Robust Non-Parametric Techniques to Estimate the Growth Elasticity of Poverty

Nagwa Mohammed Albehery

Department of Mathematics and Applied Statistics, Faculty of Commerce
Helwan University, Ain Helwan, Cairo, Egypt

This article is distributed under the Creative Commons by-nc-nd Attribution License. Copyright © 2019 Hikari Ltd.

Abstract

The growth elasticity of poverty is one of the most important issues that policymakers and researches are interested in many countries. The growth elasticity of poverty can be defined as the percentage change of poverty level rates respected to the percentage changes in some variables such as income mean and income inequality. The study of this growth is very important to reduce the poverty levels. Many studies are often used the parametric techniques to estimate the change of poverty levels. In this paper, the rank – based estimation will be used as a robust non-parametric technique to estimate the growth elasticity of poverty in Egypt. The real data sets that have been collected by CAPMAS "Central Agency for Public Mobilization and Statistics" in Egypt from 1990/1991 to 2014/2015 will be used to estimate the change of poverty levels in Egypt.

Keywords: Poverty measures, Poverty elasticity, income inequality, Least square estimation, Rank – based estimation, Weighted Wilcoxon estimation

1. Introduction

The first global goal for sustainable development that set in 2015 for the year 2030 by the United Nations is "No poverty". Many programs are developed by decision makers to decrease and alleviate poverty all around the world. The growth of poverty elasticity has become one of the most important issues for poverty reduction. Traditional least squares method as a parametric technique is the most widely used to estimate the linear or nonlinear regression model to measure the relationship between the change of poverty levels that is associated with the changes of income mean and Gini index as a measure of income inequality see i.e.,

(D. Champers and S. Dhongde, 2011) estimate the growth elasticity of poverty using non-parametric panel method. In general, the parametric techniques like the least square method is the best method under regularity assumptions such as normality. But the least square will be not good method in case of violation in some of required assumptions and the nonparametric and robust nonparametric techniques will be better. For non-parametric techniques, there are no strict parametric assumptions and these techniques can handle violation in assumptions but sometimes cannot deal with the outliers in the data. Then the robust non-parametric technique can be used to make the regression estimators are not sensitive to outliers.

Robust as a statistical term can be defined as insensitive against violation of a specified assumption, since 1960 there are robust and resistant methods have been developed (P.J.Huber and E.M. Ronchetti, 2009). Rank-based estimators were developed as a robust non-parametric alternative to traditional least square or likelihood estimators and was introduced by (J. Jureckova, 1971) and (A. Jaeckel, 1972). A Newton step algorithm of the rank-based estimates has been presented by (T.P. Hettmansperger and J.W. Mckean, 1978). A rank-based analysis for linear models has been developed by (M. Hollander and D. A. Wolfe, 1999) and (T. P. Hettmansperger and J.W. Mckean, 2011). The rank-based analysis generalizes Wilcoxon procedures for simple models and it has the same high efficiency as that simple non-parametric procedures. The Weighted Wilcoxon (WW) as a rank-based estimator to estimate linear regression model has been presented by (J. W. Mckean, 2004). (N. M. Albehery, H. A. Auda and E. A. Hassan, 2019) used the Weighted Wilcoxon estimator to estimate the absolute poverty line in Egypt from the year 2011/2012 to the year 2014/2015.

In this paper, in section 2, a brief review of estimating the growth elasticity of poverty in previous studies will be presented. In Section 3, the robust non-parametric technique for using rank-based estimation that used Weighted Wilcoxon estimator will be used to estimate the change of poverty levels. In section 4, the change of poverty as a measure of the growth elasticity of poverty will be estimated using the least square method and the rank-based method to the real data sets collected in Egypt (from CAPMAS) form the year 1990/1991 to the year 2014/2015 and the results between these methods will be compared. The conclusion is presented in section 5.

