SMALL AREA ESTIMATION OF POVERTY IN THE PHILIPPINES

Relatore: Gianni Betti
Correlatrice: Laura Neri

Tesi di Laurea:
Kenneth R. Mauro

A.A. 2015-2016

Contents

TABLES

FIGURES

ACKNOWLEDGEMENT

This research is supported together with Professor Betti Gianni, who I greatly thank for the patient work, for suggesting some information through the computation of EBLUP estimates and to this analysis concerning the poverty in the Philippines. I would like to show my gratitude to my family, particularly to my parents and my sister, Mary Grace. I am thankful with all my heart to Mr. Vasco Caini for his financial support, giving me the opportunity to realize my dreams. My sincere thanks go also to my dear friends who have been supporting me for these years: Vittoria, Lisa, Yuliya, Anna, Dania, MJ, Enza, Giulia, Alvaro and Giuseppe.
I thank my colleagues Jessica, Lucia, Michele and Federico. In addition, I am thankful to Mr. Scagliola, to Ms. Mikiko, to Ciarrocchi and Noli families in inspiring me on how they put dedication in their daily lives and giving me a lot of encouragement. Finally, I thank Prof. Maryann for her efforts on checking the second chapter of this essay and giving me some feedbacks on my English language mistakes.

Kenneth Rayel Mauro
Born at Ilaor Sur, Oas Albay, Philippines, the December 27 1990.
kneth@libero.it

ABBREVIATIONS

Asean	Association of Southeast Asian Nations
APIS	Annual Poverty Indicator Survey
BLUP	Best Linear Unbiased predictor
CPH	Census of Population and Housing
EAs	Enumeration Areas
EB, EBP	Empirical Bayes or Empirical Best Predictor
EBLUP	Empirical Best Linear Unbiased predictor
FGT	Foster-Greer-Thorbecke or severity poverty
FIES	Family Income and Expenditure Survey
GIS	Geographic Information Systems
HCR	Head count ratio
HDI	Human Development Index
LABs	Least accessible barangays
ML	Maximum Likelihood
MS	Master Sample
MSE	Mean Squared Error
NCR	National Capital region
NSO	National Statistical Office
POPs	Barangays with peace and order problem
PPP	Purchasing Power Parity
PSUs	Primary Selection Units
REML	Residual Maximum Likelihood
SAE	Small area estimation
USD	U.S. dollars

ABSTRACT

The objective of this essay is to utilize a model of small area estimation in the Philippines. Because of direct estimates are subject to be unreliable, some of the model in practice have been proposed, in particular the EBLUP model. The variables utilized are the following: head count ratio, income gap, poverty gap, severity of poverty, average annual income and expenditure, and food threshold. EBLUP estimates will be computed for each of these indicators with their relative MSE. After the analysis has been taken place, the EBLUP estimates are very close to the direct one. The reason could rely on larger sample size utilized on each region, even though it is not calculated in this analysis. Gamma coefficients are smaller than unity so that the EBLUP estimates is showing a good model. However, the MSE gain is very small, only some regions in particular can be excluded but their interpretation change in which indicator is utilized. Only the Food threshold had some adjustment using the EBLUP model, meaning that the covariates used on the model have a strong relationship to this indicator. To obtain much gain in the MSE, the EBLUP model works much better if the sub-domains are extended to provincial and municipal level. However, it could not be implemented in this analysis as some information are not available.

Key words: Poverty, Small Area Estimation, EBLUP model, head count ratio, income gap, poverty gap, severity of poverty, average annual income, average annual expenditure, food threshold.

1. INTRODUCTION

The objective of this essay is to utilize a model of small area estimation in order to estimate the poverty indicators. Philippines is the location of study. Because of direct estimates are subject to be unreliable, some of the model in practice have been proposed. The poverty indicators are the following: head count ratio, income gap, poverty gap and severity of poverty. Other estimations have been also used: average annual income and expenditure, and food threshold. Each of these indicators will be computed a new estimates with their relative MSE. In the end, the results will be the comparison between the direct estimates and that of EBLUP.

