

Toward a Hierarchical Bayesian Framework for Modelling the Effect of Regional Diversity on Household Expenditure

¹Pudji Ismartini, ²Nur Iriawan, ²Setiawan and ²Brodjol Sutijo Supri Ulama

^{1,2}Department of Statistics, Faculty of Mathematics and Natural Sciences

Institut Teknologi Sepuluh Nopember, Jl. Arif Rahman Hakim, Surabaya 60111, Indonesia

¹Statistics Indonesia (BPS), Jakarta, Indonesia

Abstract: Problem statement: Household expenditure analysis was highly demanding for government in order to formulate its policy. Since household data was viewed as hierarchical structure with household nested in its regional residence which varies inter region, the contextual welfare analysis was needed. This study proposed to develop a hierarchical model for estimating household expenditure in an attempt to measure the effect of regional diversity by taking into account district characteristics and household attributes using a Bayesian approach. **Approach:** Due to the variation of household expenditure data which was captured by the three parameters of Log-Normal (LN3) distribution, the model was developed based on LN3 distribution. Data used in this study was household expenditure data in Central Java, Indonesia. Since, data were unbalanced and hierarchical models using a classical approach work well for balanced data, thus the estimation process was done by using Bayesian method with MCMC and Gibbs sampling. **Results:** The hierarchical Bayesian model based on LN3 distribution could be implemented to explain the variation of household expenditure using district characteristics and household attributes. **Conclusion:** The model shows that districts characteristics which include demographic and economic conditions of districts and the availability of public facilities which are strongly associated with a dimension of human development index, i.e., economic, education and health, do affect to household expenditure through its household attributes.

Key words: Household attributes, Log-Normal (LN3), hierarchical structure, estimating household expenditure, reducing poverty, development-oriented, paradigms associated

INTRODUCTION

Regional income distribution can determine the ability of the region in creating change and improvement of its people, such as reducing poverty. It is noted that inequality of regional income distribution will not create wealth for society in general, but only creates wealth for certain groups. According to BPS (2010b), inequality of income distribution can be viewed from three sides. First, the relative inequality i.e., size distribution of income disparities. Second, rural-urban income disparities which are usually caused by more development-oriented to urban areas. This Urban bias development often occurs in developing countries such as Indonesia. Third is the regional income disparity, which is generally viewed in Indonesia because of the economic development disparities between regions and inequality in the distribution of natural resources between region.

Basically the factors that affect the welfare problems can be broadly categorized into two main things. Those are behavior paradigms and policy paradigms (Akita and Pirmansyah, 2011). Behavioral paradigms related to the effort of responsibilities of each individual or household in achieving their welfare levels. In each household, there are specific factors that potentially contribute to the paradigm of such behavior. While the policy paradigms associated with economic conditions, politics and government policy. In addition, non-household factors may also affect the difference in the level of welfare. An example is community-level factors such as geography and availability of public facilities (economic, education and health facilities).

Income per capita is an economic indicator that is often used for measuring the prosperity and well-being. Analysis of household income is essential in order to formulate government policy.

Corresponding Author: Pudji Ismartini, Department of Statistics, Faculty of Mathematics and Natural Sciences Institut Teknologi Sepuluh Nopember, Jl. Arif Rahman Hakim, Surabaya 60111, Indonesia, Statistics Indonesia (BPS), Jakarta, Indonesia

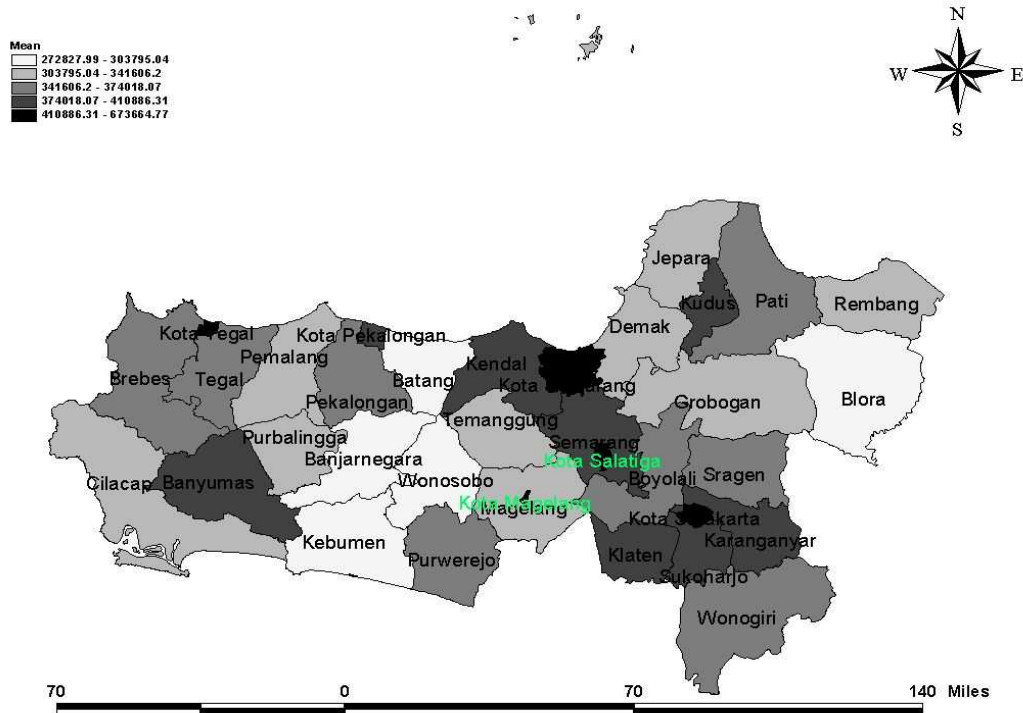


Fig. 1: Map of central Java by mean of household expenditure per capita

However, household income is generally very difficult to be measured accurately, especially in developing countries. Basically, household income and household expenditure are not the same thing. But such relationships between those two are very strong. Akita and Pirmansyah (2011) states that consumption expenditure is more reliable than income as an indicator of a household permanent income because it does not vary as much as income in the short term. For those reasons, household expenditure patterns approach is then widely used to analyze the pattern of household income.