2. A brief review of estimating the growth elasticity of poverty in previous studies

(T. Besely and R. Kanbur, 1988) presented a study to find the increasing of the income of the poorest people in a population to decrease the poverty levels. The study used the Foster index (J. E. Foster and E. A. Thorbecke, 1984) that takes the following form:
Robust non-parametric techniques

\[ P_\alpha(x, z) = \int_0^z \left(1 - \frac{x}{z}\right)^\alpha f(x) \, dx, \quad \alpha \geq 0 \] (1)

Here: \( z \) is the poverty line that is the border line between poor and non-poor people, \( x \) is the income mean of the poor; \( x \in [0, z) \) and \( z \in (0, \infty) \). \( \alpha = 0 \), the Foster’s index is called the percentage of poor, denoted by \( P_0(x, z) \). When \( \alpha = 1 \), it is called the poverty gap ratio, denoted by \( P_1(x, z) \). When \( \alpha = 2 \), it is called the distribution of income among poor, denoted by \( P_2(x, z) \).

\[ \frac{\partial P_\alpha}{\partial y} = -\frac{\alpha}{z} P_{\alpha-1}, \quad \alpha \geq 0 \] (2)

(N. Kakwani, 1990) calculated the percentage change of poverty measures in each area in a population respect to the income mean and the income inequality as the following:

\[ \frac{\partial P_i}{P_i} = \eta P_i \frac{\partial \mu_i}{\mu_i} + \xi P_i \frac{\partial G_i}{G_i}, \quad i = 1, 2, \ldots, m \] (3)

Here: \( P_i \) is the poverty measure in area number \( i \), \( i = 1, 2, \ldots, m \), \( m \) is the number of areas in a population, \( \mu_i \) is the mean income in area \( i \), \( \eta P_i \) is the elasticity of poverty measure respect to the mean income in area \( i \), \( G_i \) is the gini index in area \( i \), \( \xi P_i \) (R. Heltberg, 2002) analyzed the poverty elasticity of growth and discussed the relationship between the poverty change and the growth. The study defined the elasticity of Foster poverty measure that is defined in (1) as followed:

\[ \eta_\alpha = -\frac{\alpha(P_{\alpha-1}-P_\alpha)}{P_\alpha}, \quad \alpha \neq 0 \] (4)

(R. H. Adams, 2004) estimated the change of poverty using a panel of 60 developing countries by estimating the following regression model:

\[ \Delta P_{it} = \alpha + \beta \Delta \mu_{it} + \gamma \Delta g_{it} + \Delta \epsilon_{it} \] (5)

Here: \( P_{it} \) is the logarithm of the proportion of poor people in a country \( i \) at time \( t \) and \( \Delta P_{it} = P_{it} - P_{i(t-1)} \), \( \mu_{it} \) is the logarithm of mean expenditure of country \( i \) at time \( t \) and \( \Delta \mu_{it} = \mu_{it} - \mu_{i(t-1)} \), \( \beta \) is the growth elasticity of poverty with respect to mean expenditure, \( g_{it} \) is the logarithm of Gini index of country \( i \) at time \( t \) and \( \Delta g_{it} = g_{it} - g_{i(t-1)} \), \( \beta \) is the Gini elasticity of poverty with respect to Gini index, \( \epsilon_{it} \) is a white noise error that includes the errors in poverty measure.
Nagwa Mohammed Albehery

(D. Champers and S. Dhongde, 2011) used the non–parametric panel model as followed:

\[ P_{it} = m(\mu_{it}, g_{it}) + \varepsilon_{it} \]  \hspace{1cm} (6)

Here \( m(\cdot) \) is a smooth function and \( \varepsilon_{it} \) is identically independent random errors. The model in (6) measured the dependence of poverty measure \( P \) on mean income \( \mu \) and income inequality \( g \).

(R. Ram, 2015) estimated the growth elasticity of poverty for the Asian countries over the period 1990-2005. The study used the following exponential regression model:

\[ \ln(Y_t) = \alpha + \beta(t) + U_t \]  \hspace{1cm} (7)

Here \( Y_t \) denotes real GDP "Gross Domestic Product" per capita in year \( t \), \( \beta \) is an approximation to the rate of increase in per capita GDP during the period, \( U \) is the random error term and \( t \) is time from 1990–2005.

