Health Expenditure Modelling in Southeast of Iran: A Population-Based Setting Using Quantile Regression Perspective

Abolfazl Payandeh,1 Yadollah Mehrabi,2,* Farid Zayeri,3 and Zahra Rezaei Ghahroodi4

1Department of Biostatistics, School of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3Department of Biostatistics and Proteomics Research Center, School of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran
4Statistical Research and Training Center, Tehran, Iran

*Corresponding author: Yadollah Mehrabi, Department of Epidemiology, School of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Tel: +98-9123849707, E-mail: mehrabi@sbmu.ac.ir

Received 2016 November 20; Revised 2017 February 04; Accepted 2017 March 13.

Abstract

Background: Equity in health is the focus of attention in the world health circles in recent decades. The financing of household health expenditure is, therefore, a concern in any region.

Objectives: The current study aimed at exploring the determinants of household health expenditure in Sistan and Baluchistan province, Southeast of Iran.

Methods: The data of the current study were a part of the Iranian households’ income and expenditure survey, administered by the Iranian statistical center (ISC) from March 2013 to March 2014. A 3-stage stratified cluster sampling method was employed in the current cross sectional survey. A total of 1391 households from Sistan and Baluchistan province were randomly selected. The study employed quantile regression modelling to determine the effective explanatory variables of household health expenses. R programming language version 3.3.2 was used for data analysis.

Results: It was found that annual median per capita medical expenditure was US$ 5.36 (first quartile = 0; third quartile = 38.31). Households also spent about 3.9% of their annual income on health expenses per capita. Literacy, occupational status of household head, and family income had significant effect on the median of household health expenses.

Conclusions: Researchers and policy makers should pay more attention to household health expenses, as it is an important matter in the low-income and deprived areas.

Keywords: Health Expenditure, Household, Quantile Regression, Sistan and Baluchistan

1. Background

Health is one of the crucial factors for the economic development of countries (1). A healthy population is likely to lead to higher economic added value (2). The health care issue is the millennium development goal set by the United Nations in 2000 and signed by nearly 190 countries (2, 3). Universal health coverage is the goal in which all the people obtain the health services they need without risking financial hardship. Thus, equity in health is the focus of attention in the world health circles in recent decades (4, 5). Equity in health is defined as the absence of systematic differences in 1 or more aspects of health status across the socially, demographically or geographically defined populations or population subgroups (6). The financing of household health expenditure (HHE) is, therefore, a concern in any country. It is shown that HHE in the organization for economic cooperation and development (OECD) varies substantially over time across countries (7, 8).

High health expenses can seriously affect the living standards (1). Poor households are particularly vulnerable to high health expenses that deal with pain and trouble, while facing with the increasing costs for health services. As a consequence, they reduce other essential family expenditures (9). The investigation of determinants of HHE as a percent of gross domestic product (GDP) is increasing since the 1970s among developed countries to explain differences across population subgroups (9-11).

Sistan and Baluchistan, the widest province of Iran with 11.5% of total country area, should be investigated regarding health services as a sample of low-income regions. Several studies were devoted to the study of household expenditure (2, 11-18).

Household expenditure and income data often do not meet the assumptions needed for the majority of statistical methods in the real world. In particular, the ho-
mosedasticity assumption frequently fails, and focusing exclusively on central tendencies can fail to capture informative trends in the response distribution. Ignoring this fact may reduce the efficiency of the tests (19). Quantile regression is a very effective method to analyze such information. It models the relationship between predictors (X’s) and the conditional quantiles of response (Y) given X = x. It is especially useful in applications where extremes are important (20). The quantile regression model is a natural extension of the linear classical regression model, which does not rely on limitative assumptions. To the authors’ best knowledge, few publications can be found in the literature that address the issue of HHE evaluation and health inequalities in low and middle income regions using quantile regression modelling. This approach is utilized in many applications such as quality of schooling (21), infant birth weight (22), genetics (23), malnutrition and growth curves (24, 25), ophthalmology (26), and public health (27).

