



Bayesian Hierarchical Mix-Effects Approach to Impacts of Air Pollution and Economic Growth on Private Health Care Expenditure

Bui Hoang Ngoc¹ and Nguyen Ngoc Thach²(✉)

¹ The FEMRG Research Group, Ho Chi Minh City Open University, 97 Vo Van Tan Street, District 3, Ho Chi Minh City 70000, Vietnam
ngoc.bh@ou.edu.vn

² Banking University HCMC, Ho Chi Minh City, Vietnam
thachnn@buh.edu.vn

Abstract. Assessment of the interactive effects of specific macro factors such as air pollution, economic growth, and urbanization on private health care expenditure is vital for social sustainability policies. Nevertheless, to the best of our knowledge, previous studies on this topic focused on economies outside ASEAN and, particularly, were conducted in the out-of-date frequentist framework. Furthermore, empirical outcomes are often different, even contradictory. This research aims to revisit the effects of air pollution, economic growth, and urbanization on private health care expenditure in ASEAN using the Bayesian two-level mixed-effects regression via the Metropolis-Hastings algorithm to capture varying effects of the determinants on the response variable, decreasing model uncertainty. The Bayesian outcomes show that income per capita and air pollution positively and strongly impact private health care expenditure. But, interestingly, the impact of urbanization is ambiguous.

1 Introduction

Over the past decade, humanity has witnessed an increase in mortality and morbidity caused by environmental pollution. Air pollution adversely affects the development of the natural world, threatening the safety of human life (Currie et al. 2008; Hansen & Selte 2000; Yazdi & Khanalizadeh 2017). The pressure of improving income per capita boosts economic activities, which intensifies the destruction of environmental quality (Hassan et al. 2019; Moutinho et al. 2017; Xu 2018). Environmental pollution and climate change trigger health problems, resulting in huge health costs. The impact of air pollution on individual health is obvious, but who pays health costs is an essential policy issue. In most emerging countries, the public health care service systems are overloaded, whereas government budgets spent on health insurance are not sufficient to meet the requirements of residents.

The nexus of economic growth, air pollution and health expenditure has been discussed in the United States and the United Kingdoms. In the early 1950s, when the American Institute of Public Health published a report of the impact of air pollution on health in 1957. Nowadays, a great number of scientists have found a positive impact of air pollution and economic growth on health spending. The pioneering work of Newhouse (1977) investigating the influence of economic growth on health care expenditure in 13 Organisation for Economic Co-operation and Development (OECD) countries revealed a positive effect of income on personal health budget. Consistent with this view, Murthy and Okunade (2016) examined the effects of the macroeconomic factors on American health spending from 1960 to 2012 by applying the ARDL approach. Real income, growth rate of the population over 65 years, and high medical technology are shown to be positively correlated to health spending. Analyzing a sample of 15 OECD countries for 1995–2011, Doğan et al. (2014) revealed that the largest and smallest effects on health care expenditure are exerted by public health expenses and labor force respectively, while the influence of income is positive. However, Hansen and King (1996) found no relationship between economic growth and health care expenses for 20 OECD countries. This conclusion was reconfirmed by McCoskey and Selden (1998). Via the fixed effect and random effect models on the data of 29 OECD countries for 1995–2014, Fernandez et al. (2019) also discovered that per capita income has a insignificantly positive effect on health spending. To explain the conflicting conclusions on the mentioned connection, Newhouse (1977), Blazquez-Fernandez et al. (2019), Liang and Tussing (2019) argued that the relationship depends on the “Norm” hypothesis, according to which physicians, medical facilities, and hospitals provide health care services by formal norms. That means health care services are offered more and better if people are involved in many types of health insurance, or participate at a high level of cost.