(C. Arndt, K. Mahrt and C. Schimanski, 2017) calculated poverty growth elasticity \( PGE \) using the following equation:

\[ PGE = \frac{P_t - P_0}{P_0 \frac{GDP_t - GDP_0}{GDP_0}} \]  \hspace{1cm} (8)

Here \( P_t \) denotes the current poverty rate, \( P_0 \) denotes the initial poverty rate, \( GDP_t \) denotes the current GDP and \( GDP_0 \) denotes the initial GDP.

3. Robust non-parametric technique to estimate the change of poverty level

In this paper, the change of poverty level will be used as a measure of the growth elasticity of poverty. The following linear regression model will be used to estimate this change:

\[ \Delta P_\alpha = \beta_0 + \beta_1 \Delta \mu + \beta_2 \Delta G + \beta_3 \Delta P_{\alpha-1} + \Delta \varepsilon, \quad \alpha \geq 1 \]  \hspace{1cm} (9)

Here: \( P_\alpha \) is the foster poverty index defined in (1) and \( \Delta P_\alpha = P_\alpha - P_{\alpha(t-1)} \). \( \mu \) is the mean income and \( \Delta \mu = \mu_t - \mu_{(t-1)} \). \( G \) is the Gini index as a measure of income inequality and \( \Delta G = G_t - G_{(t-1)} \). \( \varepsilon \) is identically independent random errors. In this model, we assume that the change of foster poverty index at \( \alpha \) is dependent variable and the change of income mean, the change of Gini index and the change of foster poverty index at \( (\alpha-1) \) are explanatory variables.
Robust non-parametric techniques

(J. W. McKean, 2004) presented the robust analysis that offers a methodology for linear models similar to least square "LS" and it is not sensitive to outliers. Least square estimators based on minimize the Euclidean norm, while rank-based estimators based on minimize pseudo-norm that similar to Euclidean norm. Let a set of data where $X_1, ..., X_p$, $p$ be the explanatory variables, and $Y$ be a response variable and the regression of $Y$ on $X$ at $n$ points. Let $Y$ be the $n \times 1$ vector of responses and let $X$ be the $n \times p$ matrix. The linear regression model will be:

$$Y = 1\alpha + X_c\beta + \varepsilon$$  \hspace{1cm} (10)

where $\beta$ is $p \times 1$ vector of regression coefficients, $\alpha$ is an intercept, $X_c$ is the centered design matrix with $x_{ij} = x_{ij} - \bar{x}_j$, $\bar{x}_j = \sum_{i=1}^{n} x_{ij}/n$ and $\varepsilon$ is $n \times 1$ vector of random independent errors. The least square estimator for the regression coefficients is:

$$\hat{\beta}_{LS} = \text{Arg min}_{\beta} \| Y - X_c\beta \|^2$$  \hspace{1cm} (11)

where $\| Y - X_c\beta \|^2$ is the squared Euclidean norm and under regular assumptions. For the large sample, the distribution of the least square estimator will be:

$$\begin{pmatrix} \hat{\alpha}_{LS} \\ \hat{\beta}_{LS} \end{pmatrix} \approx N \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \begin{bmatrix} \sigma^2/n & 0 \\ 0 & \sigma^2(X_cX_c)^{-1} \end{bmatrix}$$

were, $\sigma^2$ is the variance of errors for least square estimator.

But the rank-based estimator for the regression coefficients will be as followed:

$$\hat{\beta}_{WW} = \text{Arg min}_{\beta \in \mathbb{R}^n} \| Y - X_c\beta \|_w$$  \hspace{1cm} (12)

where $\| Y - X_c\beta \|_w$ is pseudo-norm the norm of rank-based estimator. Let $V = Y - X_c\beta$, $V \in \mathbb{R}^n$, and the pseudo-norm $\| V \|_w$ will be

$$\| V \|_w = \sum_{i=1}^{n} a(R(v_i)) v_i = \sum_{i=1}^{n} a(i) v(i),$$  \hspace{1cm} (13)

$R(v_i)$ is the rank of $v_i$, a(i) is a function of $\varphi$ as followed:

$$a(i) = \varphi(i / (n + 1))$$  \hspace{1cm} (14)