Indonesia has changed its governance systems for centralized into a decentralized system since 1999. Consequently, the achievement of local government will be largely determined by the active and innovative role of local government in determining its local policy in order to achieve prosperity and welfare of its residents. Since the Indonesian area is vast and the regional conditions vary with each other, the contextual welfare analysis needed by taking into account the regional diversity in order to formulate government policy. Shahateet (2006) shows that there is regional effect of income inequality.

Central Java is one of the provinces on Java Island in Indonesia. It is known as the heart of Javanese culture because the culture of Central Java is diverse

and includes a variety of cultures from another province in Java. The total area of Central Java is 32,800.69 km², or approximately 25.34% of the total Java Island (BPS, 2010a). Its poverty rate was about 16.6% of its population in 2010 (BPS, 2010a). That number is higher than average percentage of poor people of Indonesia (13.3%) (BPS, 2010a). In 2011, the local government shows the success in declining the percentage of poor people in Central Java to around 15.76% (BPS, 2011).

Administratively, the province of Central Java is divided into 35 districts consisting of 29 regencies and 6 cities. The differences regarding the household expenditure level in the Central Java inter-district can be seen in Fig. 1. This Fig. 1 shows that the mean of household expenditure varies between districts and districts in urban areas have a higher household expenditure mean than rural areas.

Household expenditure distribution has a shape that close to a right skewed distribution such as log-normal. Battistin *et al.* (2007) state that Log-Normal distribution provides a useful theoretical model for studying certain economic population such income and expenditure distributions. Two parameter log-normal distribution however, is insufficient to capture the variation in the empirical distribution of household data

in Central Java. The three parameter Log-Normal distribution (LN3) therefore is applied to explain the variation of the data. The probability density function for LN3 is specified as follows:

$$f(y|\mu, \tau, \lambda) = \frac{1}{(y-\lambda)} \sqrt{\frac{\tau}{2\pi}} \exp\left(-\frac{\tau}{2}((y-\lambda)-\mu)^2\right) \quad (1)$$

where:

- $\mu > 0$ = The location parameter
- $\tau > 0$ = The scale parameter and
- $-\infty < \lambda < \infty$ = The threshold parameter

It is shown in Eq. 1 that LN3 has additional parameter, i.e., threshold parameter that shifts the whole of its distribution curve above zero. This characteristic represents the expenditure data which never has zero value.

Since household data is nested in its regional residence, it is classified as hierarchical platform. In this case, household expenditure can be influenced by factors from several different levels, i.e., factors at the household level and factors at the regional level.

Hierarchical models are formulated for analyzing data with complex sources of variation (Raudenbush and Bryk, 2002). Cases with complex sources of variation are frequently referred to the hierarchical structure of data (Goldstein, 1995; and Hox, 1995). Hierarchical data structure viewed data to be classified as a multilevel structure. Standard unilevel methods are not appropriate for analyzing such of hierarchical system (Maas and Hox, 2004), due to the parameter estimates are inefficient and standard error is negatively biased (Hox, 1995; Maas and Hox, 2004).

Raudenbush and Bryk (2002), Goldstein (1995) and Hox (1995) proposed hierarchical models for overcoming this kind of several different levels of hierarchical data modeling into a single statistical analysis. It is noted that hierarchical models, mostly use a classical approach in the estimation process. In the case of the complex hierarchical models, however, parameter estimation using the classical approach would be very difficult to be derived. Raudenbush and Bryk (2002) demonstrate that a hierarchical model using a classical approach works well when the data is balanced and the number of higher level unit is large. In some applications, however, this condition will not be easily hold.

MATERIALS AND METHODS

Residential conditions and facilities are frequently used as visual indicators to judge the level of socioeconomic welfare of the household. A number of studies, which have been done, show that several household attributes affect household expenditure, i.e., household size, education level of household head,

house area, types of wall, type of floor, source of drinking water, kitchen, toilet facilities and electricity (Ismartini *et al.*, 2011; Iriawan and Ismartini, 2011; Haughton and Nguyen, 2010; Mok *et al.*, 2007; and Grosh and Baker, 1995). This study will use predictors based on those previous studies, called micro variables and other predictors, called macro variables, that are investigated to influence household expenditure. Public service facilities are the example of macro variables. Since the availability of those facilities illustrates concrete steps of the local government policies in enlarging the person's welfare. The sample coverage area of data used in this study is a Central Java Province.

Preliminary analysis of the data is shown in Fig. 2 which demonstrates the pattern of simple regression lines for five districts in Central Java that have difference in both of their slopes and intercepts. This fact indicates that there are variations on district level or the presence of regional influence in which hierarchical analysis should be employed for analyzing this problem. This study proposes to model community characteristics and household attributes on household expenditure on Central Java Province, Indonesia, using a hierarchical Bayesian model based on the three parameter log-normal distribution.

Data descriptions: This study relies extensively on household expenditure data collected by the National Socioeconomic Surveys (Susenas) which have been conducted regularly by Statistics Indonesia (BPS). The dependent variable used in the model is household expenditure per capita (y). There are several household attributes as micro variables (X) and district characteristics as macro variables (W) that are considered as having affected the household expenditure per capita. Those variables are a type of house wall (X_1), type of house floor (X_2), floor area per capita (X_3), type of sources of drinking water (X_4), toilet facilities usage (X_5). Type of cooking fuel (X_6), household size (X_7), the level of household head education (X_8), whether the head of household working in agriculture (X_9), population density (W_1), ratio of primary school to primary school age children (W_2). The ratio of junior high school to junior high school age children (W_3), ratio of senior high school to senior high school age children (W_4), number of health facilities (W_5), number of medical personnel (W_6). The percentages of villages having public phone (W_7), a number of cooperative, that is an establishment that its members are people or establishments with the legal status of the cooperative and its activities based on peoples' economic movements (W_8). The number of large and medium enterprise (W_9), number of small/household industry (W_{10}), gross regional domestic product at current price per capita (W_{11}) and percentage contribution of revenue to budget revenue (W_{12}).