Like economic growth, CO₂ emissions tend to positively affect health expenditure. Chaabouni et al. (2016) analyzing a sample of 51 countries for 1995–2013 using the GMM approach showed that there is bi-directional causality between CO₂ emissions and economic growth and between health spending and economic growth for a global sample, but uni-directional causality running from CO₂ emissions to health spending, except for low-income countries. Employing the Wavelet method for the U.S, Alola and Kirikkaleli (2018) indicated that CO₂ emissions are related to healthcare. Okunade (2005) applying the ordinary least square (OLS) regression on a sample of 26 African countries found inequality and income per capita to raise health care expenditure. Mehrara et al. (2012) utilized the vector error correction model (VECM) to clarify the links between economic growth and health expenditure in 13 Middle East and North Africa (MENA) countries during 1995–2005. A bounds test confirms cointegration among the examined variables. Specifically, health care expenses have a positive impact on income growth, but with the rate of contribution declining. Using the PM10 emissions (PM10 is a particulate matter 10 μm or less in diameter, which can be drawn deep into the lungs or blood) as a proxy for environmental pollution,

Yazdi and Khanalizadeh (2017) reconfirmed the positive effect of air pollution and economic growth on health care expenditure in MENA countries. Likewise, using VECM for the Sub-Saharan African countries from 1990 to 2015, Zaidi and Saidi (2018) affirmed that economic growth positively affects health care expenses, while the effect of air pollution is negative. By using the autoregressive distributed lag (ARDL) approach to Turkey during 1975–2007, Yavuz et al. (2013) discovered that income has no effect on health care expenditure in the long-run. Furthermore, provincial output and environmental pollution have a positive impact on public health expenditure for China during 1997–2014 (Yu et al., 2016).

Recently, investigating 20 cities in China, Yang et al. (2019) found that PM_{2.5} emissions are related to mortality, morbidity and tend to cause economic loss. In analysis of the effects of macroeconomic factors on health care expenses in 18 Arab world countries during 1995–2015, Barkat et al. (2019) specified the same model for three groups: high-, upper-middle- and lower-middle-income countries. The empirical results from the pooled mean group (PMG) and the common correlated effects (CCE) suggest that income is not the only driver of health expenditure in the Arabian countries in the long-run. Similarly, Raeissi et al. (2018) reported that a 1% increase in carbon dioxide leads to an increase of 3.32% and 1.16% in public and private health expenditures in Iran. However, Zaidi and Saidi (2018) revealed that a 1% increase in CO₂ and NO₂ emissions leads to a 0.066% and 0.577% decrease in health care expenses in the Sub-Saharan African countries. Analyzing the Greece case for 2008–2015 through a probit 2SLS regression, Kyriopoulos et al. (2019) revealed that household health care expenses respond to permanent income changes more strongly than the ones arising from current income shocks.

Similar to CO₂ emissions, urbanization could positively contribute to health-care. People tend to move to cities because of many opportunities for them to find a job with a good income. Still, the environmental pollution situation in big cities is getting worse and worse, and life costs there are also much higher than in villages. Rapidly developing urban areas have intensified CO₂ emissions (Hashmi et al. 2021). Although the relationship between air pollution, economic growth and health spending is not fully exhausted by our selective review, our paper affirms that conclusions drawn from the considered works rest mainly on data-driven frequentist inference as an obsolete estimator (Kalli & Griffin 2018; Kim 2002; Norets 2015). If the coefficients of independent variables are not yet significant, the suggestion of a conclusion or implication is difficult and even impossible (Hansen and King 1996, McCoskey and Selden 1998). The issue could be solved within the more balanced and more reliable Bayesian setting, but no research employed these methods to give substantial insights into the relationships between economic growth, air pollution, urbanization and private health care expenditure in the association of southeast Asian nations (ASEAN). This is the research gap which the current study wants to address. Hence, the contributions of the investigation are summarized as follows: first, to the best of our knowledge, ours is one of the first works accessing the impacts of air pollu-

tion, economic growth, and urbanization on private health care expenditure in ASEAN, and hence provides new empirical evidence; second, the majority of earlier studies on health care expenses utilized outdated frequency-based techniques, where coefficient parameters are fixed point estimates. More importantly, examining the links between all interested variables is impossible as non-significant coefficients are dropped out of analysis. We are the first to apply the Bayesian approach through the integrated Markov chain Monte-Carlo (MCMC) sampler to give probabilistic interpretations of model uncertainty and varying effects of air pollution, economic growth, and urbanization on private health care expenditure. With the accomplishments achieved, the research methodologically and empirically contributes to the health care expenditure field.

