal and Germany. They strongly move up through 2000's. The lowest HDI values are for Bulgaria and Romania with overall means 0.7426 and 0,7850, respectively. A consistent increase draw the attention particularly in Spain, Chezh Rebublic, France and Austria.

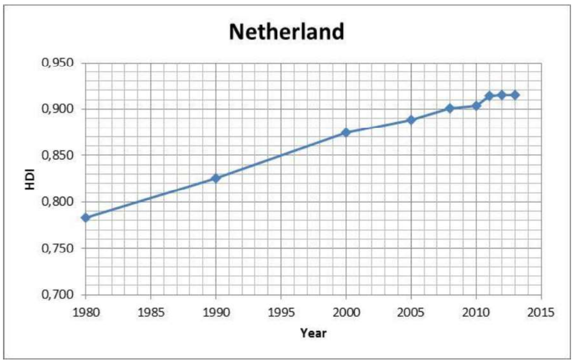

Figure 1: Human Development Index of Netherland, 1980-2013

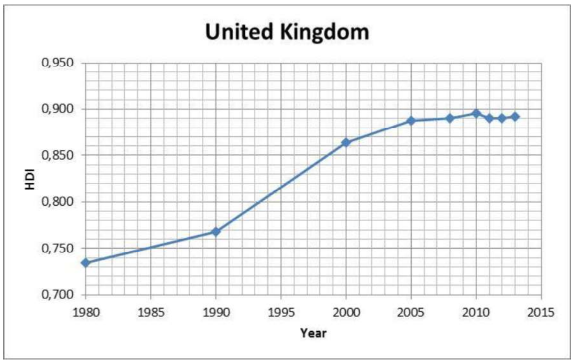

Figure 2: Human Development Index of United Kingdom, 1980-2013

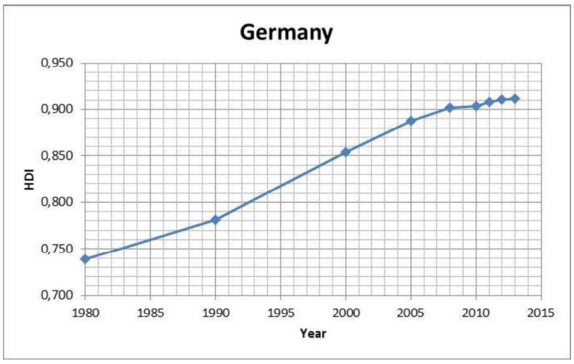

Figure 3: Human Development Index of Germany, 1980-2013

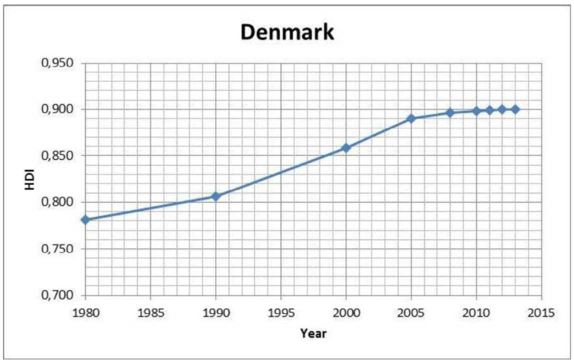

Figure 4: Human Development Index of Denmark, 1980-2013

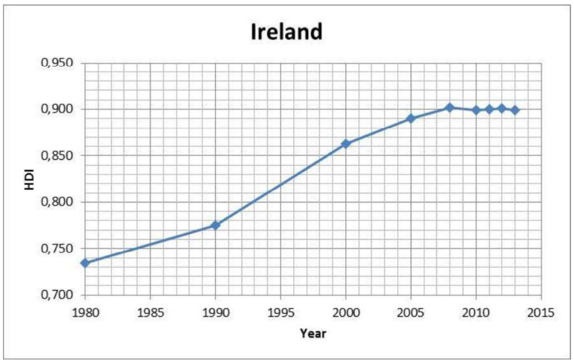

Figure 5: Human Development Index of Ireland, 1980-2013

