Population Review

Volume 53, Number 2, 2014

Type: Article pp. 47-55

Group-Based Trajectory Modeling of Longitudinal International Infant Mortality Rates

Authors: Serpil Aktaş and Stephen J. Finch
Affiliations: Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey (Aktaş); Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, NY, USA (Finch)
Corresponding author/address: Serpil Aktaş, Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey; e-mail: spxl@hacettepe.edu.tr

Abstract

The objective of this paper is to estimate the group-based trajectory models for longitudinal infant mortality rate data of United Nations countries. Four linear trajectories were detected. The infant mortality rate has a descending trend over the period 1990-2010 for each of the trajectories. The time varying covariate variable Human Development Index has a strong statistically significant negative linear relationship with Infant Mortality Rate. The trajectory groups were correlated with the wealth classification of the nations.

Keywords

Infant mortality rate, human development index, mixture models

Note: The research underpinning this paper was supported by funds from The Scientific and Technological Research Council of Turkey (TÜBİTAK).

Introduction

Infant mortality is often used as an indicator of the health of societies and is measured by the infant mortality rate (IMR), which is the number of infant deaths per 1,000 live births (World Health Organization, 2011). IMR is also a useful indicator of a country's level of health or development and is a component of the physical quality of life index. The IMR is often related to variables, such as poverty, low education level and low health services.

One important covariate of the IMR is the human development index (HDI). For example, the Pearson correlation coefficient between the HDI and the IMR was -0.921 ($P<0.001$) based on the data used in this paper.

The United Nations HDI is a composite index that aggregates three basic dimensions: an index of health, an index of education, and per capita gross domestic product. These variables measure the socioeconomic development of a country (UNDP, 2011). Lee et al. (1997) found that the HDI is a useful index for both the socioeconomic development of a country and its infant and maternal mortality rates. Since 1990, the United Nations Development Program has published a series of annual Human Development Reports in which the HDI is computed for each country. HDI is defined as a simple arithmetic average of normalized indices in the dimensions of health, education and income (Klugman et al., 2011) — specifically the life expectancy index, education index and income index. When creating these dimension indices, minimum and maximum values are set in order to transform the indicators into indices between 0 and 1. The maximums are the highest observed values in the time series from 1980 to 2011. The minimum values are set at 20 years for life expectancy, at 0 years for both education variables, and at $100 for per capita gross national income (UNDP, 2011).

This paper uses a group-based trajectory modeling to analyze a longitudinal data set that describes 194 countries for the years 1990, 1995, 2000, 2005 and 2010. Nagin (1999, 2005) developed this model, which is a special application of finite mixture modeling. Nagin and Trambley (2001) also presented the group-based method to jointly estimate developmental trajectories of two distinct but theoretically related measurement series. In recent years there have been several studies on this method. Erosheva et.al. (2014), for example, described trajectories of human behavior over time.

This paper employs the group-based trajectory model approach for explaining IMR data of United Nations countries. Data is analyzed by Proc Traj (Jones et.al. 2001), which is a SAS procedure for developmental trajectories. Its purpose is to assess whether there is homogeneity of IMR trajectories after HDI adjustment. If there is heterogeneity, we seek to compare the trajectory classes to the United Nations Report of the country's development. This group-based methodology allows researchers to test whether each country follows the same pattern, or whether there are multiple patterns. IMRs for each country generally decline over time. We hypothesize that there will be multiple patterns of decline after controlling for HDI. We further hypothesize that these patterns are strongly associated with the World Bank income groups. We seek to identify whether there are poorer countries whose improvement in IMR exceeds expectation. If these exist, we seek to identify them and assess what factors are associated with this positive outcome.

In general, we attempt to identify countries whose decline in IMR differs notably from the expectations calculated by the model. Proc Traj estimates a regression model for each discrete trajectory.

Following Arrandale et al. (2006), we address the following five questions:

1. Are there multiple patterns of change in the outcome?
2. How many patterns of change are there in the outcome?
3. What is the shape of the change over time?
4. What predicts membership in each of these groups?
5. What are the characteristics that differ (or are similar) between the different groups?