2 Methodology

In the context of an increasing crisis in standard frequentist statistics, due to the surprisingly rapid advancements of powerful computational tools, the Bayesian approach has been becoming a more and more commonly used methodology in behavioral and social sciences over the past 30 years (Lemoine 2019; Ngoc & Awan 2021). The research employs a Bayesian two-level mixed-effects method to capture the variability of initial health care expenditure across 10 ASEAN countries, due to which the precision of the estimates increases. Mixed-effects models or multilevel models are featured as incorporating both fixed effects and random effects. The former comparable to frequentist regression coefficients are estimated directly. On the contrary, the latter is summarized according to their estimated variances and covariances. Random effects may be either random intercepts or random coefficients, while the grouping structure of the data may consist of multiple levels of nested groups. We focus on random intercepts assuming that initial health care expenditure varies across the studied countries, while the effects of income, air pollution, and urbanization on private health care expenditure are the same.

We specify a random-effects model as follows:

$$\ln HE_{it} = \beta_0 + \beta_1 \ln GDP_{it} + \beta_2 \ln CO_{2,it} + \beta_3 UB_{it} + u_i + \varepsilon_{it}$$

where HE is private health care expenditure per capita (unit: US dollar), GDP is income per capita (unit: US dollar in constant 2010 prices), CO₂ is CO₂ emissions per capita (unit: metric ton), UB is urbanization rate (unit: percentage), u_i is random intercepts, ε is random error, i is country, and t is year. Annual data is collected from the World Bank database from 2000 to 2016.

Because our data sample size is sufficiently large, different prior specifications do not influence posterior results and in this situation noninformative priors are enough for modeling. For comparison purposes, informative priors for the model parameters are specified too. Accordingly, five posterior simulations are made. A sensitivity analysis to prior choice will be performed through a Bayes factor test and a model test. We assume to have models M_k parameterized by vectors

$\theta_k, k = 1, 2, \dots, r$. Through the use of Bayes's theorem, we calculate the posterior model probabilities:

$$\rho(M_k | x) = \frac{\rho(x | M_k) * \rho(M_k)}{\rho(x)}$$

Since it is challenging to calculate $\rho(x)$, a popular practice is to compare two models, for example, M_k and M_l via posterior odds ratio:

$$POR_{kl} = \frac{\rho(M_k | x)}{\rho(M_l | x)} = \frac{\rho(x | M_k) * p(M_k)}{\rho(x | M_l) * p(M_l)}$$

In the case of all equally plausible models, the posterior odds ratio is transformed into the Bayes factor:

$$BF_{kl} = \frac{p(x | M_k)}{p(x | M_l)}$$

Information criteria such as Akaike information criterion (AIC), Bayesian information criterion (BIC), and the deviance information criterion (DIC) are utilized to determine the most suitable model among candidate models that fits the data best. The drawback of all these criteria, however, is that they either ignore prior distributions or suppose only noninformative prior distributions. They are therefore not appropriate for Bayesian sensitivity analysis, when comparison models have the same parameters but different priors. In accessing the Bayesian framework, Bayes factors are preferred to model-selection criteria as, contrast to BIC, AIC, and DIC, they contain all information about prior distributions. Focus on prior information is crucial for Bayesian sensitivity analysis, when research compares models with the same parameters but various priors. The Bayes factor of two models is just the ratio of their marginal likelihoods calculated on the same data. Bayes factors usage, nevertheless, is often criticized in some venues for a difficulty in calculating marginal likelihoods. Besides, being relative quantities, Bayes factors cannot be adopted to evaluate goodness of fit of a model of interest unless the base model fits the data well. This is another limitation of Bayes factors.

While a Bayes model test compares posterior probabilities of specified Bayesian models to determine which model is more likely among examined models on the same dataset. The present study compares the candidate models with the same parameters but different priors.

3 Bayesian Simulation Outcomes

Model Comparison

This subsection compares five posterior regression models, where the respective Gaussian prior distributions are $N(0,1)$, $N(0,10)$, $N(0,100)$, $N(0,1000)$, and $N(0,10000)$.