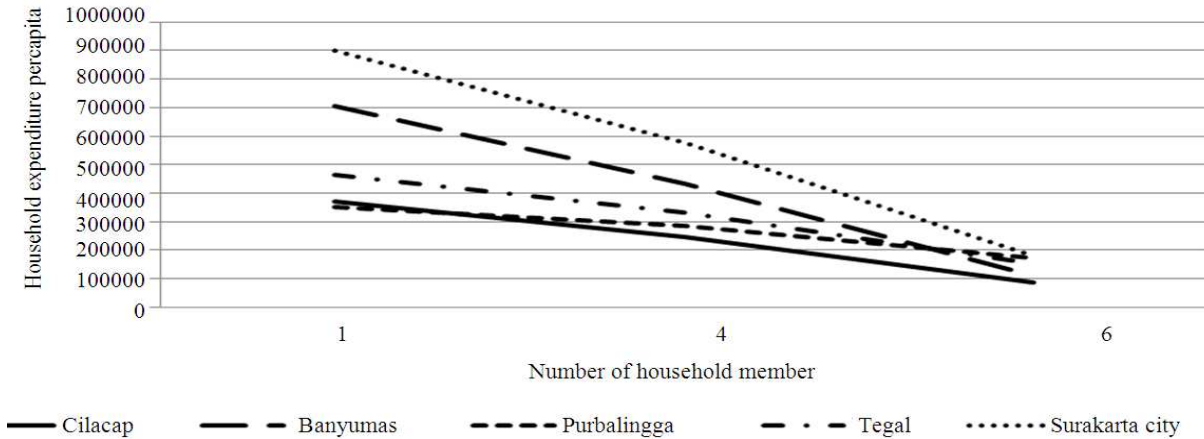


Fig. 2: Simple regression lines for five districts in central Java

Log-normal hierarchical models: A hierarchical model is formed by two sub-models, i.e., micro models (the models at a lower level) and macro models (models at higher levels) (Goldstein, 1995). For the two level hierarchical models of household expenditure in Central Java, the micro model investigates the association between household expenditure and various household attributes, while the macro model examines the relation among coefficients of micro model and district characteristics.

Suppose N is the number of households which is sampled from m districts and n_j is the number of households which is sampled in j^{th} districts, so

$\sum_{j=1}^m n_j = N$. Suppose y_j as a response in micro model and

x_j as micro variables where $j = 1, 2, \dots, m$. y_j is $n_j \times 1$ vector and X_j is $n_j \times p$ a matrix where $p = k+1$ and k represent a number of micro variables. Since $y_j \sim \text{LN3}(\mu_{[y_j]}, \tau_{[y_j]}, \lambda_j)$, the micro models based on Log-Normal distribution is specified as follows (Stata, 2009):

$$y_j = \exp(X_j \beta_j) \times r_j \text{ or } y_j^{\mathcal{L}} = X_j \beta_j + r_j^{\mathcal{L}} \quad (2)$$

where, $y_j^{\mathcal{L}} = \ln(y_j)$, r_j is the residual vector of micro models and $r_j^{\mathcal{L}} = \ln(r_j)$. β_j is $p \times 1$ coefficient vector of micro models. The macro models are, therefore can be specified as follows:

$$\beta_j = W_j \gamma + u_j, \quad (3)$$

where, W_j is $p \times q$ the matrix of macro variables with $q = l+1$ and l represent a number of macro variables, γ is the coefficient vector of macro models and u_j is the residual vector of macro models. The single equation models for Eq. 2 and 3 can be specified as follows:

$$y_j^{\mathcal{L}} = X_j W_j \gamma + X_j u_j + r_j^{\mathcal{L}} \quad (4)$$

Refer to Eq. 2, 3 and 4, the two level hierarchical Bayesian models for household expenditure in Central Java are defined as follows:

$$y_{ij}^{\mathcal{L}} = \beta_{0j} + \sum_{k=1}^9 \beta_{kj} X_{kij} + r_{ij}^{\mathcal{L}} ; i = 1, 2, \dots, n_j, j = 1, 2, \dots, 35 \quad (5)$$

$$\beta_{pj} = \gamma_{p0} + \sum_{l=1}^{12} \gamma_{pl} W_{lj} + u_{pj} ; p = 0, 1, 2, \dots, 9, l = 1, 2, \dots, 12 \quad (6)$$

Bayesian inference: Consider Bayes' Theorem (Box and Tiao, 1992; Gelman and Hill, 2007):

$$p(\theta | z) = \frac{p(z | \theta)p(\theta)}{p(z)} \quad (7)$$

where, θ and z are both random, θ is parameter vector and z denotes vector of observations from the sample. $p(z)$ is defined as normalized constant with respect to θ . Then, the posterior can be represented as a proportional form as follows:

$$p(q | z) \propto f(q | z)p(q), \quad (8)$$

It is shown in Eq. 8, the posterior is proportional to the combination of prior information and current information of data. All information about the unknown parameter of interest is included in their joint posterior distribution.

