NORMALIZING CONSUMER SURPLUS DATA FOR KOSOVO’S WTP FOR A MANDATORY HEALTH INSURANCE

Edmond Muhaxheri  
Faculty of Economy, University of Tirana, Albania  
e-mail: edmond.muhaxheri@gmail.com

Patrycja Szkudlarek  
Faculty of Economics and Sociology, University of Lodz, Poland  
e-mail: patia.1@wp.pl

Abstract: The purpose of this paper is to show that consumer surplus for Kosovo’s expected mandatory health insurance fund do not follow a normal distribution. It shows the rationale used in obtaining the initial aggregate consumer surplus, the development Surplus-to-Exploitation, and Potential Entry Threshold indicators. It also provides the logic behind individualized data set which is used in normality testing. Normality is achieved through a Johnson Transformation; with Anderson-Darling test statistic being used to test this claim.

Keywords: mandatory health insurance fund, Kosovo

INTRODUCTION

As Kosovo government pushes towards the initiating of its mandatory health insurance fund (HIF) voted in by the parliament, this paper expands on the [Muhaxheri 2015].

Database for Kosovo’s willingness-to-pay (WTP) database [Dialogue Programe in Health 2012]; which provides comprehensive information on Kosovo’s demographic indicators, and more importantly includes information on Kosovo’s people’s willingness-to-pay for health insurance – a survey conducted through a contingent valuation model (Double Clustered Dichotomous Choice); is used to establish the consumer surplus (CS) which is shown not to follow a normal distribution, and hence in data are normalized using Johnson Transformation method.
SUMMARY OF PREVIOUS FINDINGS

The most appropriate way to enhance the meaning of the previous findings is to present them in its original form. However, first the assumption and definitions are highlighted are presented:

A starting point is the proposed contribution rate ($p$) from Kosovo Law on Health Insurance (LHIF) which stands at 7%; where:

$$p = 7\% \rightarrow 3.5\% (q = \text{employeecontribution}) + 3.5\% (employercontribution)$$

Consumer surplus ($CS$) is defined as the difference between maximum willingness-to-pay and actual amounts covered employed persons are required to pay. This leads to the following:

- $m_i = \text{median of respective household monthly-income bracket}$
- $h_i = \text{respective household size}$
- $W_{\text{MAX}} = \text{maximum (MAX) willingness-to-pay (WTP) amount ($x$) per person per individual households ($i$) per month}$
- $N = \text{total number of covered persons through the mandatory scheme}$
- $n = \text{sample size (used in estimating maximum WTP)}$
- $P_{\text{life}} = \text{total amount of premiums received by private health insurance providers during a specific period, prior to the implementation of HIF.}$

**Consumer Surplus ($CS$)**

$CS$ is estimated by calculating the average difference between maximum WTP ($W_{\text{MAX}}$) (WTP Survey data provided monthly information and therefore a coefficient of 12 is used to annualize $CS$. Surveys with different frequencies can be adjusted accordingly to represent annual figures.) and actual contribution rates defined by law on LHIF, using the following annualized formula:

$$CS = 12N \sum \frac{W_{\text{MAX}} - qm_i}{h_i}$$

(1)

**Surplus-to-Exploitation Ratio - SER**

$SER$ is used as a two-fold indicator [Glendinning 1998]: (i) to measure the ratio of consumer surplus to total amount of premiums received by private health insurance providers during a specific period; and (ii) to signal private insurance providers if further opportunities are available in the market – though this achieved by defining a potential entry threshold. A range of factors affect
Normalizing consumer surplus data for Kosovo’s …

the level at which $SER$ is; which are implicitly embedded in $CS$, as they result
directly from WTP results.

Therefore, $SER$ is developed by dividing $CS$ into $P_{life}$, to obtain:

$$SER = \frac{\sum_{i}^{MAX} x_{WTP}^{i} - q_{m} h_{i}}{12N} \cdot \frac{n}{P_{life}} = \frac{CS}{P_{life}}$$ (2)

Range Values of $SER$ and Their Implications

First obvious statement resulting from (2) is that $SER>0$; which
acknowledges that $CS$ is always positive (provided the law of demand holds) and
assumes that the private health insurance market exists. Resulting ranges and
meaning of $SER$ provide the following information with respect to the level
of ratio:

- $SER<1$ – This level of $SER$ indicates that the current $CS$ has been fully exploited
  by the private health insurers.
- $SER>1$ – Values of $SER$ at this level indicate that the current $CS$ has not been
  fully exploited.