Poverty is one of the most important issues and it is often discussed how to eradicate it. In order to do that, it is necessary to give a definition of poverty. In the literature, there are so many concepts that can be attributed to it. Therefore, the first chapter concerns its general definition to identify who are "poor" to have better results on policy-decision. Because of that, we need to quantify it firstly and take into consideration some factors, which are the unit of measure, the equivalence of scale, the poverty line and the poverty indicators. We begin to the unit of analysis, which will be counted on the sample. It can be a singular person, family or household. Then it is necessary to define the equivalence of scale and the poverty line. Each country has its own definition so that it can be difficult to make comparison among them. The equivalence of scale is introduced by Engel, namely also as food ratio method. Another kind of this subject is the OECD scale, describing a certain weight of real income or consumption for each member of the family The poverty line is the quantification of income or expenditure to verify if a certain family can sustain its main necessity. It is divided on absolute, relative or subjective poverty. The absolute poverty is based on basic food, while the relative one relies on income or consumption and the subjective one depends on the minimum income question. Their advantages and downside will be also illustrated. We go ahead by defining the axioms, which consist of desirable properties of the poverty indexes: monotonicity, transferability, subgroup consistency, symmetry and so on. They will be described better further. After these steps that have done, it will be discussed the poverty indicators: headcount ratio, income gap, poverty gap, severity poverty and other indexes. Each of them has a different meaning and we can have a more detailed poverty information by putting them together. In addition, the unidimensional and multidimensional poverty will be presented, and which problems appear by using one of this definition. The unidimensional utilizes only monetary variables, income or expenditure. However, this latter cannot give an exhaustive explanation of poverty. Therefore, the multidimensional poverty is relevant because it is based on social and non-economic variables such as the general health, education and infant mortality. Finally, the inequality measures are introduced. They are computed using different methods. It can be income percentiles and Gini index as shown in the last paragraph of this section.

In the next chapter, the small area estimation (SAE) will be discussed and its relevance on measuring poverty because a national indicator may summarize too restrictive. Therefore, it is used the sub domain level in order to estimates the spatial heterogeneity among regions, provincials or municipalities. We can find different methods to estimate SAE, either by the direct or the indirect one. However, the direct estimates are not efficient as the indirect one. It is possible to find small domains in which the sample size is very little so that the direct estimates has a high standard error. We start by describing the dataset sources: CENSUS, surveys and auxiliary data.

Then, it is introduced the general model to construct SAE: design-based model, model linking and linear mixed model. The last model will be described extensively because it allows us to derive EBLUP. The following paragraph will be described how to derive the Empirical Best Linear Unbiased Predictor (EBLUP). From the linear mixed model, Henderson has developed a BLUP estimator, in which the BLUP estimator of θ_i is obtained. Then, the nested-error model will be al so presented. Moreover, MSE will be estimated and the procedure of the parametric bootstrap will be shown. The second model that is considered here is the Empirical Bayes (EB) or Empiric al Best Predictor (EBP). Its MSE is also estimated. The last model is the poverty mapping using data from a survey, which is combined together to the CENSUS. Once the erratic term and the household level effects are calculated with the estimation of household specific variance, we can go on to the simulated values of the consumption expenditure for each household. Then, the loss in precision using census aggregate is presented. Finally, the measurement of poverty is computed after the simulation.

The third chapter introduced a general situation of poverty in the Philippines.

The forth chapter concerns about the poverty analysis in the Philippines. The direct estimates are headcount ratio, income gap, poverty gap, severity of poverty. Food threshold, average annual income and expenditure are added to the analysis. Then, we collect covariates from the macroeconomic indicators. The model utilized is the EBLUP to obtain more efficient estimates. Using R program, gamma coefficient will be computed. It allows us to find the EBLUP estimates and to compare it to the direct one by calculating their relationship. The ratio of MSE will show us if there is any gain of efficiency and how many percentage it is.