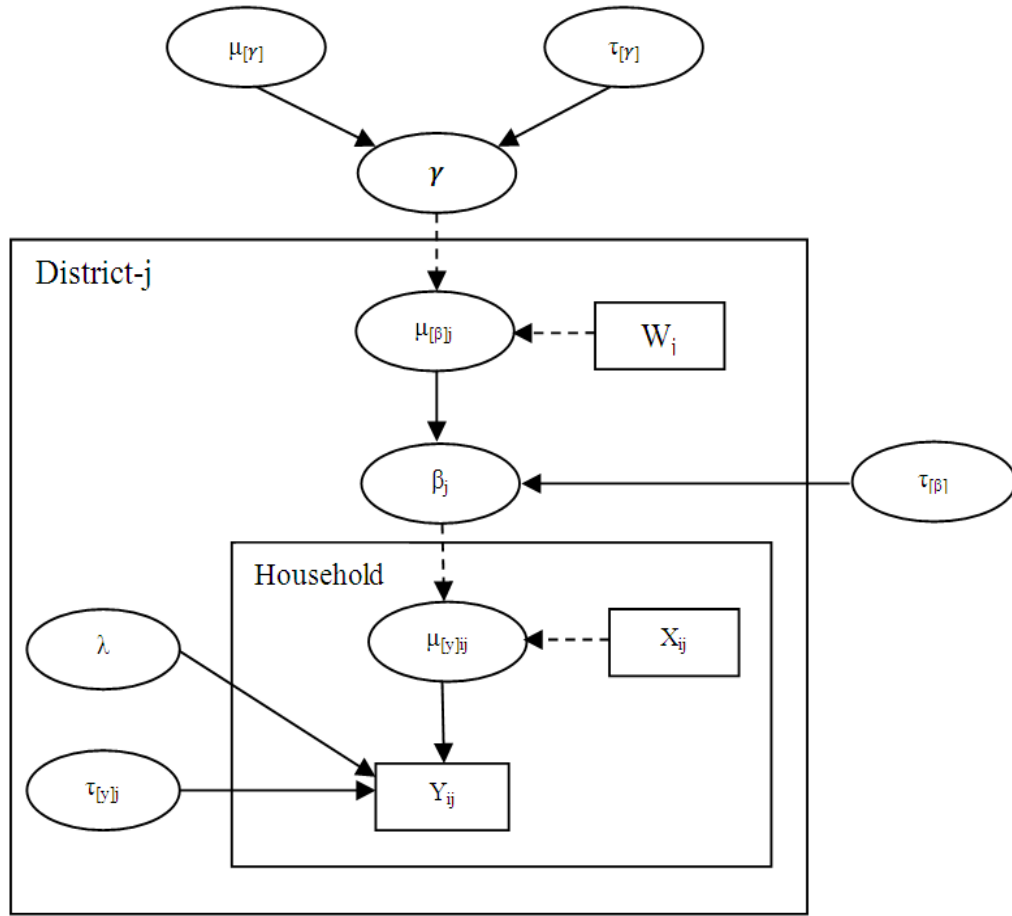


Fig. 3: DAG of two levels hierarchical Bayesian model for household expenditure in Central Java

Based on Eq. 7, the joint posterior distribution of the two level hierarchical models for household expenditure can be expressed as:

$$p(\beta, \gamma, \lambda, \tau_{[y]}, \tau_{[\beta]} | y) = \frac{f(y | \beta, \lambda, \tau_{[y]}) p_1(\beta | \gamma, \tau_{[\beta]}) p_2(\gamma, \lambda, \tau_{[y]}, \tau_{[\beta]})}{p(y)} \quad (9)$$

With:

$$p(y) = \int \int \int \int f(y | \beta, \lambda, \tau_{[y]}) p_1(\beta | \gamma, \tau_{[\beta]}) p_2(\gamma, \lambda, \tau_{[y]}, \tau_{[\beta]}) \pi(\beta) \pi(\gamma) \pi(\lambda) \pi(\tau_{[y]}) \pi(\tau_{[\beta]}),$$

or as in Eq. 8, the proportional form of Eq. 9 can be represented as:

$$p(\beta, \gamma, \lambda, \tau_{[y]}, \tau_{[\beta]} | y) \propto f(y | \beta, \lambda, \tau_{[y]}) p_1(\beta | \gamma, \tau_{[\beta]}) p_2(\gamma, \lambda, \tau_{[y]}, \tau_{[\beta]}). \quad (10)$$

where, $p_1(\beta | \gamma, \tau_{[\beta]})$ is a first stage prior for random parameters and $p_2(\gamma, \lambda, \tau_{[y]}, \tau_{[\beta]})$ is a second stage prior or hyper prior for hyper parameter. Eq. 10 is a proportional form of posterior for two level hierarchical model.

In Bayesian inference, all parameters need prior distribution. The nature of proposed prior distributions in this study is treated as independent prior distributions (Box and Tiao, 1992; Carlin and Chib, 1995) which are comprised combination of conjugate and informative prior distributions and pseudo prior.

Inference about the subset of focal parameters of interest is derived using its marginal conditional distribution. The marginal conditional distribution is calculated by integrating Eq. 10 with respect to auxiliary unknown parameters, which tend to complex numerical integration. To overcome that problem, Bayesian method is taking repeated samples from the full conditional posterior distribution using MCMC and Gibbs Sampling (Gelman *et al.*, 2004; Gelman and Hill, 2007; Ntzoufras, 2009).

The estimation of parameters of interest is implemented in WinBUGS 1.4 as a computational power of recent software for Bayesian computation.