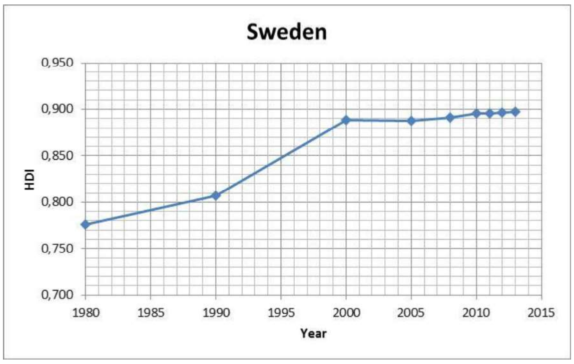

Figure 6: Human Development Index of Sweden, 1980-2013

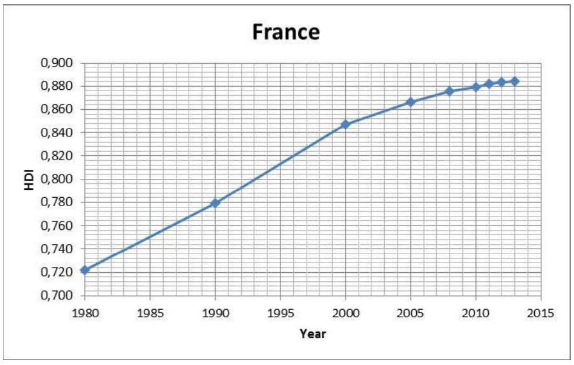

Figure 7: Human Development Index of France, 1980-2013

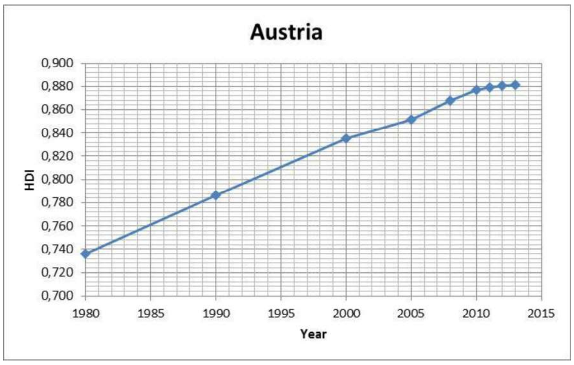

Figure 8: Human Development Index of Austria, 1980-2013

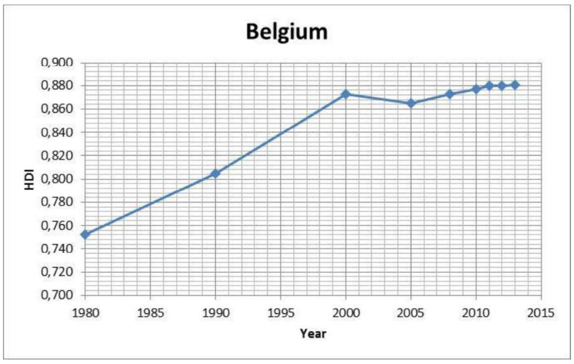

Figure 9: Human Development Index of Belgium, 1980-2013

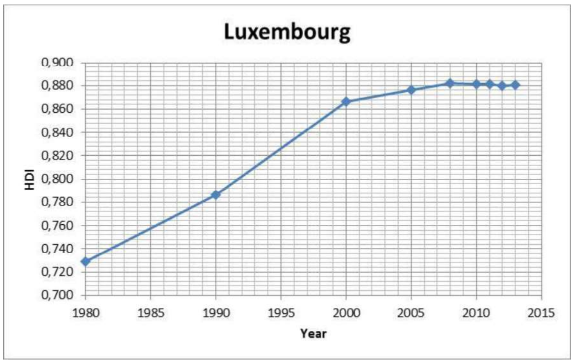

Figure 10: Human Development Index of Luxembourg, 1980-2013

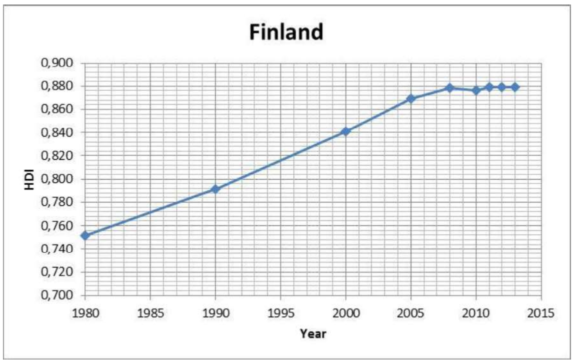

Figure 11: Human Development Index of Finland, 1980-2013

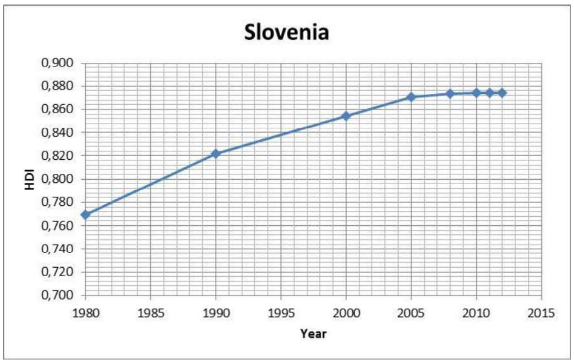

Figure 12: Human Development Index of Slovenia, 1980-2013

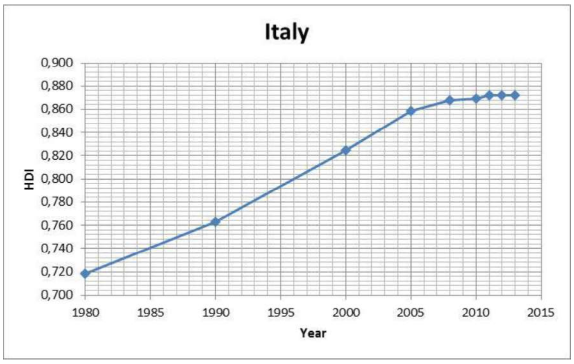