2. Objectives

Based on the presented approach, the current study aimed at exploring the determinants of HHE using quantile regression methodology in a population-based setting.

3. Methods

3.1. Study Design

The data used in the current analysis were a part of the households’ income and expenditure survey (HIES), a cross sectional study administered by the Iranian statistical center (ISC) from March 2013 to March 2014. The HIES was implemented since 1963 and 1968 in rural and urban areas of Iran, respectively. It generally aimed at estimating the average expenditure and income for households nationwide.

3.2. Study Population and Sampling Methods

The HIES target population includes all privately settled and collective households. A 3-stage stratified cluster sampling method was employed in the study. The census areas, the urban and rural blocks, and the households were classified and selected in the 1st, 2nd, and 3rd stages, respectively. The sample size of 1391 households was optimized to estimate the average annual income and expenditure of a household based on the goal of the study.

3.3. Inclusion and Exclusion Criteria

All privately settled and collective households selected as samples were included in the study, unless the household did not cooperate for any reasons mainly, unwillingness, mental disabilities, absence of the family during sampling period, and houses with no residents (28).

3.4. Data Collection and Construction of Main Variables

3.4.1. Data Collection

The questionnaire included different sections that covered demographic characteristics, education, marital status, primary life equipment, access to facilities, housing, monthly and yearly food and non-food expenditure, and yearly income of the household. Health expenses section covered expenditure questions on treatment, medicines, laboratory tests and diagnostics, hospitalization, healthcare products, remedial equipment, outpatient services, visits to traditional healers, surgical operations and instruments, dentistry, addiction therapy, and other health-related expenditures for all household members. The earnings of all household members who worked to obtain the household income were aggregated. It should be noted that the study focused on out-of-pocket health expenses in the analysis. Insurance premium and governmental financial supports on HHE were ignored. According to the purchasing power parity (PPP), in average, one U.S dollar was equal to 22370 Rls (Iranian currency) in the period of data collection (29).

3.4.2. Dependent Variable

Health expenditure was calculated as the sum of all expenditure for treatment, health care, and medical equipment of all family members in the past 12 months before the interview.

3.4.3. Predictor Variables

Predictive variables were chosen based on available data and related literatures. The variables were as follow: 1, Household head variables: Age, gender (male / female), literacy status (literate / illiterate), and occupational status (employed / unemployed / unemployed, but has income / student / housewife / etc.); 2, Household variables: living area (rural / urban), annual income, and size of the family (number of household members). For simplicity, household head occupational status variable was recategorized into 2 new categories (with income (employed + unemployed, but has income / student + housewife + etc.). Age was also divided by 10 for better interpretation. The HHE per capita and income per capita was used through analyses and the variable weight was used for sample size determination and estimation precision. Raw data were transferred into the Microsoft Access 2010. Then, variables of interest were converted into the Microsoft Excel 2010, and statistical programming environment R version 3.2.2.

3.5. Statistical Analysis

For categorical variables, contingency tables with frequency and percentage were constructed in each cell.
Payandeh A et al.

continues variables, mean, standard deviation (SD), first quartile (Q1), median (Q2), and third quartile (Q3) values were estimated. Note that the analyses were based on the sampling weights. The right-skewed distribution of the HHE and income were propounded logarithmic transformation for modeling purposes. Quantile regression approach was employed to model the relationship between HHE and predictor variables. This modelling methodology is preferred when distribution of outcome of interest is skewed and/or the modelling of response extremes investigated. In this approach, regression model to estimate $\tau$th quantile of HHE can be written as follows:

$$Q_y(\tau \mid x) = x^T \beta(\tau) + \varepsilon$$  \hspace{1cm} (1)

Where $\tau \in (0, 1)$, $\beta \in R$, and $x$ represents the matrix of explanatory variables. All computations were performed using the package quantreg (30) for the statistical programming environment R (31).