2. POVERTY

2.1. GENERAL DEFINITION OF POVERTY

Poverty can be defined as somebody who is homeless, cannot afford basic food and clothing. Similarly, if a person who is retired and cannot make ends meet or does not have money to buy goods or services that are commonly necessary to sustainability. However, which is exactly the correct concept?
In practice, many definitions can be attributed to poverty. It is important to choose a specific one in order to compute the level of poverty and to know how to address the policy-maker. Different treatments of these arguments bring various results, so that it may change the poverty effects on the population considered. Therefore, to divert to these problems, it is important to analyze some factors, which are listed below:

1. Unit of analysis;
2. Choice of equivalence scales;
3. Poverty lines;
4. Poverty indicators;
5. Unidimensional and multidimensional approach.

It is important to add the advantages and critical points of view of each definition. Every country has its own concept, determinants and methodology to compute poverty, so that there are some difficulties regarding the comparison between the results. Another element that causes this relative problem, is the cost of living of each nation. The last one can be partly resolved using the *Purchasing Power Parity (PPP)* and transform it into the current U.S. dollars. Besides, more concern must be given to who lives on less than USD 2 per day and to a smaller account, which implies that their expenses is completely for survival needs. In the end, the main purpose is to show the trend of that relative country to address the right policy to deprive poverty.
Another issue to take in consideration is the inequality. If we analyzed all the changes of absolute poverty, we can say that the percentage has decreased, even if it remains a considerable number. The disparities between countries is getting larger due to many factors over time, which brings us to our situation nowadays. These factors are mainly political, social, cultural and economic causes. Inequalities measure are determined by Lorenz Curve, Gini equation and Income percentile.

2.2. UNIT OF ANALYSIS

It is common to utilize the household units, which is a group of people who are living in the same dwelling and share their expenses. For instance, an income family distribution can indicate the well-being of the population considered. On the other hand, individual units are used in other types of research and it is a singular person who has specific requirements to be included in the sampling. It expresses the generic character of a nation, such as population’s average number and gender. Mostly, the family size is more convenient because the income is extended to all the members. Then, this latter will be supplemented using equivalence scale.

2.3. CHOICE OF EQUIVALENCE SCALE

The equivalence scale is relevant because every family member has a different share in the overall expenses, given a certain income distribution. Adults have a higher quota than children do, so by defining it, then it is possible to see each average of their contribution. It changes in every country and until now, it is the subject of numerous studies to be more appropriate and apply then to their relative nation. Its definition to transfer into computation may be quite difficult because of the differences of characteristics concerning the number of family's composition, their ages and preferences. Moreover, many surveys are using household units, so that a vector of coefficients reflecting the income's individuals is necessary. Hence, it is possible to obtain the equivalised disposable income by the ratio of the family's income and the coefficients determined before.
The first studies were based on the food ratio method, in which Engel argued that as the family size increases, so the expenditure for food, in order to maintain the same level of well-being. He also said that the function of the expenses would tend to increase less proportionally in case of higher income. Therefore, he defined the equivalent scale as the ratio between two household incomes, given the same amount of food expenditure.
Engel's method of iso-prop, in which he considers the relevance of clothing and house's utilities as well as foods to define the utility level, is also relevant.
Another description is made by Rothbarth, who considers adult goods as a factor of well-being. In this case, the equalized income is computed by comparing one's household expenses to the general one.
Other types of concepts can be based on c*onsumer theory.*

There is not only one measure, but household's size and other factors relative to the specific population must be taken into consideration.
For instance, OECD scale-modified is applied in some countries:

	OECD scale	OECD scale-modified
first adult	1	1
additional adults	0.7	0.5
children	0.5	0.3

2.4. POVERTY LINE

The poverty line is a boundary in which the value of the poverty's level is set. Given the principal elements to identify who are poor, e.g. the variables and the units of analysis, it is then possible to determinate this quantity. After measuring it, a household or a family is said to be poor if they do not across this boundary.
In practice, it is relative to a certain income's level or expense function.
Mostly, it is possible to aggregate it into three groups, which are widely used:

1. Absolute poverty
2. Relative poverty

3. Subjective poverty

Absolute poverty considers the importance of the basic needs and sustenance of each family. Therefore, it is necessary to determine which goods and services are highly important, which will be part of the basket of basic needs. At this point, there will be a list by Province or by Region with each value and updated every year, considering the changes in expense cost. This relies on many factors like the region they are living, economic and social group, and their different group characteristics. In the end, the poverty line is set to whom are not capable to provide the minimum value. This method is called Cost of Basic Needs, which is computed as the sum of survival food needs, basic non-food needs and basic food needs for economic and social activity.
An alternative method is the Food-Energy Intake, which aims to find the monetary value of a certain predetermined food bundle for the maintenance human body energy and to set the poverty line. Factors such as individual character, time relative to the observation and food energy are the reason of different result to this latter. It is not necessary to adjust the price index.
To sum up, the definition of this part can be distinguished as a concept of sustenance or as a minimum acceptable level of life. The former influences life itself due to lack of resources, whereas the latter is somebody who cannot manage to pay for goods and services to have a life's level acceptable in the area considered.
Hence, absolute poverty is an economic condition of incapability to purchase basic needs, irrespectively of the mean standard life of the overall population considered.

Relative poverty, however, is based on income or consumption. By doing so, poverty is measured through the comparison of the overall of these variables considering a certain group of population. Hence, here the relevance is on the mean living standard and every country has a fixed percentage of media or median of the income/consumption in order to determine the poverty line. It relies on choosing the variable to estimate and to construct its distribution every year. It is important to take note that this involving only that specific country as the others have different distribution of income/expenditure. Problems may appear about the comparability between time and space. While the latter has been argued before, the former is due to the real representation of income distribution. For instance, it is possible to have a period of positive economic cycle, where everybody benefits from its results. In these cases, the consequence on poverty does not changed if these benefits have been equally distributed among the entire household. This effect is also attributed during a recession period, where the family's income become lesser.
Thus, the relative poverty line is correlated to the standards of predominant life of that relative group and it is dependent to that social, economic and cultural characteristic. Because of the problems showed above and the effects on the changes of distribution, it is preferable to compare it with the definition of absolute poverty.

The **subjective poverty** is based on personal family's judgment of their actual situation. It is based on minimum income question (MIQ), whether they have enough income to meet the basic needs. By these answers, it is possible to set the subjective poverty line, where if a family is satisfied with their condition of income level, then they will set above the poverty line; otherwise, they will set below.

2.5. POVERTY INDICATORS

Before going ahead, it is well to define the axioms that provide good properties to the indexes and choose among them which are preferable.

Monotonicity Axiom – If there is a reduction in the income for those below the poverty line, the index measurement should go up.

Transfer Axiom – If there is some transfer of income from the individual considered poor to who is rich, the poverty measure should increase.

Subgroup Consistency Axiom – the overall level of poverty must decrease if the poverty of some within a subgroup of a population goes down whereas the between group remain unchanged.

Symmetry Axiom – the index remains unchanged even if the poor's income is inverted to the rich one.

Mean Independence Axiom – it is irrelevant to the index if the overall income is multiplied at the same constant.

Population Size Independence Axiom – if the population has changed, the poverty measurement will be at the same level as initially.

Decomposability Axiom – the index can be decomposed in different population groups.

After describing the axioms, the main indexes are presented in order to quantify poverty and how severe it is.

Headcount Ratio

The index most used is the headcount ratio for its simple comprehension and because of the way it is computed. It represents the percentage of family or households who are below the poverty line.

$$H = \frac{q}{N}$$

It is the ratio between the households who do not cross the poverty line (q) and the total population (N).
However, there are weaknesses to be considered because some desirable properties are not satisfied: monotonicity, transfer and subgroup consistency axioms. Another limit is the dichotomization between poor and non-poor, without considering how severe the poverty is. Consequently, it may cause some problems because every poor has the same weight, even to those who are closer to the poverty line. Lastly, it must be noted that this index is mostly expressed in household units. On the other hand, using individual units has a better result, without considering whether if it has a smaller or a larger household.