Where $\varphi(u)$ is the score function and it is non-decreasing function on $(0, 1)$. (J. W. McKean, 2004) used the linear Wilcoxon score function and the rank-based estimator becomes Weighted Wilcoxon estimator (WW) with score function $\varphi(u)$ as the following:

$$\varphi(u) = \sqrt{12}(u - 0.5)$$  \hspace{1cm} (15)

And the estimate of $\alpha$ using the median of residuals will be:

$$\hat{\alpha}_W = \text{med}_{1\leq i \leq n} \{ Y_i - x'_c \hat{\beta}_W \}$$  \hspace{1cm} (16)

For the large sample, the distribution of the Wilcoxon estimator will be:
\[
\left( \hat{\alpha}_W \right. \left. \hat{\beta}_W \right) \approx N \left( \left( \begin{array}{c} \alpha \\ \beta \\ \frac{\tau^2}{n} \\ 0 \\ \tau^2(XcXc)^{-1} \end{array} \right) \right)
\]

where \( \tau^2 \) and \( \tau_s^2 \) are the scale parameters are estimated by (T. P. Hettmansperger and J. W. McKeen, 1998), \( \tau^2 \) is the variance of error for Wilcoxon estimation where,

\[
\tau = \left( \sqrt{12} \int f^2(t)dt \right)^{-1}, \tau_s = (2f(0))^{-1}
\]

Where \( f(t) \) is the probability distribution function of the errors.

To compare between the coefficients that estimated using different methods, the asymptotic relative efficiencies (ARE) (J.W. McKeen, 2004) can be used by using their variances. ARE between \( \hat{\beta}_{LS} \) that is the estimation of least and \( \hat{\beta}_{WW} \) that is the estimation of Weighted Wilcoxon is:

\[
\text{ARE} \left( \hat{\beta}_{LS}, \hat{\beta}_{WW} \right) = \frac{\text{var} (\hat{\beta}_{LS})}{\text{var} (\hat{\beta}_{WW})} = \frac{\sigma^2}{\tau^2}
\]

4. Applied example

In this paper, data for Urban and Rural Egypt for 1990/1991, 1995/1996, 1999/2000, 2003/2004, 2007/2008, 2010/2011, 2012/2013 and 2014/2015 that are collected from the Household income, expenditure and consumption survey will be used. This survey was presented by "Central Agency for Public Mobilization and Statistics (CAPMAS)" the government agency to collect data in Egypt. The relative poverty line "z" is used as the border line between poor and non-poor people in Egypt. This line is defined as a percentage of median of expenditure. (N. M. Albehery and T. Wang, 2011) estimated the upper limits of the relative poverty line \( Z = 2/3 \) of the median of annual household expenditure. This upper limit will be used as the poverty line in this paper.

The Foster poverty measure that is defined in (1) at \( \alpha = 1 \); \( P_1(x, z) \) will be used as the measure of poverty level in Egypt. And the change of this measure will be the dependent variable in the linear regression model defined in (9). The change of mean expenditure, the change of percentage of poor people and the change of Gini index will be the independent variables. Where the Gini index estimated using the following form:

\[
G = 1 - \sum_{i=1}^{n} (h_i - h_{i-1})(E_i + E_{i-1}),
\]

Where: \( h \) is the relative cumulative of total number of households in interval \( i \), \( E \) is the relative cumulative of total expenditure of households in interval \( i \), \( i = 1, 2, \ldots, n \) and \( n \) is the number of intervals. The following table contains the measures that are estimated from Urban and Rural Egypt data.
Table (1): Estimated poverty line \( z \), poverty gap \( P_1 \), percentage of poor \( P_0 \), mean of expenditure \( \mu \) and the Gini index \( G \) for urban and rural Egypt from 1990/1991 to 2014/2015.