4. Results

4.1. Descriptive Statistics

The overall measurement results are summarized in Table 1. About 14.4% of households were female headed, which spent more on health than their male counterparts. A large number of heads (36.4%) were illiterate. The results also revealed that 21.7% of heads had no income. The average number of family members were 4.6 (3.6 for Iran) and 25% of family heads were above 53 years old.

4.2. Distribution of Household Expenditure

It was found that the annual median per capita of medical expenses and income were SUS 5.36 (Q1 = 0 and Q3 = 38.31) and 774.12 (Q1 = 495.17 and Q3 = 1407.60) SUS, respectively. Table 2 shows the mean ± SD for HHE and income per capita by expenditure deciles. The proportion of annual income per capita spending on total and health expenses were increased with expenditure deciles.

4.3. Results of Quantile Regression Modeling

Table 3 outlines the results of linear quantile and classic regression models. Quantile regression results depicted that predictors had different impacts on HHE as quantile changes, while classic regression estimations showed a fixed effect for each predictor. The effects of different quantiles on each predictor were illustrated in more details in Figure 1 and Table 3. It is one of the main characteristics of quantile regression. It means that predictors had different effects on the households with low, middle, and high health expenditure. The current study findings indicated that income, literacy, and occupational status were the significant determinants at the 50th percentile of HHE in quantile model; while age and income were identified as substantial factors of HHE in the classic regression model. Figure 1 plots quantile model coefficients of each predictor versus quantiles, and illustrates the changes of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region, Freq. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>597 (42.0)</td>
<td>114 (8.3)</td>
<td>711 (49.7)</td>
</tr>
<tr>
<td>Urban</td>
<td>596 (43.6)</td>
<td>84 (6.1)</td>
<td>680 (50.3)</td>
</tr>
<tr>
<td>Literacy of the household head, Freq. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literate</td>
<td>819 (60.1)</td>
<td>44 (3.5)</td>
<td>863 (63.6)</td>
</tr>
<tr>
<td>Illiterate</td>
<td>374 (25.5)</td>
<td>154 (10.9)</td>
<td>528 (36.4)</td>
</tr>
<tr>
<td>Occupational status of the household head, Freq. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With income</td>
<td>949 (69.7)</td>
<td>120 (8.7)</td>
<td>1069 (78.3)</td>
</tr>
<tr>
<td>Without income</td>
<td>244 (16.0)</td>
<td>78 (5.7)</td>
<td>322 (21.7)</td>
</tr>
</tbody>
</table>

Table 2. Mean (SD) for HHE and Income Per Capita for Sampled Sistan and Baluchistan Province Households by Total Expenditure Decile

<table>
<thead>
<tr>
<th>Expenditure Decile</th>
<th>HHE (1000 Rls)</th>
<th>Income (1000 Rls)</th>
<th>% (HHE/Income)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>387 (949)</td>
<td>15171 (10012)</td>
<td>2.8</td>
</tr>
<tr>
<td>2</td>
<td>489 (948)</td>
<td>18700 (10012)</td>
<td>2.6</td>
</tr>
<tr>
<td>3</td>
<td>542 (167)</td>
<td>20830 (10207)</td>
<td>2.5</td>
</tr>
<tr>
<td>4</td>
<td>714 (1488)</td>
<td>23789 (14938)</td>
<td>3.7</td>
</tr>
<tr>
<td>5</td>
<td>1058 (2783)</td>
<td>39380 (29433)</td>
<td>3.7</td>
</tr>
<tr>
<td>6</td>
<td>1262 (2405)</td>
<td>36244 (24770)</td>
<td>5.5</td>
</tr>
<tr>
<td>7</td>
<td>1174 (2250)</td>
<td>46504 (44530)</td>
<td>3.7</td>
</tr>
<tr>
<td>8</td>
<td>1325 (2040)</td>
<td>38330 (24814)</td>
<td>4.0</td>
</tr>
<tr>
<td>9</td>
<td>2637 (5004)</td>
<td>51897 (37787)</td>
<td>6.4</td>
</tr>
<tr>
<td>10</td>
<td>3926 (1134)</td>
<td>60487 (38252)</td>
<td>19.3</td>
</tr>
<tr>
<td>Total</td>
<td>874 (2852)</td>
<td>26858 (27472)</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Abbreviations: Q1, First Quartile; Q2, Median; Q3, Third Quartile; SD, Standard Deviation.