HDI is added to the model because the model extension allows the statistical testing of hypotheses about whether such individual-level characteristics distinguish trajectory group membership.

Methods

The group-based trajectory model assumes that the underlying population is composed of distinct groups defined by their trajectories.

Let $Y_i = (y_{i1}, y_{i2,...}, y_{iT})$ represent the longitudinal IMR measurements on country i over t=5 time points , t=1,2,…,5 (1990, 1995, 2000, 2005, 2010). The group based trajectory model is a mixture of J underlying trajectory groups such that

$$P(Y_i) = \sum_j \pi_j P(Y_i \mid J = j) \quad (1)$$

where $P(Y_i \mid J = j)$ is the probability of Y_i, given membership in group j , and π_j is the probability of group j (Nagin, 2005).

Proc Traj allows a country trajectory model that can be up to a fourth order polynomial. Proc Traj can also model both linear and non-linear trajectories within the same model.

We assume that the IMR at each time point can be modeled as a censored normal distribution, called the CNORM model in Proc Traj. We model the IMR as

$$y_{it}^{j} = \hat{\beta}_0^{j} + \hat{\beta}_1^{j} HDI_{it} + \hat{\beta}_2^{j} TIME , \quad (2)$$

where y_{it}^{j} is expected IMR value of country i at time t, given membership in group j, HDI_{it} is the ith country's HDI value at time t.

We used the Bayesian Information Criteria (BIC) to select the number of trajectory groups so that the number of trajectory groups selected the largest BIC value.

We also examined linear and quadratic trends. Since no trajectory group had a significant quadratic association, we reported the models using a linear trend. Proc traj uses the censored normal model (CNORM) for calculating the trajectories.

According to the World Bank income classification (http://data.worldbank.org/country), countries are classified into four income groups as follows: low income countries (GNI per capita of US$1005 or less= 1); lower middle income countries (GNI per capita between US$1006 and US$3,975 =2); upper middle income countries (GNI per capita between US$3,976 and US$12,275= 3); high income countries (GNI above US$12,276 =4). These classifications will be compared to the predicted model groups.

Data

IMR data was obtained from the World Health Organization (WHO) web page (http://apps.who.int/ghodata/?vid=160). In this data set, countries were repeatedly measured on their IMR values over the years 1990, 1995, 2000, 2005 and 2010. Furthermore, Human Development Indices for corresponding five years were obtained from the Human Development Report, 2011 (UNDP, 2011).

The data set consisted of the values for 194 Member States of the United Nations observed at each of five time points for a total of 970 values. Andorra, Bhutan, Comoros, Djibouti, Grenada, Guinea-Bissau, Saint Kitts and Nevi, Saint Lucia, Saint Vincent and the Grenadines and Vanuatu were excluded in the analysis due to multiple missing values.

Results

Table 1 contains the BIC values for the models starting from one group to ten groups. In the analysis, BIC is maximized for four trajectory groups.

Table 1. Bayesian Information Criteria values

Number of groups	BIC
1	-3024.41
2	-2846.82
3	-2789.83
4	-2715.58
5	-2728.62
6	-2631.22
7	-2677.93
8	-2684.86
9	-2664.90
10	-2678.00

Figure 1 (next page) shows the graphical representation of trajectories for the four group model. Dashed lines represent the expected and solid lines represent the observed trajectories. A linear decrease in IMR over time is observed for each group (i.e. IMR is decreasing over time).

The numbering of the groups from one to four is consistent with the ordering from the lowest IMR rates from lowest average IMR to highest. The lowest IMR trajectory group contains an estimated 48.6 per cent of countries. There are 31.7 per cent of countries in the second group, 14.1 per cent in the third group and 5.5 per cent in the highest IMR group.