Figure 13: Human Development Index of Italy, 1980-2013

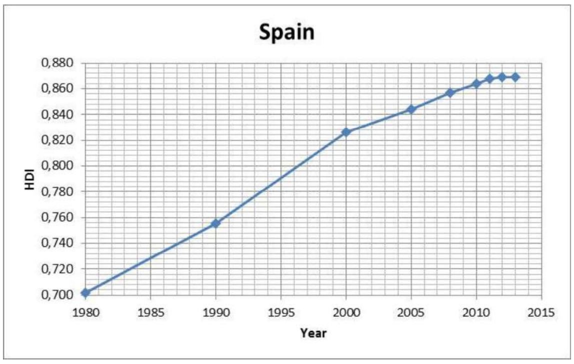

Figure 14: Human Development Index of Spain, 1980-2013

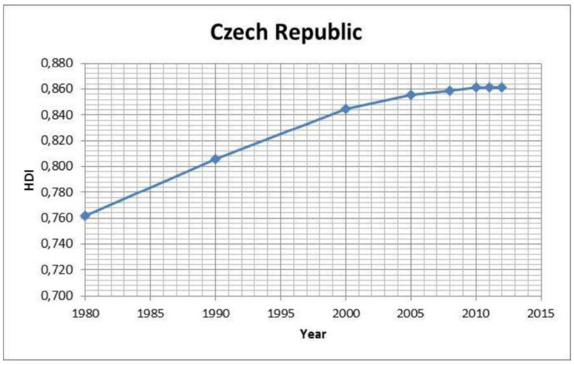

Figure 15: Human Development Index of Czech Rep., 1980-2013

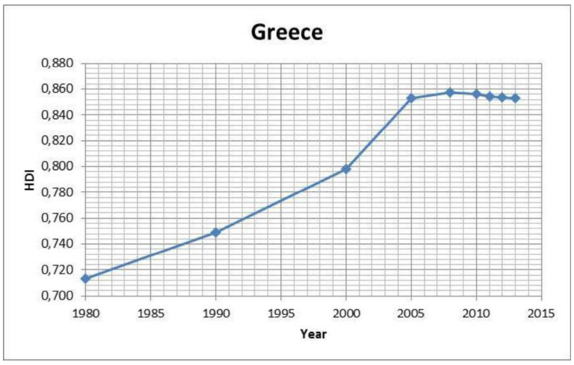

Figure 16: Human Development Index of Greece, 1980-2013

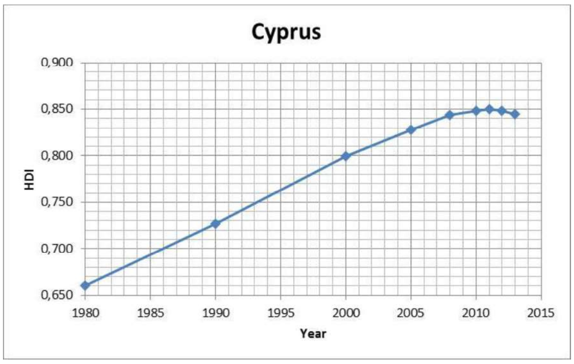

Figure 17: Human Development Index of Cyprus, 1980-2013

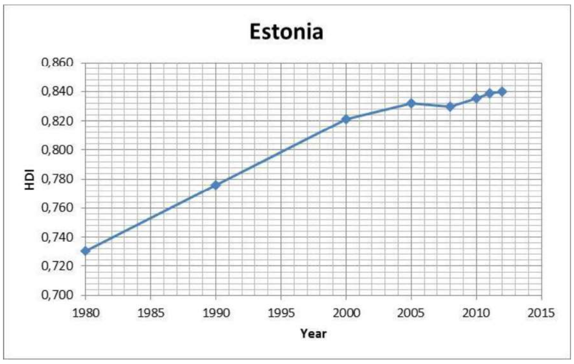

Figure 18: Human Development Index of Estonia, 1980-2013

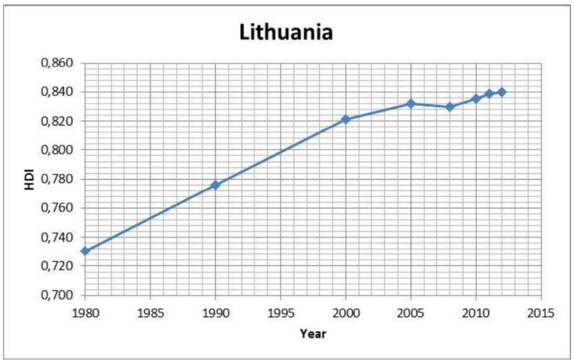

Figure 19: Human Development Index of Lithuania, 1980-2013

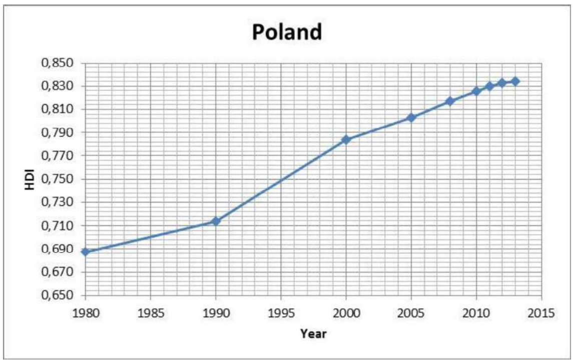

Figure 20: Human Development Index of Poland, 1980-2013

