DOES GOVERNMENT MATTER TO HEALTH OF ITS PEOPLE? A CROSS-COUNTRY LINKAGE BETWEEN HEALTH INDICATORS, HEALTH EXPENDITURE, AND THE MACROECONOMY

Angela Nalica
Paolo Victor Redondo
Stephen Jun Villejo
Erniel Barrios
University of the Philippines
School of Statistics

ABSTRACT

We develop a panel multilevel count data model. The model is estimated using the backfitting algorithm imbedded with bootstrap, forward search, and maximum likelihood estimation. With health indicators as the outcome, its linkage with some indicators like per capita income, economic growth, government expenditures on health, etc. is established using the postulated model. The study focuses on modelling incidence of tuberculosis (TB) using the Poisson Autoregressive or PAR(1) model with health expenditures, economic and health indicators as covariates among Southeast Asian countries. Results show that TB can be explained by expenditures for health, GDP growth of the country, access to clean water, population density and adult male mortality rate.

Keywords: panel data, multilevel count model, backfitting, Poisson autoregression

1. Introduction

Despite decades of economic growth and development in countries that belong to World Health Organization (WHO) Southeast Asian Region, most countries in this region still have a high burden of communicable diseases. Gupta and Guin (2010) consider that the Millennium Development Goals for health, economic indicator of progress per country, as well as the impact of global financial crisis on communicable diseases. The analysis indicates that focus of funding should include less-discussed but high-burden diseases often related to inadequacies in the health sector and the particular development paths that the different countries pursue. Although regional collaboration and integration is important, it is vital to take into the differences in the needs of each country in the health sector. Effectiveness of future policies can only be assured if these policies were developed by policy-makers familiar with the health priorities of each country.

One of the communicable diseases that persists to be a major global health problem is tuberculosis (TB). It remains to be one of the leading causes of death and disability for many decades. According to the Global Tuberculosis Report 2013, the number of deaths due to TB is unacceptably large given that they are the most preventable. Although victory has been achieved because of decreasing levels of those afflicted with TB through the years but the decline is not considered fast enough.

Wu and Dalal (2012) found that socioeconomic determinants and health system development have a significant effect on the control of TB in Asia and the Pacific. Tuberculosis incidence, prevalence and mortality rate were higher in countries with lower human development index, corruption perception index, gross domestic product (GDP) per capita and countries with more people under minimum food supplements. Among the health system variables, total health expenditure per capita, governmental health expenditure per capita, hospital beds, and access to improved water and sanitation were strongly associated with tuberculosis.

As the third Sustainable Development Goal states 'Ensure healthy lives and promote well-being for all at all ages." The United Nations made a statement to end the epidemics of AIDS, tuberculosis, malaria and neglected tropical diseases and combat hepatitis, water-borne diseases and other communicable diseases by 2030. The aim is to achieve universal health coverage for all and provide access to safe and affordable medicines and vaccines.

In line with this goal, the study aims to develop a panel multilevel count data model on incidence of TB. Data among 7 countries in the Southeast Asia from 1995 to 2014 were gathered. Health indicators and expenditures and economic indicators are considered as covariates. This research likewise intends to establish a single model for the whole Southeast Asian Region, which includes Philippines, Indonesia, Malaysia, Thailand, Cambodia, Lao and Vietnam (excluding Myanmar due to data constraint).

Definition of Variables

The following give a detailed description of the variables used in the study. All definitions are obtained from the World Health Organization (WHO) website.

Incidence of Tuberculosis (per 100,000 people)

"Incidence of tuberculosis is the estimated number of new and relapse tuberculosis cases arising in a given year, expressed as the rate per 100,000 population. All forms of TB are included, including cases in people living with HIV. Estimates for all years are recalculated as new information becomes available and techniques are refined, so they may differ from those published previously."

Government Health Expenditure per Capita (current US$)

"Total health expenditure is the sum of public and private health expenditures as a ratio of total population. It covers the provision of health services (preventive and curative), family planning activities, nutrition activities, and emergency aid designated for health but does not include provision of water and sanitation. Data are in current U.S. dollars."

Improved Water Source (% of population with access)

"Access to an improved water source refers to the percentage of the population using an improved drinking water source. The improved drinking water source includes piped water on premises (piped household water connection located inside the user's dwelling, plot or yard), and other improved drinking water sources (public taps or standpipes, tube wells or boreholes, protected dug wells, protected springs, and rainwater collection)."

Population Density (people per sq. km. of land area)

"Population density is midyear population divided by land area in square kilometers. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship--except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of their country of origin. Land area is a country's total area, excluding area under inland water bodies, national claims to continental shelf, and exclusive economic zones. In most cases the definition of inland water bodies includes major rivers and lakes."