Income Gap

It indicates how the percentage of poor's income is below to the poverty line. It satisfies the monotonicity and the transfer axiom, where the last one must be under constrained to those who change their status by becoming non-poor.

$$I = \frac{1}{q}\sum_{i=1}^{q}\left(\frac{z - y_i}{z}\right)$$

In this formulae, y_i is the consumption or income of the i^{th} family and z is the poverty line. However, it does not reflect the real results in case of somebody that jump the limit defined. This is because it is an indicator which considers only the mean income of poor. Therefore, it has an opposite result due to that relative person that is not counted anymore.

Poverty Gap

To deviate to this last problem, we must take in consideration the index with respect to the total population. This indicator is the poverty gap, expressed as a proportion of the poverty line.

$$P = \frac{1}{N}\sum_{i=1}^{q}\left(\frac{z - y_i}{z}\right)$$

It satisfies the monotonicity, but not the transfer axiom.

Severity Poverty (Foster-Greer-Thorbecke)

It expresses the mean distance from the poverty line and it is calculated to the total population. It also represents the mean value to transfer to the poor people in order to have the same income as the limit defined. It satisfies the main properties of poverty index.

$$FGT = \frac{1}{N}\sum_{i=1}^{q}\left(\frac{z - y_i}{z}\right)^{\alpha}$$

When α = 0, then it is the headcount ratio; if α = 1, then it represents the poverty gap ratio. Greater α means a better understanding on how severe the poverty is, giving a higher value to the poorest people. In this case, α is a measure of adversity to the poverty. It is called Poverty Severity when α = 2. It also has an advantage in that these three results can be compared to each other and give a wide knowledge of the poor's average and poverty gap together. For instance, it is possible to analyze if some groups are numerously poor but not deeply so and vice versa.

Sen Index

Sen has proposed a combination among indexes studied above and the Gini coefficient.

$$S = HG_q + P(1 - G_q)$$

Every factor on this formula varies between 0 and 1. Therefore, if each of them is equal to 0, then the whole household is not considered poor. Otherwise, if nobody of them do not across the poverty line, then Sen index is equal to 1. Another result is the case of income distributed equally ($G_q = 0$), so that S = P.

Watt's Measure

An additional preferable index is Watt's measure. Although it satisfies the monotonicity and transfer axioms, it is not widely used on the practical computation.

$$W = \frac{1}{N}\sum_{i=1}^{q}[\ln(z) - \ln(y_i)]$$

Using measure, in which it is then divided to growth rate (g), it is possible to obtain how many years is needed to exit from the poor area.

$$t \approx \frac{\ln(z) - \ln(y_i)}{g} = \frac{W}{g}$$

Critical points of view

These measurements are very important to see the trend of poverty of each country. However, as mentioned before, the poverty cannot be dichotomized between poor and non-poor.
In addition, the determinants of income or expenditure shows partly the real situation of poverty. It must be supplemented with other variables such as health, life expectancy and education. This analysis is about the multidimensional approach and it will be discussed at the next paragraph.

2.6. UNIDIMENSIONAL AND MULTIDIMENSIONAL APPROACH

The unidimensional approach is based on monetary variables: income is frequently used as well as expenditure. Income is the total revenue from different types of jobs, land or capital businesses. In many cases, the real amount is difficult to compute because of the taxes on it. The disposable income, on the other hand, is the means by which a family satisfy their needs.
Expenditure is a measure of poverty that can be considered more reliable than income, since it represents permanent income that reflects their life standard in long terms. However, it must be pointed out that expenditure is relative to a particular social life, basic needs and other geographical characters.
Therefore, by using one or more of these variables, it is possible to identify poverty, given the definition of absolute or relative showed before. It is called "poor" who cannot across the value of the poverty line and has not enough resources to the basic expenses, relatively to the overall population.