Table 1: Estimated Coefficients regression and the Standard Deviation of Micro Model by Districts

Districts	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
Cilacap	12.120 (0.065)	0.181 (0.0404)	0.132 (0.0463)	0.006 (0.0008)	0.297 (0.0586)	0.169 (0.0372)	0.164 (0.0381)	-0.134 (0.0121)	0.339 (0.0505)	-0.137 (0.036)
Banyumas	12.110 (0.0729)	0.237 (0.0417)	0.190 (0.0578)	0.008 (0.0012)	0.421 (0.0589)	0.125 (0.039)	0.123 (0.0397)	-0.163 (0.0137)	0.542 (0.0497)	-0.022 ¹ (0.0428)
Purbalingga	11.780 (0.0774)	0.182 (0.0469)	0.158 (0.0551)	0.012 (0.0014)	0.203 (0.0507)	0.151 (0.0404)	0.186 (0.0417)	-0.062 (0.0139)	0.308 (0.0554)	-0.039 ¹ (0.0394)
Banjarnegara	12.190 (0.0775)	0.179 (0.039)	0.009 ¹ (0.0414)	0.007 (0.0013)	-0.222 ¹ (0.1559)	0.230 (0.0393)	0.276 (0.055)	-0.110 (0.0134)	0.337 (0.0649)	-0.169 (0.0335)
Kebumen	12.260 (0.0682)	0.079 (0.0396)	0.120 (0.0394)	0.003 (0.0007)	0.107 ¹ (0.0957)	0.195 (0.0381)	0.246 (0.0408)	-0.139 (0.0116)	0.313 (0.0543)	-0.037 ¹ (0.035)
Purworejo	12.070 (0.0963)	0.306 (0.0574)	0.145 (0.0586)	0.005 (0.001)	0.149 (0.074)	0.487 (0.0481)	0.117 (0.06)	-0.177 (0.018)	0.269 (0.059)	-0.013 ¹ (0.0446)
Wonosobo	12.270 (0.0742)	0.111 (0.0349)	0.158 (0.0409)	0.004 (0.0009)	0.082 (0.0378)	0.178 (0.0361)	0.286 (0.0445)	-0.112 (0.0126)	0.269 (0.0652)	-0.056 ¹ (0.0346)
Magelang	11.960 (0.0707)	0.202 (0.0463)	0.195 (0.0474)	0.008 (0.0009)	-0.004 ¹ (0.0554)	0.113 (0.0369)	0.280 (0.0428)	-0.106 (0.014)	0.219 (0.0489)	-0.152 (0.0372)
Boyolali	12.460 (0.0634)	0.199 (0.0379)	0.086 (0.0388)	0.003 (0.0006)	0.173 (0.0467)	0.160 (0.0379)	0.176 (0.0434)	-0.135 (0.0125)	0.221 (0.0464)	-0.089 (0.035)
Klaten	12.370 (0.0957)	-0.185 (0.078)	0.246 (0.0535)	0.004 (0.0006)	0.114 ¹ (0.0709)	0.141 (0.0389)	0.265 (0.0383)	-0.116 (0.0142)	0.265 (0.0398)	-0.060 ¹ (0.0373)
Sukoharjo	12.260 (0.0821)	0.221 (0.0622)	0.218 (0.0602)	0.003 (0.0007)	0.311 (0.052)	0.105 (0.0408)	0.010 ¹ (0.0437)	-0.161 (0.0131)	0.395 (0.0412)	-0.107 (0.0472)
Wonogiri	12.090 (0.1062)	0.115 (0.0496)	0.227 (0.0652)	0.003 (0.0009)	0.313 (0.0817)	0.231 (0.0579)	0.413 (0.0656)	-0.134 (0.0177)	0.251 (0.0736)	-0.254 (0.0437)
Karanganyar	11.640 (0.1392)	0.594 (0.1121)	-0.008* (0.091)	0.006 (0.0011)	0.242 (0.0598)	0.101 (0.0524)	0.397 (0.0547)	-0.139 (0.019)	0.399 (0.0607)	-0.025 ¹ (0.0552)
Sragen	12.450 (0.0869)	0.069 ¹ (0.0485)	0.186 (0.0468)	0.002 (0.0006)	0.158 (0.0502)	0.118 (0.0429)	0.122 (0.0526)	-0.131 (0.0175)	0.308 (0.0584)	-0.170 (0.0423)
Grobogan	12.470 (0.0589)	0.061 ¹ (0.0535)	0.152 (0.0387)	0.004 (0.0006)	0.107 (0.0498)	0.141 (0.0346)	0.087 (0.0345)	-0.117 (0.0139)	0.195 (0.0543)	-0.224 (0.0339)
Blora	12.120 (0.0771)	0.147 (0.0616)	0.192 (0.0432)	0.003 (0.0006)	0.367 (0.0538)	0.195 (0.0408)	0.196 (0.051)	-0.144 (0.0147)	0.270 (0.0573)	-0.010 ¹ (0.0364)
Rembang	12.040 (0.074)	0.112 (0.0417)	0.215 (0.04)	0.004 (0.0006)	0.111 (0.0476)	0.079 (0.0342)	0.246 (0.0442)	-0.054 (0.0162)	0.259 (0.0589)	-0.010 ¹ (0.0348)