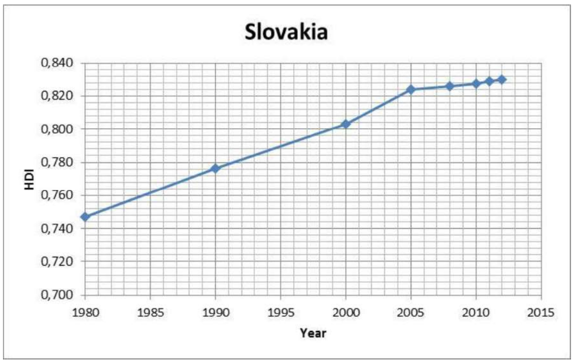

Figure 21: Human Development Index of Slovakia, 1980-2013

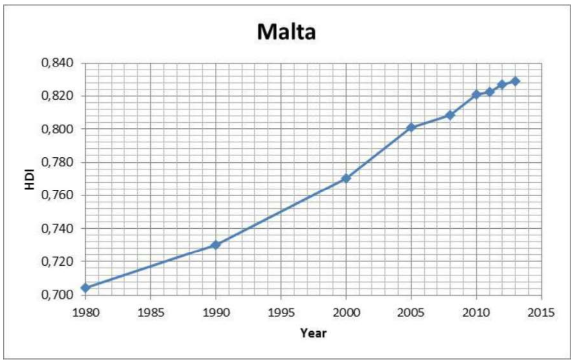

Figure 22: Human Development Index of Malta, 1980-2013

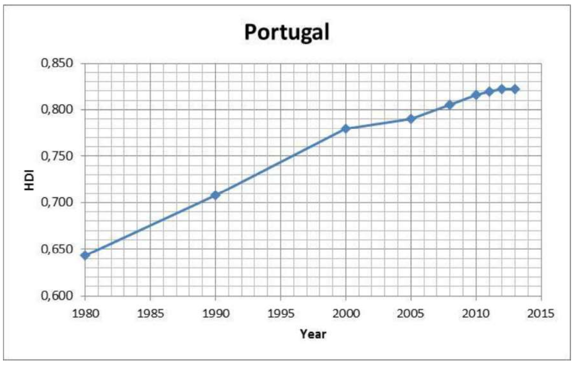

Figure 23: Human Development Index of Portugal, 1980-2013

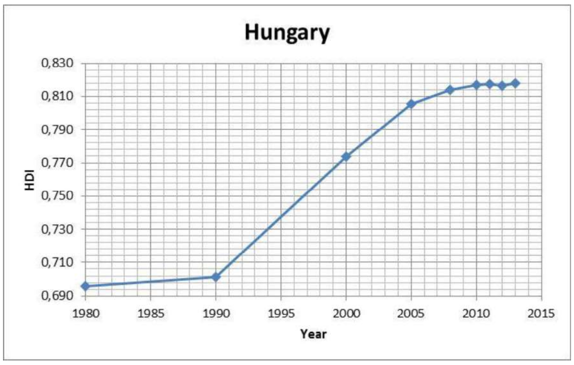

Figure 24: Human Development Index of Hungary, 1980-2013

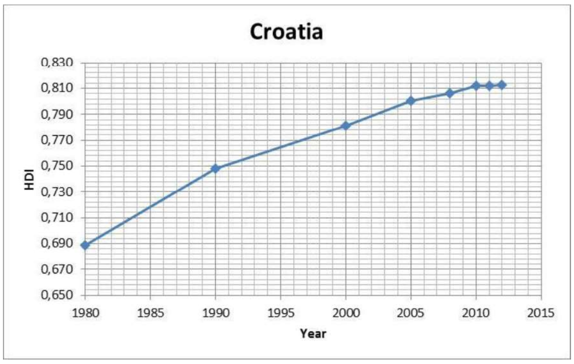

Figure 25: Human Development Index of Croatia, 1980-2013

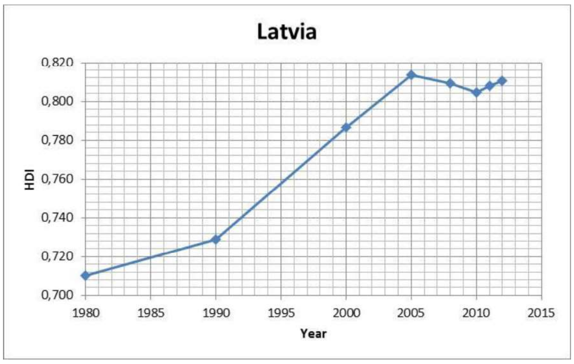

Figure 26: Human Development Index of Latvia, 1980-2013

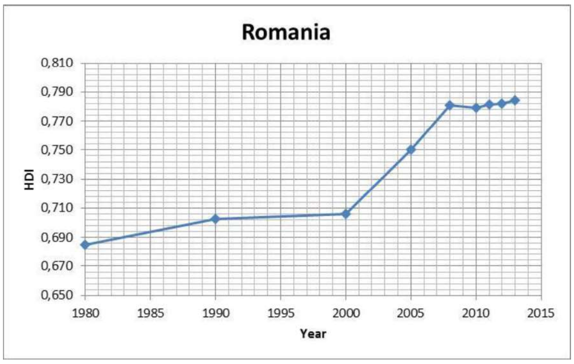

Figure 27: Human Development Index of Romania, 1980-2013