Adult Male Mortality Rate (per 1,000 male adults)

" Adult mortality rate is the probability of dying between the ages of 15 and 60--that is, the probability of a 15-year-old dying before reaching age 60, if subject to age-specific mortality rates of the specified year between those ages."

GDP Growth Rate

"Annual percentage growth rate of GDP at market prices based on constant local currency. Aggregates are based on constant 2010 U.S. dollars. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources."

2. Proposed Model

Assuming a Poisson Autoregressive PAR(1) model for the incidence of tuberculosis Yt, with covariate health expenditure (X1t), access to improved water sources (X2t), population density (X3t), adult male mortality (X4t) and GDP growth rate (X5t), that is,

$$Y_t \sim \text{Poisson}(m_t) \text{ i.e., } P(Y_t | m_t) = \frac{m_t^{Y_t} e^{-m_t}}{Y_t} \quad \text{where}$$

$$m_t = \rho_1 Y_{t-1} + (1 - \rho_1) \exp(\underline{X}_t` \underline{\delta})$$

where Y_t - count response at time *t*,

m_t - parameter governing the Poisson data generating process at time *t*

ρ_1 - autoregressive parameter for lag 1

$\underline{X}_t$ - vector of covariates considered

$\underline{\delta}$ - vector of coefficients of each covariate

Proposed Estimation Algorithm

Component1:

An iterative estimation procedure in a backfitting framework that is proposed to estimate a PAR(1) model for each country.

For the kth country:

Step1: Ignore the autoregressive term and fit Y_t with covariates using Poisson Regression to obtain $\delta^{(0)}$

Step2: Define residual as $\mathrm{R1} = Y_{k,t} - exp(\underline{X}_{k,t}`\underline{\delta}_k{}^{(0)})$

Step3: Fit residual (*R1*) as AR(1) model to obtain $\rho^{(0)}$

Step4: Define new residual as $\mathrm{R2} = (Y_{k,t} - \rho_{k,1}^{(0)} Y_{k,t-1})/(1 - \rho_{k,1}^{(0)})$

For i = 1, 2, 3, …

Step5: Fit new residual (*R2*) with covariates using Poisson Regression to obtain $\delta^{(i)}$

Step6: Define new residual as $\mathrm{R1} = Y_{k,t} - \left(1 - \rho_{k,1}^{(i-1)}\right) exp(\underline{X}_{k,t}`\underline{\delta}_k{}^{(i)})$

Step7: Fit new residual (*R1*) as AR(1) model to obtain $\rho^{(i)}$

Step8: Define new residual as $\mathrm{R2} = (Y_{k,t} - \rho_{k,1}^{(i)} Y_{k,t-1})/(1 - \rho_{k,1}^{(i)})$

Iterate from Step5 defining residuals using the updated estimates for ρk and δk and fitting residuals R1 and R2 using Poisson Regression and as AR(1) model respectively until the convergence criterion, all percent change in parameter estimates are less than 0.1%, is met.

Therefore, we obtain estimates in the form of a matrix similar below.

Country\Parameter	$\rho_{k,1}$	$\delta_{k,1}$	$\delta_{k,2}$	$\delta_{k,3}$	$\delta_{k,4}$	$\delta_{k,5}$
Philippines (1)	$\rho_{1,1}$	$\delta_{1,1}$	$\delta_{1,2}$	$\delta_{1,3}$	$\delta_{1,4}$	$\delta_{1,5}$
Indonesia (2)	$\rho_{2,1}$	$\delta_{2,1}$	$\delta_{2,2}$	$\delta_{2,3}$	$\delta_{2,4}$	$\delta_{2,5}$
Malaysia (3)	$\rho_{3,1}$	$\delta_{3,1}$	$\delta_{3,2}$	$\delta_{3,3}$	$\delta_{3,4}$	$\delta_{3,5}$
Thailand (4)	$\rho_{4,1}$	$\delta_{4,1}$	$\delta_{4,2}$	$\delta_{4,3}$	$\delta_{4,4}$	$\delta_{4,5}$
Cambodia (5)	$\rho_{5,1}$	$\delta_{5,1}$	$\delta_{5,2}$	$\delta_{5,3}$	$\delta_{5,4}$	$\delta_{5,5}$
Lao (6)	$\rho_{6,1}$	$\delta_{6,1}$	$\delta_{6,2}$	$\delta_{6,3}$	$\delta_{6,4}$	$\delta_{6,5}$
Vietnam (7)	$\rho_{7,1}$	$\delta_{7,1}$	$\delta_{7,2}$	$\delta_{7,3}$	$\delta_{7,4}$	$\delta_{7,5}$

Component2:

From Component1 of the proposed estimation procedure:

Step1: Obtain the mean of all autoregressive parameters, denoted by $\hat{\rho}$.