Figure 1. Trajectories of IMR under CNORM model, from 1990 to 2010

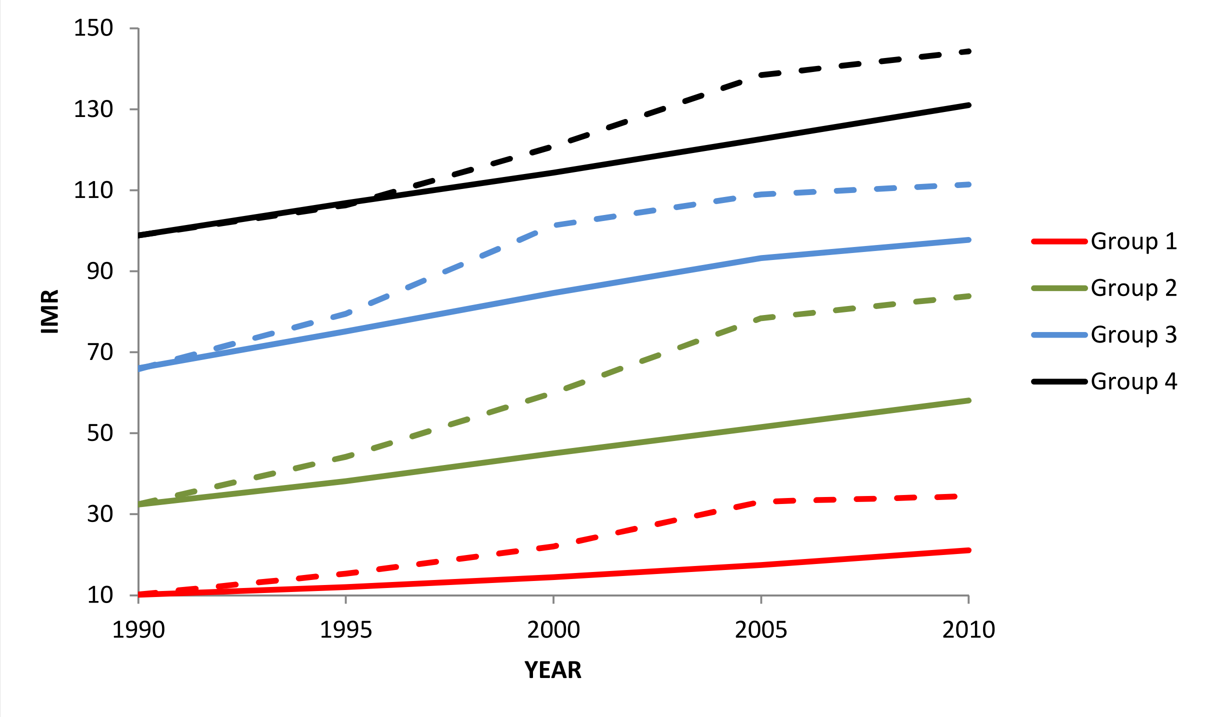

Table 2 contains the maximum likelihood estimates (MLEs) of the model parameters. The HDI variables added as a time varying covariate was highly significant. The MLEs of the parameters are reported in table 2. The coefficient of HDI was significantly negative in each group, with smallest T-value equal to -5.906. That is, within each IMR trajectory group, countries with higher HDI had a

Table 2. MLE of parameters under covariate model

Group	Parameter	Estimate	SE	T value	Prob>|T|
1	Intercept	72.9597	2.8779	25.351	0.000
	Linear	-1.0409	0.8865	-1.174	0.241
	HDI	-78.1491	3.6622	-21.339	0.000
2	Intercept	128.4422	2.6073	49.262	0.000
	Linear	-3.4891	1.1588	-3.011	0.003
	HDI	-136.3208	4.5267	-30.115	0.000
3	Intercept	143.2909	3.2531	44.046	0.000
	Linear	-3.3582	1.6955	-1.981	0.048
	HDI	-118.1073	6.6180	-17.846	0.000
4	Intercept	157.4661	4.4857	35.104	0.000
	Linear	-7.6681	2.8329	-2.707	0.007
	HDI	-90.2380	15.2786	-5.906	0.000

lower IMR than other in the group with lower HDI. For instance, for the lowest IMR trajectory group, the parameter estimate of HDI was equal to -78.15, meaning that a unit increase in HDI was associated with a decrease IMR of -78.15 units. The coefficient of HDI was largest for the second lowest IMR trajectory group. The coefficients in the third and highest IMR trajectory groups were also large.