The above information is summarized graphically in Figure 1.

Figure 1. $SER$, Consumer Surplus-to-Exploitation Ratio

Potential Entry Threshold ($PET$) and Market Possibilities

The first step is to start with the meaning of $SER$ and investigate its
implication with respect to market possibilities for the number of providers
of private health insurance. This is achieved by assuming that current providers
make up 100% (or 1 if expressed as a decimal) of the total share of the industry.
Further, let $k =$ the number of private health insurance providers, and assume that:
all providers have an equal \((1/k)\) share of the current market,
have equal access to consumer surplus, and
consumer surplus will be utilised equally.

Next, let define \(PET\) such: \(PET = 1 + 1/k\) (3) where: \(1\) is the total current share and \(1/k\) is average share of potential entries \((ASPE)\), that is:

\[
ASPE = PET - 1
\]

As more market participants join the market, \(k\) increases, implying that the \(ASPE\) decreases and the number of potential new entrants’ increases. And in the limit we get:

\[
\lim_{k \to \infty}(1 + 1/k)
\]

That is, \(PET = 1\) is a natural (asymptotic) lower boundary. So, \(PET\) exists in the interval \((1, SER)\), labelled ‘Market Possibilities Region’. Graphically, we can present this information in Figure 2.

Figure 2. PET Analysis

Source: own elaboration

Therefore, the following conclusion is reached with respect to the potential entry threshold and market possibilities:

i. \(1 < SER \leq PET\) - The market is over saturated, implying that there is no room for additional providers to join the market, or for existing providers to provide additional services at extra costs.

ii. \(SER > PET\) - The market is unsaturated, implying that there is room for additional providers to join the market, or for existing providers to provide additional services at extra costs.

Further analysis of (ii) leads to the following:

• Market possibilities exist only in the region bounded by \(SER\) and \(PET=1\)
Ceteris paribus, and assuming that \( CS \) is completely utilised, total number of potential new entrants (\( N_t \)) is defined as the ratio of \( SER \) with \( ASPE \), that is:

\[
N_t = \frac{SER}{ASPE}
\]  

(4)

\( N_t \) also includes current non-life providers that wish to join the ‘life’ market, and current ‘life’ providers that want to offer extra services at additional costs. Unless the estimation of \( N_t \) results in a whole number, \( N_t \) should always be rounded down.

DATA TRANSFORMATION

Many studies have shown that many processes follow normal distribution. This distribution stands from the rest in its simplicity which makes it amongst the most recognized, understood, and therefore easily utilized in a world where data are becoming more available than ever before [Gilbert 1994].

Normal distribution has its appeal in requiring only to parameters (mean and standard deviation) in order to describe it, and infer conclusions. In essence here is the consumer surplus which is established in the previous section, and is an integral part of all subsequent analysis. However, where equation (1) provides an intuitive aggregate \( CS \), a more individual-data driven approach is required. Therefore, this is extracted from (1) and defined as (5):

\[
CS_i = x_{WPTi}^{MAX} - \frac{qm}{h_i}
\]  

(5)

Since, the individual data contain negative values, an appropriate transformation is the Johnson Transformation. Under this transformation, one of the following three distributions (Minitab is utilized for the transformations, and table representing the formulas has been adopted from its guides.) is optimally selected (which is then used to transform the data into a normal distribution). Johnson family distribution are presented in Table 1. Variable \( CS_i \) is presented by \( x \) in the functions.