Pati	11.990 (0.082)	0.047 ¹ (0.0474)	0.293 (0.0482)	0.007 (0.001)	0.105 (0.0463)	0.139 (0.0472)	0.200 (0.0432)	-0.110 (0.0165)	0.355 (0.057)	-0.053 ¹ (0.0399)
Kudus	11.550 (0.1202)	0.370 (0.0978)	0.302 (0.0717)	0.012 (0.0013)	0.245 (0.0622)	0.113 (0.0526)	0.095 (0.0449)	-0.085 (0.0174)	0.336 (0.0546)	-0.175 (0.05)
Jepara	11.940 (0.0649)	0.222 (0.0464)	0.188 (0.0404)	0.009 (0.0013)	0.151 (0.0634)	0.192 (0.0339)	0.164 (0.0379)	-0.102 (0.0116)	0.200 (0.0486)	-0.080 (0.0366)
Demak	12.160 (0.0522)	-0.007 ¹ (0.0349)	0.094 (0.037)	0.007 (0.0009)	0.138 (0.0339)	0.146 (0.0355)	0.138 (0.0341)	-0.075 (0.0094)	0.360 (0.0469)	-0.102 (0.0318)
Semarang	12.250 (0.0702)	0.247 (0.0491)	0.027 ¹ (0.0502)	0.005 (0.0009)	0.261 (0.0633)	0.108 (0.0436)	0.283 (0.0442)	-0.078 (0.0132)	0.248 (0.0549)	-0.162 (0.0373)
Temanggung	12.240 (0.0711)	0.264 (0.0443)	-0.023 ¹ (0.0471)	0.005 (0.0009)	0.134 (0.0504)	0.191 (0.0402)	0.220 (0.0483)	-0.109 (0.0142)	0.290 (0.0593)	-0.158 (0.0377)
Kendal	12.250 (0.0785)	0.168 (0.0412)	0.166 (0.0455)	0.006 (0.0008)	0.229 (0.0474)	0.014 ¹ (0.0418)	0.328 (0.0454)	-0.126 (0.0152)	0.282 (0.0565)	-0.072 ¹ (0.0429)
Batang	12.230 (0.0612)	0.150 (0.0447)	0.190 (0.0444)	0.001 (0.0005)	0.254 (0.0709)	0.131 (0.0405)	0.318 (0.0475)	-0.137 (0.0121)	0.125 (0.0615)	-0.075 (0.0366)
Pekalongan	12.040 (0.0658)	0.183 (0.0478)	0.160 (0.0495)	0.008 (0.0011)	-0.068 ¹ (0.0616)	0.119 (0.0373)	0.264 (0.036)	-0.116 (0.0107)	0.339 (0.0559)	-0.102 (0.0381)
Pemalang	12.450 (0.0618)	0.097 (0.0438)	0.068 ¹ (0.0423)	0.006 (0.0009)	0.068 ¹ (0.0508)	0.132 (0.0319)	0.173 (0.034)	-0.155 (0.0106)	0.266 (0.0504)	-0.075 (0.0332)
Tegal	12.180 (0.0826)	0.255 (0.0687)	0.068 ¹ (0.0571)	0.006 (0.001)	0.314 (0.0523)	0.115 (0.0332)	0.162 (0.035)	-0.107 (0.0132)	0.371 (0.0522)	-0.157 (0.0376)
Brebes	12.080 (0.0655)	0.252 (0.0459)	0.041 ¹ (0.0422)	0.008 (0.001)	0.094 (0.0427)	0.153 (0.0355)	0.162 (0.0368)	-0.094 (0.0112)	0.167 (0.0578)	-0.050 ¹ (0.0354)
Magelang City	12.310 (0.1242)	0.026 ¹ (0.0798)	0.393 (0.0756)	0.008 (0.0011)	0.083 ¹ (0.0478)	0.059 ¹ (0.055)	0.087 ¹ (0.0772)	-0.127 (0.0159)	0.456 (0.042)	-0.176 ¹ (0.1261)
Surakarta City	12.720 (0.1201)	0.246 (0.0712)	0.127 ¹ (0.1213)	0.004 (0.0006)	0.131 (0.0419)	0.166 (0.0512)	-0.163 ¹ (0.0566)	-0.145 (0.0123)	0.455 (0.0417)	-0.187 ¹ (0.1898)
Salatiga City	12.290 (0.0877)	0.213 (0.0651)	0.229 (0.084)	0.005 (0.0008)	0.308 (0.0458)	-0.025 ¹ (0.0627)	0.128 (0.0602)	-0.138 (0.0133)	0.457 (0.0502)	0.004 ¹ (0.0984)
Semarang City	12.270 (0.1057)	0.271 (0.0632)	0.247 (0.0769)	0.009 (0.0008)	0.264 (0.0446)	0.194 (0.0498)	-0.116 ¹ (0.0635)	-0.078 (0.0126)	0.440 (0.0422)	-0.197 (0.0919)
Pekalongan City	12.060 (0.1134)	0.212 (0.0748)	0.265 (0.0782)	0.009 (0.0009)	0.060 ¹ (0.0427)	0.119 (0.0496)	-0.025 ¹ (0.071)	-0.149 (0.0126)	0.422 (0.0447)	-0.100 ¹ (0.086)
Tegal City	12.540 (0.1581)	0.112 ¹ (0.1134)	0.285 (0.0631)	0.005 (0.0007)	-0.097 ¹ (0.0599)	0.136 (0.0429)	0.200 (0.0655)	-0.150 (0.0139)	0.321 (0.0426)	0.060 ¹ (0.065)