41.3 \pm (51.20)

model coefficients based on quantile changes. For example, income had a descending effect on HHE as percentile increased. The exponentials of regression coefficients ($\beta$) were also reported in Table 3 for simplicity of interpretations. For instance, for a specified family with HHE about median ($US \ 5,36$), if the age of household head increased 10 years, its HHE increased about $e^{0.03} = 1.03$ times ($\approx 3\%$). Likewise, if annual income per capita of a specified family with illiterate head increased $US \ 447.03$ (~$10$ million Rls), it spent $1.12$ times more (or $12\%$ increase) on HHE. This expenditure also enhanced $1.04$ times (or $4\%$ increase) in households with literate heads.

Since a high portion of households did not report any out-of-pocket expenditure on health, estimations for lower quantiles ($< 0.4$) were not reported. Thus, the model for selected quantiles (i.e., $\geq 0.4$) was reported.

5. Discussion

The current study mainly aimed at drawing attentions toward the analysis of annual HHE data. Quantile regression results indicated that income, literacy level, and occupational status had significant effects on median of HHE. As shown in the previous section, the effects of predictors differed in sign and size for distinct quantiles. It is one of the key features of the quantile regression method over classical regression models.

Age was a positively effective predictor on upper quantile of annual HHE, which concurred with other studies (13, 14). It is well understood that elderly populations require more health services, which results in higher health expenditure (13). Previous studies showed that the effect of income was in doubt. The current study results revealed that income had a strong impact on HHE in the whole range of quantiles, which was in agreement with the results of Li et al. (13). Descriptive analysis showed that the low HHE was observed in low-income households. Policymakers and authorities should enhance health equipment and insurance in low-income areas. Recessions and economic instability have a potentially adverse effect on health.

The present study considered the effect of household head gender on HHE. The results showed that average annual HHE for female headed households were more (about $17\%$) than that of their male counterparts. The unhealthy lifestyle of female headed families, which is due to literacy level and income, might be the main reason.

Although the effect of living area was not substantial, the results indicated that urbanization makes families spend more on HHE in upper quantiles. The reason may be due to the difficulties to access the health services by rural residents. Access to health services has 3 dimensions: physical, financial, and acceptability. Geographical barrier and distance from village to town health centers may lead to ignore health seeking in the early stage of diseases (32). People’s ability to pay for services without financial hardship provides the impetus for changes in the way services are used. These financial charges are also taken into account of indirect costs (e.g., the costs of transportation to and from services and taking time away from work). Affordability is influenced by health financing system and by household income. Therefore, access improves by reducing poverty and income inequalities (14, 32). In addition, residents of urban areas have the higher average of annual income and their patients tend to be visited by more skillful health providers (33) in the early stage of diseases, which costs them lower expenses. In contrast, as mentioned earlier, rural patients do not refer to city hospitals except in serious stages of the diseases.

The current study findings indicated that families with illiterate heads spend less on HHE than their literate counterparts. The possible reason could be that literate and educated individuals tend to follow a healthy lifestyle and spend more on their health (1, 14).

The main limitation of the present study was that although the HIES data were recorded using educated persons and was carried out precisely with ISC, the income and expenditures information are self-reported for the year prior to the interview. There might be recall bias in reports especially in expenditures. It cannot be ascertained that possible inaccuracies in recall occur similarly for income or expenditures, which were not verifiable from other sources. In addition, zeroes were quite prevalent in HHE data. It might be due to insurance coverage and government supportive policies in deprived areas.