<table>
<thead>
<tr>
<th>Johnson Family</th>
<th>Transformation Function</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bounded ( S_B )</td>
<td>( \gamma + \eta \ln[(x - \varepsilon) / (\lambda + \varepsilon - x)] )</td>
<td>( \eta, \lambda &gt; 0; -\infty &lt; \gamma &lt; \infty; -\infty &lt; \varepsilon &lt; \infty; \varepsilon &lt; x &lt; e + \lambda; )</td>
</tr>
<tr>
<td>Lognormal ( S_L )</td>
<td>( \gamma + \eta \ln(x - \varepsilon) )</td>
<td>( \eta &gt; 0; -\infty &lt; \gamma &lt; \infty; -\infty &lt; \varepsilon &lt; \infty; \varepsilon &lt; x; )</td>
</tr>
<tr>
<td>Unbounded ( S_U )</td>
<td>( \gamma + \eta \sinh^{-1}[(x - \varepsilon) / \lambda] )</td>
<td>( \eta, \lambda &gt; 0; -\infty &lt; \gamma &lt; \infty; -\infty &lt; \varepsilon &lt; \infty; -\infty &lt; x &lt; \infty; )</td>
</tr>
<tr>
<td>( \sinh^{-1} )</td>
<td>( \ln[x + \sqrt{1 + x^2}] )</td>
<td></td>
</tr>
</tbody>
</table>

Source: own elaboration
Under the transformation it is found that the best function to transform the data is $S_u$; under which the data are normalized [Carol 2008]. Parts of Minitab output for the data transformation is presented below on Figure 3.

Figure 3. Probability Plot of Original CS Data

![Probability Plot for Original Data](image)

Source: own elaboration

Figure 4. Probability Plot of Transformed CS Data

![Probability Plot for Transformed Data](image)

Source: own preparation

An indicator of whether data follow a normal distribution is if the probability plot of the data follows the indicated linear lines from bottom left to top right (Figure 4, transformed data). Clearly, the original data show a distinct deviation
from normality (Figure 3). Furthermore, Anderson-Darling (AD) test is used to test whether data follow a normal distribution; with the following hypothesis:

\[ H_0: \text{Data follow a normal distribution} \]
\[ H_1: \text{Data do not follow a normal distribution} \]

Under AD test, if p-value is smaller than the level of significance (\( \alpha \), typically 0.05), then \( H_0 \) is rejected; otherwise it is not rejected. Visual inspections of data distributions are reinforced by p-values presented in Figures 3 and 4; where original data have a p-value less than 0.005; whereas transformed data have a p-value of 0.124.

Table 2 includes the actual function and its estimated coefficient that are used in transforming the original data. Table 3 provides comparative descriptive statistics for original and transformed data.

Table 2. Johnson Transformation Analysis

<table>
<thead>
<tr>
<th>P-Value for Best Fit</th>
<th>0.124205</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z for Best Fit</td>
<td>0.91</td>
</tr>
<tr>
<td>Best Transformation Type</td>
<td>SU</td>
</tr>
<tr>
<td>Transformation function equals</td>
<td>-0.585609 + 0.864371 × Arcsinh(( X + 2.50571 ) / 6.81815 )</td>
</tr>
</tbody>
</table>

Source: own calculations

Table 3. CS original versus Transformed Data – Summary Statistics

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>Transformed Data</th>
<th>Original Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.04035</td>
<td>6.35591</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.03165</td>
<td>0.68037</td>
</tr>
<tr>
<td>Median</td>
<td>-0.06272</td>
<td>1.87500</td>
</tr>
<tr>
<td>Mode</td>
<td>-0.28961</td>
<td>-0.12500</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.00846</td>
<td>21.67600</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.04980</td>
<td>3.70301</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.04980</td>
<td>-50.00000</td>
</tr>
<tr>
<td>Maximum</td>
<td>-0.04980</td>
<td>243.87500</td>
</tr>
<tr>
<td>Confidence Level (0.95)</td>
<td>-0.04980</td>
<td>1.33510</td>
</tr>
</tbody>
</table>

Source: own calculations

CONCLUDING REMARKS

Introduction of the mandatory health insurance scheme by the Kosovo government has opened up opportunities for further research into possibilities available to private health insurance providers. An intuitive approach is used to develop a framework for estimated consumer surplus (CS) is provided by
[Muhaxheri 2015]. This forms the basis for extrapolating an individualized data set approach (CS). The data are transformed using Johnson Transformation, and then both sets of data tested for normality using Anderson-Darling statistic. The analysis show that original data do not follow a normal distribution, however normality is achieved upon transformation.

REFERENCES

Muhaxheri E. (2015) Mandatory Health Insurance Fund for Kosovo! What now for private insurers?. A human being in space and time…., TOM III, University of Lodz, Poland.