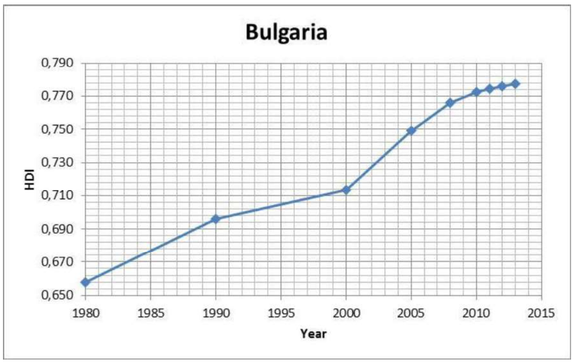

Figure 28: Human Development Index of Bulgaria, 1980-2013

The overall means by years along with their standard errors and 95% confidence interval are given in Table 2 and Figure 29 shows the trend by year. HDI has a steady upward trend after 2008. A sharp increase from 1980 to 2000 and a gradual increase after 2000 can be seen in Figure 3. For all countries except Cyprus, the HDI is the highest in 2013, even though the mean HDI for 2012 seems to equal with the HDI for 2013. Romania started to move up in 2000's. Latvia reached the peak in 2005. Long-term progress can be usefully assessed relative to other countries.

Table 3: Overall means by year

YEAR	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1980	0.719	0.009	0.702	0.737