In practice, poverty can be measured using other definitions, which consist of the multidimensional and the subjective one. The former relies on many factors, such as longevity, health, occupation, education, housing, accessibility on water and electricity, and many main types of amenities to be considered well-being or not. This definition came after Sen's theories of capabilities and functionings[1]. Whereas the subjective one is based on personal judgment of respondents and different variables that a researcher include on the questionnaire. This latter can follow either the statistical method, so that the results are efficient and reliable, or not, because of problems due to time, data collection or space.

It can be proved that the unidimensional approach can be useful. For instance, an analysis in Luxemburg is taken, where the poverty line is set at 19,400 € and the overall population on the sampling is 13,423. It is shown how poverty is getting worse without the social transfers and it is quite close to the half percentage of the population that will be deemed "poor" if it is not counted any transfer payments and retirement benefits on the disposable income.

		Headcount Ratio (HCR)
case 1	disposable income	14.50%
case 2	disposable income, without transfer payments	29.06%
case 3	disposable income, without transfer payments an d retirement benefits	45.02%

Source: Essay on poverty's analysis in Luxembourg based on monetary variable (income), a group elaboration with Jessica Solari and Lucia Cafaro.

Mostly, in high-income countries, these services are guaranteed and its data are available, whereas in other nations they there may be fewer of these services or none of them. If this method is then applied to a middle or low-income country, it does not show the social effects effectively and HCR has a unique meaning. Hence, this measure and this approach are not enough. To sum up, there are two problems concerning the unidimensional approach: the use of monetary variables and how to determine the poverty line.
To resolve these problems, a multidimensional approach is used, which is based on many determinants for a better and exhaustive explanation of poverty. It describes a wide concept of poverty because it uses social and non-economic monetary factors. And finally, there is no dichotomization between poor and non-poor.
Because of the numerous variables, some of which have been mentioned before, it is necessary to verify if they are measurable. Moreover, it introduces the problem of a complex analysis as the number of determinants are getting larger. Therefore, it is possible to attribute a relative weight to each of them or to use a cluster analysis. Nonetheless the latter cuts off some of the variables. We can tolerate some loss of information because some of it will be overlapped. Eigenvalue, cumulative portion or scree plot are the methods to use in order to choose which factors will be analyzed. For example, the cumulative ratio will express the share of total variability that is explained by the principal components. In the end, it is possible to interpret the output of the Factor Model, on which the main variables that describe it better are associated.

"Development is the desired change from a life with many sufferings and few choices to a life with satisfied basic needs and many choices, made available through the sustainable use of natural resources."
In some aspects, it is similar to the definition of poverty, except that its development is a concept

[1] Commodities are the means to satisfy the essential needs. He argued that the main components is not itself the monetary variable, but how to obtain our life's purposes (functionings), by the commodities and the capabilities.

more huge, in which include the poverty and how to eradicate it.
Nowadays, Sen's method is applied to Human Development Index (HDI), which indicates that poverty is measured by longevity, education and standard of living. Each of them has an equal weight to the total indicator. However, it must be noted that a higher income is not always associated with a better social and developed life. In fact, some of the determinants may be inefficient because of the government system or lack of investment, despite their increasing GNP. Moreover, the different characteristics among countries must be consider, so that it is possible to add further adjustment indicators.
An additional study is attributed to John Rawls, who defines well-being as based on "justice". This latter is expressed by liberty, equality and reward for services contributions.
Other international measurements can be found in Multidimensional Poverty Index, Gender Development Index and Inequality-adjusted Human Development Index (IHDI), apart from HDI.
Some authors, like Rosling, argued that the division between two different groups of poor and non-poor is not possible anymore. A new taxonomy is provided also by the World Bank, which divides them according to the gross national income: family in a high-income country (from 9,076 USD), middle income-country (2,936 – 9,075 USD), low-income country (lower to 736 USD) and collapsed country.
The development indexes do not to substitute the poverty indicators based on monetary variable that have been proposed above, but to give a wider interpretation of poverty as a whole and to address the right policies for that relative determinant which is deprived towards a targeted group.