¹: Not significant () Standard deviation of the coefficient regression

Table 2: Estimated Coefficients regression and the standard deviation of Macro Model by estimating coefficients of micro model

Coefficient	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	γ_9	γ_{10}	γ_{11}	γ_{12}
β_0	12.180 (0.099)	0.014 (0.004)	-0.001 (1E-04)	0.069 (0.01)	0.073 (0.01)	-0.020 (0.01)	-0.075 (0.01)	0.004 (3E-04)	0.002 (0.001)	-0.631 (0.01)	0.659 (0.01)	-0.005 (3E-04)	0.001 (1E-04)
β_1	0.149 (0.01)	-0.002 (0.001)	0.021 (0.01)	-0.017 (0.001)	-0.017 (0.001)	-0.002 ¹ (0.001)	-0.001 ¹ (0.001)	0.002 (3E-04)	-0.099 (0.001)	-0.202 (0.011)	0.001 (3E-04)	0.006 (3E-04)	0.001 (1E-04)
β_2	0.080 (0.010)	-0.015 (0.001)	-0.004 (0.001)	0.035 (0.010)	0.025 (0.010)	0.002 (0.010)	0.082 (0.010)	0.004 (3E-04)	1E-04 (3E-05)	0.067 (0.010)	-0.385 (0.010)	-0.718 (0.010)	0.001 (1E-04)
β_3	0.001 (3E-04)	-0.0003 (1E-04)	-0.774 (0.001)	0.003 (0.001)	-0.002 (0.001)	0.064 (0.001)	0.025 (0.002)	0.070 (0.010)	-0.004 (3E-04)	-0.001 ¹ (0.002)	0.001 (3E-04)	0.082 (0.010)	-0.001 (1E-04)
β_4	0.221 (0.100)	0.005 (0.002)	0.037 (0.010)	-0.049 (0.010)	0.048 (0.010)	-0.001 (0.010)	-0.001 (1E-04)	0.009 (0.003)	3E-5 ¹ (1E-04)	-0.164 (0.011)	-3E-04 (3E-05)	-0.666 (0.01)	0.002 (1E-04)
β_5	0.176 (0.032)	0.001 ¹ (0.001)	0.006 (0.001)	0.006 (0.001)	-0.019 (0.001)	-0.771 (0.001)	-0.220 (0.010)	0.004 (3E-04)	0.325 (0.010)	0.097 (0.009)	3E-04 (3E-05)	-0.001 (3E-04)	0.001 (1E-04)
β_6	0.114 (0.001)	-0.003 (0.001)	0.014 (0.001)	-0.037 (0.010)	0.035 (0.010)	0.002 (0.010)	-0.146 (0.001)	-0.008 (0.003)	0.001 (3E-04)	-0.008 (3E-04)	0.406 (0.010)	0.001 (3E-04)	0.001 (1E-04)
β_7	-0.166 (0.003)	-0.005 (0.001)	-0.007 (0.001)	0.030 (0.010)	-0.021 (0.010)	0.001 (0.010)	0.202 (0.010)	0.002 (3E-04)	-0.151 (0.003)	0.031 (0.003)	-0.096 (0.003)	0.0004 (3E-05)	0.0002 ¹ (1E-04)
β_8	0.280 (0.100)	-0.001 ¹ (0.001)	-0.025 (0.010)	0.007 (0.001)	0.022 (0.010)	0.003 (0.010)	-0.050 (0.010)	0.893 (0.003)	-0.216 (0.010)	-1E-04 (3E-05)	0.386 (0.010)	0.002 (3E-04)	0.001 (1E-04)
β_9	-0.122 (0.010)	-0.001 ¹ (0.001)	-0.010 (0.001)	0.003 (0.001)	0.036 (0.010)	0.693 (0.010)	0.238 (0.010)	-0.006 (3E-04)	-0.665 (0.010)	-0.072 (0.010)	-0.045 (0.010)	-0.001 (3E-04)	0.002 (1E-04)

¹Not significant () Standard deviation of the coefficient regression

This software is an interactive windows version of the BUGS program for Bayesian analysis by implementing MCMC techniques and Gibbs sampling. The algorithm to generate the estimated parameter $\beta, \gamma, \lambda, \tau_{[y]}$, and $\tau_{[\beta]}$ are:

Step 1: Choose initial value for all focal parameters of interest $\beta^{(0)}, \gamma^{(0)}, \lambda^{(0)}, \tau_{[y]}^{(0)}$, and $\tau_{[\beta]}^{(0)}$

Step 2: Run sequential process below iteratively to generate a T sample of the parameter of interest using full conditional distribution after the equilibrium distribution is reached. The equilibrium distribution is the target posterior distribution of interest.

For $a = 1$ to T, where T is the number of iterations

Generate $\beta^{(a)}$ from $p(\beta | \gamma^{(a-1)}, \lambda^{(a-1)}, \tau_{[y]}^{(a-1)}, \tau_{[\beta]}^{(a-1)}, y)$

Generate $\gamma^{(a)}$ from $p(\gamma | \beta^{(a-1)}, \lambda^{(a-1)}, \tau_{[y]}^{(a-1)}, \tau_{[\beta]}^{(a-1)}, y)$

Generate $\lambda^{(a)}$ from $p(\lambda | \beta^{(a-1)}, \gamma^{(a-1)}, \tau_{[y]}^{(a-1)}, \tau_{[\beta]}^{(a-1)}, y)$

Generate $\tau_{[y]}^{(a)}$ from $p(\tau_{[y]} | \beta^{(a-1)}, \gamma^{(a-1)}, \lambda^{(a-1)}, \tau_{[\beta]}^{(a-1)}, y)$

Generate $\tau_{[\beta]}^{(a)}$ from $p(\tau_{[\beta]} | \beta^{(a-1)}, \gamma^{(a-1)}, \lambda^{(a-1)}, \tau_{[y]}^{(a-1)}, y)$

The concept of that iterative estimation process is generated by Winbugs derived from Directed Acyclic Graph (DAG) of the hierarchical model. Figure 3 shows DAG of two level hierarchical Bayesian model for household expenditure in Central Java as the implementation of Eq. 5 and 6.

RESULTS

The two level hierarchical Bayesian model for household expenditure in Central Java Province is

developed based on DAG in Fig. 3. As the starting step, the modeling focuses to determine an appropriate prior for parameter and hyperparameter of the model. The prior distribution of $\beta, \gamma, \lambda, \tau_{[y]}$, and $\tau_{[\beta]}$ are as follows:

$$\begin{aligned} \beta &: N(\mu_{[\beta]}, \tau_{[\beta]}^{-1}) \\ \lambda &: N(\mu_{[\lambda]}, \tau_{[\lambda]}^{-1}) \\ \gamma &: N(\mu_{[\gamma]}, \tau_{[\gamma]}^{-1}) \\ \tau_{[y]} &: \text{Gamma}(0.1, 0.001) \\ \tau_{[\beta]} &: \text{Gamma}(1, 0.001) \end{aligned}$$

where, $\mu_{[\beta]}, \tau_{[\beta]}^{-1}, \mu_{[\lambda]}, \tau_{[\lambda]}^{-1}, \mu_{[\gamma]},$ and $\tau_{[\gamma]}^{-1}$ are fixed values.