1990	0.760	0.009	0.742	0.778
2000	0.821	0.011	0.798	0.844
2005	0.845	0.010	0.824	0.865
2008	0.856	0.009	0.837	0.875
2010	0.860	0.009	0.842	0.878
2011	0.862	0.009	0.844	0.880
2012	0.863	0.009	0.845	0.881
2013	0.863	0.009	0.845	0.881

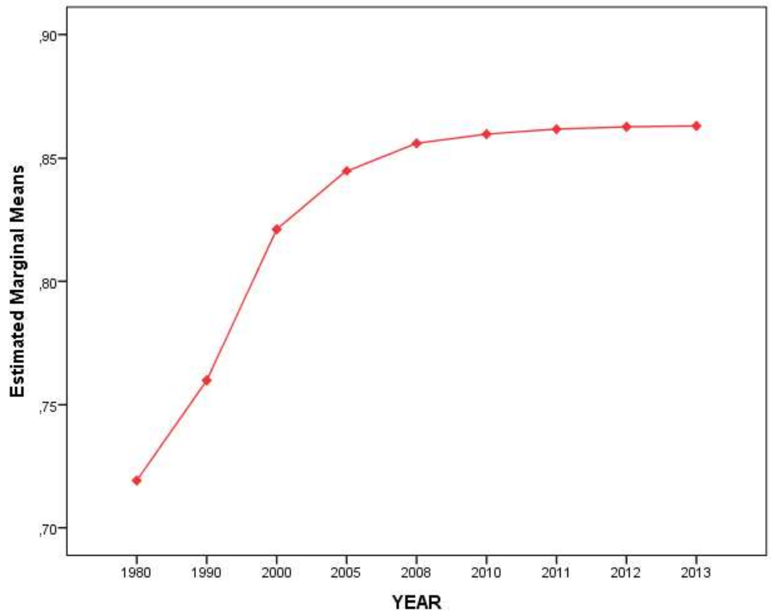

Figure 29: Marginal means of HDI by year

IBM SPSS 20 was used for the analysis. Generalized Linear Model menu includes techniques of Generalized Linear Models and Generalized Estimating Equations. Table 3 summaries the result of GEE analyses with an unstructures correlation structure. Test of model effects evaluates each of the model variables with the appropriate degrees of freedom. Intercept and year are statistically significant (P<0.01).