Table 1. Bayes factor test

Model	Gaussian distribution	DIC	log(ML)	log(BF)
Model 1	N(0,1)	335.6943	-184.7604	*
Model 2	N(0,10)	336.1227	-189.1019	-4.3415
Model 3	N(0,100)	336.1339	-193.9300	-9.1696
Model 4	N(0,1.000)	336.1047	-198.5439	-13.7835
Model 5	N(0,10.000)	36.8694	-80.2871	104.4733

Note: * means reference logBF

Table 2. Bayesian model test

Model	Gaussian distribution	log(ML)	P(M)	P(My)
Model 1	N(0,1)	-184.7604	0.2000	0.0000
Model 2	N(0,10)	-189.1019	0.2000	0.0000
Model 3	N(0,100)	-193.9300	0.2000	0.0000
Model 4	N(0,1.000)	-198.5439	0.2000	0.0000
Model 5	N(0,10.000)	-80.2871	0.2000	1.0000

The model comparison results are presented in Tables 1 and 2. In general, the smaller the DIC value, the larger the log(ML) and log(BF) estimate, the better a model fits the data (Table 1). P(My) denotes the posterior model probability. The higher P(My), the better a posterior model (Table 2). Consequently, model 5 is the best.

MCMC Convergence Test

In the application of a MCMC algorithm, convergence checks are needed before proceeding to inference. Once chain convergence is established, the model parameters have converged to equilibrium values. To avoid pseudo convergence, we simulate three MCMC chains and verify whether the results satisfy the convergence rule. This is because pseudo convergence takes place when the chains have seemingly converged, but indeed, they explored only a portion of the domain of a posterior distribution. As demonstrated in Table 3, the maximum Gelman-Rubin statistic Rc of 1.01 is close to 1.1, indicating MCMC convergence.

Table 3. Gelman-Rubin convergence diagnostic

Max Gelman-Rubin Rc = 1.0075	
Dependent variable: <i>lnHE</i>	Rc value
<i>lnCO₂</i>	1.0028
<i>lnGDP</i>	1.0006
<i>UB</i>	1.0020
<i>Intercept</i>	1.0075
<i>U₀ : sigma2</i>	1.0071
<i>sigma2</i>	1.0004
Convergence rule: Rc < 1.1	

The model summary reports rate of acceptance and algorithm efficiency as initial indicators of MCMC convergence. The acceptance rate is the number of proposals accepted in the total proposals, whereas algorithm efficiency is the mixing properties of MCMC sampling. Concerning the chosen model 5, the acceptance rate of 0.83 is larger than the minimum level of 0.1, whereas average efficiency is equivalent to 0.19, which is more than the acceptable level of 0.01.

Bayesian Simulation Outcomes

Table 4 exhibits the simulation summaries of model 5. The variables $lnCO_2$, $lnGDP$, and UB are of our interest. With a probability of mean between 0.7 and 1, $lnGDP$ and $lnCO_2$ exert strongly positive effects on $lnHE$, while the 54% probability denotes that the impact of UB is ambiguous. The 95% credible intervals also point to similar results. The lower Monte-Carlo standard error (MCSE) values, the more accurate posterior mean estimates. For MCMC algorithms, MCSE estimates close to one decimal are acceptable. Moreover, standard deviations for all the parameters are small, indicating the preciseness of parameter estimates. Importantly, because the variations of private health care expenditure between the researched countries are captured in a Bayesian mixed-effects model, its variance estimates decrease in comparison with those from maximum likelihood estimation.

Compared to frequentist statistics, credible Bayesian intervals have direct and intuitive probabilistic interpretation. Thus, we can state that the coefficient for $lnCO_2$ belongs to the interval $[-0.139, 0.275]$ with a 95% probability. Similar interpretations can be offered for the remaining parameters of the model.