The linear time coefficients were negative for each IMR trajectory groups, documenting a decreasing IMR after controlling for HDI. The linear time coefficient for the lowest IMR trajectory, while negative, was not significant (t=-1.174, p=0.2407). The linear time coefficient was significantly negative for all other IMR trajectory groups.

Table 3 contains the confidence intervals for the expected IMR of each IMR trajectory group at each time point. In the table, LL and UL are the lower limit and the upper limit values, and E is the estimated expected value.

Table 3. Confidence limits for estimated IMR

	Group							
	1		2		3		4	
Time Points	LL	UL	LL	UL	LL	UL	LL	UL
1990	31.01	37.93	80.01	87.58	106.68	116.1	94.09	103.64
1995	29.80	36.41	74.99	81.81	104.78	113.12	102.43	110.14
2000	19.51	24.51	57.06	62.84	97.21	105.40	115.26	126.40
2005	13.44	36.41	41.94	46.46	76.49	82.50	131.69	145.19
2010	8.64	37.93	30.27	34.62	62.32	69.25	137.37	151.12

Countries that are classified in the highest IMR trajectory group are described in table 4. Mean values and their standard deviations of the eleven countries are given along with mean model residuals in the table.

Trajectory groups were than compared with World Bank income level classifications. The 4 x 4 cross-classification table is given in table 5 (next page). World Bank Income group starts from “wealthy” to “poorest”; trajectory group starts from “adjusted low IMR” to “highest adjusted IMR”.

Table 4. Mean IMR and their standard deviation and model residuals

Country	Mean IMR	St.Dev.	Mean Residuals
Afghanistan	112,0	15.8272	38.7915
Angola	120,0	18.2071	41.1844
Central African Republic	111.6	4.0373	25.1319
Democratic Republic of the Congo	116,0	2.2361	24.7758
Equatorial Guinea	98.8	14.7207	37.8139
Guinea-Bissau	107.6	12.9923	13.7258
Liberia	114.6	32.3311	23.9624
Mali	114.2	12.6767	22.6720
Mozambique	118.8	21.0404	36.9939
Nigeria	109.8	15.9122	29.8126
Sierra Leone	141,0	20.3961	61.0123

The Weighted Cohen Kappa between the two classifications was 0.3513 (0.0418), indicating moderate agreement between the two classifications. This could possibly be due to misclassification. Eritrea was the country in the World Bank poorest group that was classified in the lowest IMR trajectory group. The next poorest World Bank group (3) had 20 countries in the lowest IMR trajectory group

Table 5. Cross table of trajectory groups by World Bank income groups

	World Bank income groups (from wealthy to poorest)					
		1	2	3	4	Total
Trajectory groups	1	42	27	20	1	90
	2	3	16	19	19	57
	3	2	2	8	14	26
	4	0	0	1	10	11
	Total	47	45	48	44	184

Conclusions

A mixture model is a probabilistic model for modeling subpopulations. Group-based modeling identified four groups of homogenous trajectories of IMR and estimated the trends in the outcome over time.

The results showed that IMR was decreasing over time in each group and was negatively associated with HDI. The lowest IMR trajectory group did not have a significant decrease in IMR after controlling for HDI. These countries were among the wealthiest. This result suggests that further decreases in IMR will be relatively small in this group and dependent on improvement of HDI. The next lowest trajectory group had the largest coefficient of HDI, suggesting that the improvement in IMR in this group would be projected to be large since IMR improvement had a continuing strong decline over time and a strong dependence on HDI. Similar expectations hold for the third IMR trajectory group and the highest IMR trajectory group.

With respect to average values between 1990 and 2010, the five countries with the highest infant mortality rate were Sierra Leone, Angola, Mozambique, Democratic Republic of Congo and Liberia. The IMR trajectory of Eritrea is notably unusual.