<table>
<thead>
<tr>
<th>Year</th>
<th>Urban Egypt</th>
<th>Rural Egypt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( Z )</td>
<td>( Z )</td>
</tr>
<tr>
<td>1990/1991</td>
<td>3919.9</td>
<td>2775.7</td>
</tr>
<tr>
<td>1995/1996</td>
<td>4319.3</td>
<td>3442.9</td>
</tr>
<tr>
<td>1999/2000</td>
<td>6203.2</td>
<td>4550.5</td>
</tr>
<tr>
<td>2003/2004</td>
<td>7271.5</td>
<td>5609.9</td>
</tr>
<tr>
<td>2007/2008</td>
<td>11275.1</td>
<td>9070.0</td>
</tr>
<tr>
<td>2010/2011</td>
<td>14319.9</td>
<td>11278.7</td>
</tr>
<tr>
<td>2012/2013</td>
<td>16570.5</td>
<td>13704.5</td>
</tr>
<tr>
<td>2014/2015</td>
<td>22570.2</td>
<td>14054.5</td>
</tr>
</tbody>
</table>

Since the measures in the above table are calculated using the data available for the years 1990/1991, 1995/1996, 1999/2000, 2003/2004, 2007/2008, 2010/2011, 2012/2013 and 2014/2015, then the measures between the given years are estimated using the interpolation method to calculate the change of these measures between the years from 1990/1991 to 2014/2015. The linear model that is defined in (9) is estimated using least square regression method and weighted Wilcoxon method for Urban and Rural Egypt and the results are included in the following table:

Table (2): The estimated coefficients and standard error of model (9) by using Least Square and Weighted Wilcoxon estimation for Urban and Rural Egypt.

<table>
<thead>
<tr>
<th></th>
<th>Least square estimation</th>
<th>Weighted Wilcoxon estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\beta}_{LS} )</td>
<td>( \hat{\beta}_{WW} )</td>
</tr>
<tr>
<td>( \hat{\beta}_0 )</td>
<td>7.913e^{-3}</td>
<td>6.629e^{-3}</td>
</tr>
<tr>
<td>( \hat{\beta}_1 )</td>
<td>-6.063e^{-6}</td>
<td>-6.791e^{-6}</td>
</tr>
<tr>
<td>( \hat{\beta}_2 )</td>
<td>6.205e^{-2}</td>
<td>1.185e^{-1}</td>
</tr>
<tr>
<td>( \hat{\beta}_3 )</td>
<td>-2.327e^{-1}</td>
<td>-5.837e^{-1}</td>
</tr>
<tr>
<td>( \hat{\sigma} )</td>
<td>0.01226</td>
<td>0.009307</td>
</tr>
</tbody>
</table>

The asymptotic relative efficiencies (ARE) between the estimation of least square \( \hat{\beta}_{LS} \) and the estimation of Weighted Wilcoxon \( \hat{\beta}_{WW} \) is estimated as followed:
For Urban Egypt: \( \text{ARE}(\hat{\beta}_{LS}, \hat{\beta}_{WW}) = \frac{\sigma^2}{\tau^2} = \frac{(0.01226)^2}{(0.00730)^2} = 2.76 \)

For Rural Egypt: \( \text{ARE}(\hat{\beta}_{LS}, \hat{\beta}_{WW}) = \frac{\sigma^2}{\tau^2} = \frac{(0.00930)^2}{(0.00459)^2} = 4.11 \)

7. Conclusion

The growth elasticity of poverty is very important to develop some plans and programs to reduce poverty levels in a population. So, using a good technique to estimate the rate change of poverty is required. In this paper, the change of poverty level will be used as a measure of the growth elasticity of poverty. The linear regression model of the change of poverty gap measure on the change of expenditure mean, the change of Gini index and the change of percentage of poor people is defined. The least square estimation as a parametric technique and the Weighted Wilcoxon estimation as a robust non-parametric technique are used to estimate the coefficients of the linear model for Urban and Rural Egypt using real data sets from years 1990/1991 to 2014/2015. The asymptotic relative efficiencies (ARE) is used to compare between these different estimation methods and we conclude that the Weighted Wilcoxon estimation is better than the least square estimation 2.76 times for Urban Egypt and 4.11 times for Rural Egypt.

References


Received: March 21, 2019; Published: April 14, 2019