Table 4. Bayesian simulation outcomes

Variables	Mean	Std.Dev.	MCSE	Probability of mean > 0	Equal-tailed [95% Cred.Interval]
$lnCO_2$	0.0674	0.1054	0.0062	0.72	[-0.1392, 0.2747]
$lnGDP$	1.9764	0.1393	0.0048	1	[1.6994, 2.2441]
UB	0.0015	0.0102	0.0005	0.56	[-0.0182, 0.0218]
<i>Intercept</i>	-12.701	1.2346	0.1716	1*	[-15.139, -10.278]
$U_0 : sigma2$	5.8076	3.9268	0.2612	1	[1.8123, 15.615]
$sigma2$	0.0673	0.0088	0.0001	1	[0.0533, 0.0855]

Note: * Probability of mean < 0

Discussion

Bayesian outcomes report the following findings. Per capita income, air pollution, and urbanization positively affect private health expenditure. However, their impact magnitude is different in terms of probability, with a strongly positive effect from income and a moderate effect from CO_2 emission, but an ambiguous relationship between urbanization and private health care expenditure. Importantly, our estimation results are economically plausible. Compared with the previous studies, our outcome is in line with Yazdi and Khanalizadeh (2017), Zaidi and Saidi (2018), Barkat et al. (2019), reconfirming a positive impact of air pollution on health care expenditure. Regarding the effect of income on health care expenses, our result is consistent with Gerdtham et al. (1992) Hitiris

and Posnett (1992), Bhat and Jain (2006), Wang and Rettenmaier (2007). The reasons for the outcomes could be as follows: first, the rise of income allows to obtain access to better health care services at a high-cost level (Rao et al. 2009); second, the consciousness of people on their health positively changes resulting from income improvement.

4 Concluding Remarks

The study estimates the effects of income, air pollution, and urbanization on private health care expenditure employing the Bayesian hierarchical mixed-effects regression through the hybrid Metropolis-Hastings sampler on a panel data of 10 ASEAN countries over the period 2000 to 2016. The Bayesian mix-effects regression allows for variance reduction and so increases the accuracy of the estimates. According to the simulation outcomes, we claim in view of the probability that economic growth strongly and positively affects private health care expenditure, CO₂ emissions have a moderate positive impact, while the effect of urbanization is ambiguous.

Based on the obtained empirical results, two main policy implications are suggested: *First*, ASEAN countries should plan detailed policies to propagate on the harmful effects of environmental pollution, encouraging residents to use clean energy, enterprises to invest in high and green technologies; *Second*, medical waste is a factor in destroying environmental quality. Medical garbage should be well controlled by governments and hospitals.

The main limitation of this research is that we did not incorporate random coefficients due to the high dimensionality of multilevel models, which may lead to non-convergence for some model parameters.

Acknowledgments. We thank the anonymous reviewers and Editor-in-chief for helpful comments and suggestions. All errors are our.

References

- Alola, A.A., Kirikkaleli, D.: The nexus of environmental quality with renewable consumption, immigration, and healthcare in the US: wavelet and gradual-shift causality approaches. *Environ. Sci. Pollut. Res. Int.* **26**(34), 35208–35217 (2019)
- Barkat, K., Sbia, R., Maouchi, Y.: Empirical evidence on the long and short run determinants of health expenditure in the Arab world. *Q. Rev. Econ. Financ.* **73**, 78–87 (2019)
- Bhat, R., Jain, N.: Analysis of public and private healthcare expenditures. *Econ. Pol. Wkly* **41**(1), 57–68 (2006)
- Blazquez-Fernandez, C., Cantarero-Prieto, D., Pascual-Saez, M.: On the nexus of air pollution and health expenditures: new empirical evidence. *Gac. Sanit.* **33**(4), 389–394 (2019)
- Chaabouni, S., Zghidi, N., Ben Mbarek, M.: On the causal dynamics between CO₂ emissions, health expenditures and economic growth. *Sustain. Urban Areas* **22**, 184–191 (2016)