Table 4: Test of model effects

Source	Type III		
	Wald Chi-Square	df	Sig.
Intercept	1095.364	1	0.000
Year	37.905	1	0.000

Table 5 includes the regression coefficients for each of the variables along with standard errors, p-values and 95% confidence intervals for the coefficients and Exp(B). The coefficient for year is 0.024 .The model in Equation (1) can be represented by $\hat{Y}$= 0.705 + 0.024 Year.

This means that the expected change in HDI for a one-unit change in time is 0.024. In other words, the beta parameter can be interpreted as: 1-unit increase in year is associated with a 0.024 increase in HDI and a significant positive beta coefficient here would mean the change in year has changes in HDI correspondingly.

Table 5: Parameter estimates for GEE

Parameter	Beta	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
Intercept	0.705	0.0213	0.663	0.747	1095.364	1	0.000	2.024	1.941	2.110
Year	0.024	0.0039	0.016	0.032	37.905	1	0.000	1.024	1.016	1.032
(Scale)	0.005									

Working correlation matrix across all nine time periods under unstractured covariance matrix assumption is given below (Table 6). A working correlation structure is a correlation matrix for repeated or clustered measurements from each individual. An unstructured working correlation matrix has no explicit pattern. In the GEE method, if the working correlation matrix is correctly specified, the parameter estimates become more reliable.

Table 6: Working correlation matrix

Measurement	Measurement								
	1	2	3	4	5	6	7	8	9
1	1	0.271	0.197	0.158	0.195	0.351	0.525	0.694	0.915
2	0.271	1	0.358	0.315	0.289	0.266	0.249	0.226	0.200
3	0.197	0.358	1	0.661	0.530	0.372	0.209	0.034	-0.134
4	0.158	0.315	0.661	1	0.504	0.348	0.189	0.022	-0.174
5	0.195	0.289	0.530	0.504	1	0.343	0.252	0.152	-0.053
6	0.351	0.266	0.372	0.348	0.343	1	0.343	0.342	0.189
7	0.525	0.249	0.209	0.189	0.252	0.343	1	0.548	0.462
8	0.694	0.226	0.034	0.022	0.152	0.342	0.548	1	0.742
9	0.915	0.200	-0.134	-0.174	-0.053	0.189	0.462	0.742	1

Conclusions

GEEs provide a practical method with good statistical properties to model data that exhibit association but cannot be modeled as multivariate normal. Ordinary linear regression ignores the correlation between subjects but GEE takes into account the dependency of observations by specifying a working correlation structure. The main advantage of GEEs resides in the robust estimation of parameters' standard errors, even when the correlation structure is misspecified.Therefore using GEE would be considered a better alternative for clustered data and outperforms the classical regression. It could be presumably misleading to compare the HDI rankings with those of previously published reports, because the calculation method has changed. United Nations Development Programe data ensure as much cross-country comparability as possible.
However a progress in the HDI can be observed for all countries. During the period between 1980 and 2013, countries experienced different degrees of progress in terms of their HDIs.

Results also suggest that changes in HDI over years are statistically significant. A significant positive coefficient for time would mean the change in year has changes in HDI correspondingly.

References

Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994), *Analysis of Longitudinal Data*, Oxford: Clarendon Press.

Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *Longitudinal data analysis* (2nd ed.). Oxford, UK: Oxford University Press.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using Generalized Linear Models. Biometrika*, 73,* 13–22.

Hin, L.-Y., Wang, Y.-G. (2009). *Working Correlation Structure Identification in Generalized Estimating Equations*, Statistics in Medicine, 28, 642-658.

IBM Corp. Released 2011. *IBM SPSS Statistics for Windows,* Version 20.0. Armonk, NY: IBM Corp.

McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models*, New York: Chapman and Hall.

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42,* 121–130.

United Nations Development Programme (UNDP) (2011). "Human Development Report 2011, Technical Notes."

Twisk, Jos W. R. (2003). Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide. Cambridge University Press.