The top five countries with the lowest infant mortality rate were Iceland, Japan, Singapore, Sweden and Finland. Two common features of these countries stood out: 1) each had a very high literacy rate, and 2) each had an average life expectancy that was above average.

The lowest IMR group had relatively small decreases in IMR over time after adjustment for HDI. For example, the IMR of United States of America remained constant over the years 2000, 2005 and 2010.

Compared to the years 1990 and 2010, we note that small reductions of IMR were observed in Niue, Somalia, Nauru, Democratic People's Republic of Korea and Barbados. Timor-Leste, Maldives, Malaysia and Liberia experienced large reductions in IMR over the past twenty years, although the 2010 IMR rates were high.

The evidence presented in this paper suggests that significant percentage reductions in infant mortality might be possible for countries after controlling the HDI. Hence it can be said that the results are not completely with the World Bank income level classifications.

Maternal and child health policies of countries will contribute more to mortality decline, especially infant and child mortality. During the last century, IMR has declined markedly in most of the countries under study. This improvement may due to better nutrition, better prenatal and postnatal care and advanced medical technology. The fact that the IMR and the HDI are so highly correlated supports the idea that the causes of infant mortality are strongly related to many factors, such as economic development, general living conditions, social well being, and the quality of the environment. Furthermore, because IMR is sensitive to changes, HDI becomes an important indicator.

The presence of a dramatic downward trend in IMR in the 20th century, after controlling for HDI, reflects the fact that HDI is a very efficient determinant for IMR.

It should be noted that some countries have no data (e.g. Somalia) and are thus excluded from the analysis. It is clear that there is a high negative correlation IMR and HDI, which has become an important alternative to the traditional one dimensional measure of development. Why? Because HDI is based on the indices of health, education, and per capita gross domestic product.

In this paper, we built a group-based trajectory model that documented the relationship between HDI and IMR. With this approach, we were able to visualize and quantify the variability between subjects. Furthermore, analysis of the longitudinal data set over time allowed us to consider flexible models of the evolution of responses.

HDI, which is a useful measure for socioeconomic development, is an effective predictor of infant mortality rate. The findings of this study are consistent with those of previous studies on the subject.

References

Arrandale, V., M. Koehoorn, Y. MacNab, S.M. Kennedy (2006). "How to use SAS proc traj and SAS proc glimmix in respiratory epidemiology" [M.Sc. thesis]. University of British Columbia; 2006. [Available: http://circle.ubc.ca/handle/2429/834].

Erosheva, Elena, A. Matsueda, L. Ross, and Donatello Telesca(2014). "Breaking Bad: Two Decades of Life-Course Data Analysis in Criminology, Developmental Psychology, and Beyond". *Annual Review of Statistics and Its Application*, 1: 301-32.

Lee, K., S. Park, B. Khoshnood, H. Hsieh and R. Mittendorf (1997). **"**Human development index as a predictor of infant and maternal mortality rates". *The Journal of Pediatrics*, 131(3): 430-433.

Jones, B.L., D.S. Nagin and K. Roeder (2001). "A SAS procedure based on mixture models for estimating developmental trajectories." *Sociological Methods & Research* 29: 374-393.

Jones, B. L., and D.S. Nagin (2007). "Advances in Group-based Trajectory Modeling and a SAS Procedure for Estimating Them." *Sociological Research and Methods*, 35: 542-571.

Klugman, J., F. Rodriguez and H.J. Choi (2011). "United Nations Development Programme Human Development Reports Research Paper, April 2011"

Nagin, D.S. (1999). "Analyzing Developmental Trajectories: Semi-Parametric, Group-Based Approach." *Psychological Methods* 4:39-177.

Nagin, D.S. (2005). "Group-Based Modeling of Development." Cambridge, MA: Harvard University Press.

Nagin, D.S. and R.E. Tremblay (2001). "Analyzing Developmental Trajectories of Distinct but Related Behaviors: A Group-Based Method.'' *Psychological Methods* 6 (1): 18-34.

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

World Health Organization (2011). "World Health Statistics, 2011." [http://www.who.int/gho/publications/world_health_statistics/EN_WHS2011_Full.pdf].