- Currie, J., Neidell, M., Schmieder, J.: Air pollution and infant health: Lessons from New Jersey. National Bureau of Economic Research (Nber Working Paper, No. 14196) (2008)
- Doğan, İ, Tülüce, N.S., Doğan, A.: Dynamics of health expenditures in OECD countries: panel ARDL approach. *Theor. Econ. Lett.* **4**(8), 649–655 (2014)
- Gerdtham, U.-G., Sjøgaard, J., Andersson, F., Jönsson, B.: An econometric analysis of health care expenditure: a cross-section study of the OECD countries. *J. Health Econ.* **11**(1), 63–84 (1992)
- Hansen, King, A.: The determinants of health care expenditure: a cointegration approach. *J. Health Econ.* **15**(1), 127–137 (1996)
- Hansen, A.C., Selte, H.K.: Air pollution and sick-leaves: a case study using air pollution data from Oslo. *Environ. Resource Econ.* **16**(1), 31–50 (2000)
- Hashmi, S.H., Fan, H., Habib, Y., Riaz, A.: Non-linear relationship between urbanization paths and CO₂ emissions: a case of South, South-East and East Asian economies. *Urban Climate* **37**, 100814 (2021)
- Hassan, S.T., Xia, E., Khan, N.H., Shah, S.M.A.: Economic growth, natural resources, and ecological footprints: evidence from Pakistan. *Environ. Sci. Pollut. Res. Int.* **26**(3), 2929–2938 (2019)
- Hitiris, T., Posnett, J.: The determinants and effects of health expenditure in developed countries. *J. Health Econ.* **11**(2), 173–181 (1992)
- Kalli, M., Griffin, J.E.: Bayesian nonparametric vector autoregressive models. *J. Econ.* **203**(2), 267–282 (2018)
- Kim, J.-Y.: Limited information likelihood and Bayesian analysis. *J. Econ.* **107**(1–2), 175–193 (2002)
- Lemoine, N.P.: Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos* **128**(7), 912–928 (2019)
- Liang, L.L., Tussing, A.D.: The cyclicalities of government health expenditure and its effects on population health. *Health Policy* **123**(1), 96–103 (2019)
- McCoskey, S.K., Selden, T.M.: Health care expenditures and GDP: panel data unit root test results. *J. Health Econ.* **17**(3), 369–376 (1998)
- Mehrrara, M., Fazaeli, A.A., Fazaeli, A.A., Fazaeli, A.R.: The relationship between health expenditures and economic growth in Middle East & North Africa (MENA) countries. *Int. J. Bus. Econ. Res.* **3**(1), 425–428 (2012)
- Moutinho, V., Varum, C., Madaleno, M.: How economic growth affects emissions? An investigation of the environmental Kuznets curve in Portuguese and Spanish economic activity sectors. *Energy Policy* **106**, 326–344 (2017)
- Murthy, V.N.R., Okunade, A.A.: Determinants of U.S. health expenditure: evidence from autoregressive distributed lag (ARDL) approach to cointegration. *Econ. Model.* **59**, 67–73 (2016)
- Newhouse, J.P.: Medical-care expenditure: a cross-national survey. *J. Hum. Resour.* **12**(1), 115–125 (1977)
- Ngoc, B.H., Awan, A.: Does financial development reinforce ecological footprint in Singapore? Evidence from ARDL and Bayesian analysis. *Environ. Sci. Pollut. Res. Int.* **29**(6), 24219–24233 (2021)
- Norets, A.: Bayesian regression with nonparametric heteroskedasticity. *J. Econ.* **185**(2), 409–419 (2015)
- Raeissi, P., Harati-Khalilabad, T., Rezapour, A., Hashemi, S.Y., Mousavi, A., Khodabakhshzadeh, S.: Effects of air pollution on public and private health expenditures in Iran: a time series study (1972–2014). *J. Prev. Med. Public Health* **51**(3), 140–147 (2018)

- Rao, R.R., Jani, R., Sanjivee, P.: Health, quality of life and GDP: an ASEAN experience. *Asian Soc. Sci.* **4**(4), 70–76 (2009)
- Wang, Rettenmaier, A.J.: A note on cointegration of health expenditures and income. *Health Econ.* **16**(6), 559–578 (2007)
- Xu, T.: Investigating environmental Kuznets curve in China-aggregation bias and policy implications. *Energy Policy* **114**, 315–322 (2018)
- Yang, S., Fang, D., Chen, B.: Human health impact and economic effect for PM2.5 exposure in typical cities. *Appl. Energy* **249**, 316–325 (2019)
- Yazdi, K.S., Khanalizadeh, B.: Air pollution, economic growth and health care expenditure. *Econ. Res.-Ekonomiska Istraživanja* **30**(1), 1181–1190 (2017)
- Zaidi, S., Saidi, K.: Environmental pollution, health expenditure and economic growth in the Sub-Saharan Africa countries: panel ARDL approach. *Sustain. Urban Areas* **41**, 833–840 (2018)