Step2: Define new residual as R3 = $(Y_{k,t} - \hat{\rho}Y_{k,t-1})/(1-\hat{\rho})$ for all countries

Step3: Fit the log of the new residuals [log(R3)] with covariates using OLS Regression with all data points from different countries as one data set to obtain $\underline{\hat{\delta}} = (\hat{\delta}_1, \hat{\delta}_2, \hat{\delta}_3, \hat{\delta}_4, \hat{\delta}_5)'$.

Thus, estimates obtained from Component2 of the algorithm will be used to formulate the single model for the Southeast Asian Region.

Bootstrap Significance of Parameters

To validate the adequacy of the included variables, bootstrap replicates of the actual data using AR Sieves framework are obtained where the proposed algorithm is applied.
AR Sieve Bootstrap Procedure:

Step1: From the initial estimates for the "regional" model, obtain the residuals as

R4 = $Y_{k,t} - \hat{Y}_{k,t}$ where $\hat{Y}_{k,t} = \hat{\rho}Y_{k,t-1} + (1-\hat{\rho})\exp(\underline{X_{k,t}}`\underline{\hat{\delta}})$

Step2: Resample residuals R4 using SRSWR with the same size as the overall time points in all countries.

Step3: Use the sampled residuals R4 to obtain a new time series $\{Y^*_{k,t}\}$ that has the same structure as the original time series $\{Y^*_{k,t}\}$ where $Y^*_{k,t}$ is defined as

$Y^*_{k,t} = \hat{\rho}Y^*_{k,t-1} + (1-\hat{\rho})\exp(X_{k,t}`\hat{\delta})$ + R4

Significance:

Note that 100 replicates of the original time series are produced and each replicate is estimated using the proposed estimation procedure, thus, going through Component1 and

Component2. For the 100 set of estimates obtained, the 2.5th and 97.5th percentiles are calculated to test for the significance of each parameter estimate.

3. Results and Discussion

The proposed model and algorithm give the following results:

Variable	Parameter	Estimate	2.5th Percentile	97.5 Percentile
Tuberculosis	$\hat{\rho}$	0.949205686	0.944591	0.948508
Health Expenditure	$\hat{\delta}_1$	-0.001602306	-0.001771	-0.00112
Clean Water	$\hat{\delta}_2$	-0.011038476	-0.012287	-0.009375
Population Density	$\hat{\delta}_3$	0.001036708	0.000726	0.001087
Adult Male Mortality	$\hat{\delta}_4$	0.004947556	0.004767	0.005299
GDP Growth	$\hat{\delta}_5$	-0.01118079	-0.009396	-0.006574

From the table above, the estimated "regional" model is given by

$$Y_t \sim \text{Poisson}(m_t) \quad \text{where}$$

$$m_t = 0.9492*Y_{t-1} + 0.0508*\exp(-0.0016*X_{1,t} - 0.0110*X_{2,t} + 0.0010*X_{3,t} + 0.0049*X_{4,t} - 0.0112*X_{5,t})$$

The estimated model indicates a strong positive effect of the immediate past on the prevalence of tuberculosis to its present value. Since treatment for TB should run for six months of taking medicines, it is expected that the immediate past would still show an influence on the current TB prevalence as data comes annually. Also, government health expenditure, access to clean water and GDP growth have a negative relationship with the prevalence, that is, higher value for these variables suggest lower tuberculosis cases. On the other hand, population density and adult male mortality suggest otherwise. An increase in population density would have a corresponding increase in TB cases since this is a communicable and highly infectious disease. For adult male mortality which can be viewed as a proxy variable for health spending, an increase in this covariate would mean less health spending which can suggest an increase in TB prevalence. Furthermore, since the calculated 95% bootstrap confidence interval does not contain zero (0) for any of the parameters, all variables included in the model have a significant effect on the prevalence of tuberculosis.

4. Conclusions and Recommendation

The proposed model gives us a unified model for all Asian countries considered in the study. Government expenditure on health, growth rate as a measure of the economic progress of a country as well as other health indicators are found to be significant contributors to the incidence of TB. Thus, government matters to the health of its people.

Although regional collaboration and integration is critical, it is recommended to work on a model that would incorporate country effect to address varying regional needs. Each country would have differences in its burden of communicable diseases and would warrant varying and diverse measures to combat prevalence of a particular disease.