It can be concluded that more attention should be paid to HHE by researchers and policy makers, since it is an important matter in low-income and deprived areas. Robust statistical models, such as linear quantile regression method, can be employed to make reliable inferences about determination of effective predictors on HHE.

Acknowledgments

The authors are grateful the Iranian statistical center (ISC) for the data. The manuscript is part of a PhD thesis (No. 93/65). Authors also would like to thank the reviewers for their valuable comments.

Footnotes

Authors’ Contribution: Yadollah Mehrabi, manuscript idea developing, analyzing the data, and editing the final version of the manuscript; Abolfazl Payandeh, data analysis and writing sections of the first draft; Farid Zayeri and

### Table 3. Results of Quantile Regression Models for Sampled Households

<table>
<thead>
<tr>
<th>Predictors</th>
<th>40% Quantile Regression Model</th>
<th>50% Quantile Regression Model</th>
<th>75% Quantile Regression Model</th>
<th>90% Quantile Regression Model</th>
<th>Linear Classical Multiple Regression Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β (SE)</td>
<td>β (SE)</td>
<td>β (SE)</td>
<td>β (SE)</td>
<td>β (SE)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-10.86 (2.91)^a</td>
<td>-10.86 (2.91)^a</td>
<td>-10.86 (2.91)^a</td>
<td>-10.86 (2.91)^a</td>
<td>-10.86 (2.91)^a</td>
</tr>
<tr>
<td>Age</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Log. income. peC</td>
<td>2.75 (0.26)^a</td>
<td>2.75 (0.26)^a</td>
<td>2.75 (0.26)^a</td>
<td>2.75 (0.26)^a</td>
<td>2.75 (0.26)^a</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Region (rural)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Literacy (illiterate)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Occupation (without income)</td>
<td>10.86 (2.91)^a</td>
<td>5.25 (2.91)^a</td>
<td>5.25 (2.91)^a</td>
<td>5.25 (2.91)^a</td>
<td>5.25 (2.91)^a</td>
</tr>
<tr>
<td>Log. income. peC × occupation</td>
<td>-2.75 (0.26)^a</td>
<td>-2.75 (0.26)^a</td>
<td>-2.75 (0.26)^a</td>
<td>-2.75 (0.26)^a</td>
<td>-2.75 (0.26)^a</td>
</tr>
<tr>
<td>Region × literacy</td>
<td>-0.04 (0.14)</td>
<td>-0.04 (0.14)</td>
<td>-0.04 (0.14)</td>
<td>-0.04 (0.14)</td>
<td>-0.04 (0.14)</td>
</tr>
<tr>
<td>Log. income. peC × literacy</td>
<td>-1.04 (0.20)^a</td>
<td>-1.04 (0.20)^a</td>
<td>-1.04 (0.20)^a</td>
<td>-1.04 (0.20)^a</td>
<td>-1.04 (0.20)^a</td>
</tr>
<tr>
<td>Age × occupation</td>
<td>-0.15 (0.20)^a</td>
<td>-0.15 (0.20)^a</td>
<td>-0.15 (0.20)^a</td>
<td>-0.15 (0.20)^a</td>
<td>-0.15 (0.20)^a</td>
</tr>
<tr>
<td>Gender × occupation</td>
<td>-0.31 (0.20)^a</td>
<td>-0.31 (0.20)^a</td>
<td>-0.31 (0.20)^a</td>
<td>-0.31 (0.20)^a</td>
<td>-0.31 (0.20)^a</td>
</tr>
</tbody>
</table>

*Indicate significance at 0.01% levels, respectively.

---

**Figure 1.** Estimation of Quantile Regression Coefficients and 95% Confidence Limits at Different Quantiles in Model with Main Effects
Zahra Rezaei Ghahroodi, interpretation of the results and editing the manuscript. All authors read and approved the final manuscript.

Funding/Support: There was no financial support for the research.

References

28. SCI . Summary results of the Iranian urban and rural household income and expenditure survey. Iran: Statistical Center of Iran; 2014.

Payandeh A et al.