The results in Table 1 and 2 show the significant estimated coefficients of micro model and macro model, respectively.

DISCUSSION

The two hierarchical Bayesian model shows that household welfare levels in Central Java, generally, can be indicated by several household attributes. First, the household welfare can be specified from housing condition such as, a good type of wall and floor and size of floor area per capita. Second, in majority, the welfare can also be identified by the availability of daily needs facilities such as, clean water sources, toilet ownership and a good cooking fuel. Third, Human capital of household for instance, the number of people in the household and level of education of household head affect the household welfare as well. The fact in 18 districts shows that household which generally

economically active in agriculture sector has lower welfare level than others. According to BPS (2010c), those 18 districts mainly have a high percentage of wetland area and poverty level compare to other districts. For example, Brebes has almost 37.73% of its area is dominated by wetland and its percentage of poor people stands the fifth highest percentage among districts in Central Java (24.39%).

District characteristics do affect positively to household welfare through the specific household attributes. Those districts characteristics are demographic and economic conditions of districts and the availability of public facilities, i.e., Economic, education and health which are strongly associated with a dimension of human development index. This relation shows that better availability of those public facilities yields higher welfare of the people. In terms of the economic dimension, number of small/household industry has also a positive effect on household welfare. This is reasonable since industry can create job opportunities for the people therein.

CONCLUSION

This study has already demonstrated the work of the developed model for estimating household expenditure in order to measure the effect of regional diversity by taking into account district characteristics and household attributes using a hierarchical Bayesian approach based on the three parameters of the log-normal distribution. The result shows that the regional diversities do affect the household expenditure therein. The local government effort in providing public facilities statistically can improve its people welfare. Other interesting future research perspective is to investigate other specific district characteristics and household attributes that might affect household expenditure.

ACKNOWLEDGMENT

The researchers would like to express a sincere gratitude to Statistics Indonesia for providing funds for this research as a part of the PhD grant research program and for its human resource development projects in cooperation with Institut Teknologi Sepuluh Nopember, Indonesia.

REFERENCES

Akita, T. and A. Pirmansyah, 2011. Urban inequality in Indonesia. IUJ Research Institute.
Battistin, E., R. Blundel and A. Lewbel, 2007. Why is consumption more log normal than income? gibrat's law revisited. Institute Fiscal Studies.

Box, G.E.P and G.C. Tiao, 1992. Bayesian Inference in Statistical Analysis. 1st Edn., Wiley, New York, ISBN-10: 0471574287 pp: 588.
BPS, 2010a. Statistical pocketbook of Indonesia. BPS.
BPS, 2010b. Analysis of poverty, employment and income distribution. BPS.
BPS, 2010c. Central Java in figures 2010. BPS.
BPS, 2011. Statistics official news BPS central Java Province. BPS.
Carlin, B.P. and S. Chib, 1995. Bayesian model choice via Markov chain Monte Carlo methods. *J. Royal Stat. Soc.*, 57: 473-484.
Gelman, A. and J. Hill, 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. 1st Edn., Cambridge University Press, New York, ISBN-10: 0521867061 pp: 625.
Gelman, A., J.B. Carlin, H.S. Stern and D.B. Rubin, 2004. Bayesian Data Analysis. 2nd Edn., Chapman and Hall/CRC Press, Boca Raton,, ISBN-10: 158488388X pp: 668.
Goldstein, H., 1995. Multilevel Statistical Models. 2nd Edn., Arnold, London, ISBN-10: 111995682X, pp: 382.
Grosh, M.E. and J.L. Baker, 1995. Proxy Means Tests for Targeting Social Programs: Simulations and Speculation. 1st Edn., World Bank, Washington, DC., ISBN-10: 0821333135, pp: 49.
Houghton, D. and P. Nguyen, 2010. Multilevel models and inequality in Viet Nam. *J. Data Sci.*, 8: 289-306.
Hox, J.J., 1995. Applied Multilevel Analysis. 1st Edn., TT-Publikaties, Amsterdam, pp: 119.
Iriawan, N. and P. Ismartini, 2011. On the modeling of district policy effects to household expenditure: a hierarchical bayesian approach. Proceeding of the 7th Conference of the Asian Regional Section of the IASC, Dec. 16-19, Academia Sinica, Taipei, Taiwan.
Ismartini, P., N. Iriawan, Setiawan and B.S.S. Ulama, 2011. A hierarchical structured data analysis for household expenditure using a bayesian hierarchical model. Proceedings of the International Conference on Mathematics and Sciences, Oct. 12-13, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.
Maas, C.J.M. and J.J. Hox, 2004. Robustness issues in multilevel regression analysis. *Stat. Neerlandica*, 58: 127-137. DOI: 10.1046/j. 0039-0402.2003.00252. x
Mok, C., C. Gan and A. Sanyal, 2007. The Determinants of Urban Household Poverty in Malaysia. *J. Soc. Sci.*, 3: 190-196. DOI: 10.3844/jssp.2007.190.196

- Ntzoufras, I., 2009. Bayesian Modeling Using WinBUGS. 1st Edn., Wiley, New Jersey, ISBN-10: 1118210352, pp: 520.
- Raudenbush, S.W. and A.S. Bryk, 2002. Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd Edn., Sage Publications, Thousand Oaks, ISBN-10: 076191904X pp: 485.
- Shahateet, M.I., 2006. How Serious Regional Economic Inequality in Jordan? Evidence from Two National Household Surveys. *Am. J. Applied Sci.*, 3: 1735-1744. DOI: 10.3844/ajassp.2006.1735.1744
- Stata, C., 2009. Stata Survival Analysis and Epidemiological Tables Reference Manua Release 11. 1st Edn., StataCorp LP, College Station, Texas, ISBN-10: 1597180610 pp: 548.