2.7. INEQUALITY MEASURE

To go deeper into poverty, the inequality measure should also be taken into account. It is the disparities of resources on the entire population.
To represents this indicators, the income's variable as well as the expenditure is used. The income percentile and the Gini coefficient is most widely used.
Income percentile divides the population into fifth (quintiles) or into tenth (deciles) ranks from the poorest to the richest income, in which it is shown the proportion or the cumulative at each level. As a result, it is possible to see how much the income is distributed among the whole population.
Another measurement is the Gini Index (G). Graphically, it is derived from Lorenz curve, in which the income inequalities are presented and how much larger it is. Firstly, it is necessary to draw a graph, where on the abscissa axis there will be the cumulative income's percentage, while on the other there will be the cumulative population's percentage. Then, the curve is traced, which indicates the association of the cumulative population who holds correspondingly the cumulative income. Special cases are the perfect equality income distribution and the maximum concentration of the total income to a singular person. In the former case, the curve is on the diagonal line, whereas on the latter there is a flat line as the population has no income up until the highest point, denoting the total income distribution for that person. However, real situation is an intermediate between them. Gini Index provides how much divergence there is between the perfect equality and to the Lorenz curve. If G is equal to zero, then there is perfect equality, while if G is equal to unity, it represents the second special case mentioned before. It satisfies the mean independence, population size independence and symmetry properties, but not decomposability.

In conclusion, a vast definition of poverty has been discussed. Many indicators measure different aspects of poverty. One of the problems is the comparison in space, so that the relevance of the absolute definition is then applied, although the local descriptive analysis and the differences among Regions and Provinces must be also added. Inequalities measure describe the income's distributions to the whole population. Other social indicators show human development as a wider vision of poverty. Hence, it is possible to analyze poverty by using this information and to address equal policies among the population.

3. SMALL AREA ESTIMATION

3.1. DEFINITION OF SMALL AREA ESTIMATION (SAE)

After a general description of poverty and inequality, the next topic is the small area estimation. It is necessary to take into account the spatial heterogeneity because a national indicator may summarize too restrictive. Therefore, a disaggregation of different levels within the country is needed, which is also named as "domain". The small area estimation provides precise estimates of a relative variable of interest for any domain and of which are neither available in any surveys nor censuses. Domains can be either geographical (e.g. regional, provincial, municipal) or sectoral. The aim of this analysis is to point out how much poverty is distributed geographically. The other reasons for this use in practice are to identify the poverty's geographic factors and to address effective transfer payments to who are considered poor. On the other hand, the sectoral one concerns about demographic characteristics such as age, gender or race, and about the analysis of different sector activities and businesses. As we go deeper on the subdivision of administrative units, it shows how poverty and inequalities are getting wider among them. Many techniques have been applied in this issue. Precision and efficiency rely on the collection of the data sources, to their quality itself, to the sample size and to the minimum mean squared error among different methods.

The main procedure to the computation can be either the direct or the indirect estimator. However, it will be shown that it is more convenient the latter one in the practice. In fact, the small area estimation is a combination of survey and census data, with a more detailed variability on it.

Within the model-based estimation, it will be analyzed deeply the EBLUP estimator. In addition, the Empirical Best Predictor will be presented. The last method is the poverty mapping, in which will be included all the steps to construct it. These tools are important to analyze better the domains' disaggregation and to offer a simple comprehension of a relative indicator, e.g. poverty, undernourished children and other social indexes. Moreover, it will be described the loss in precision of the second method defined before, whereas the first one concerns about a re-estimation of the direct indicator and gives more accurate results, in which the main purpose is to lessen the poverty's standard error where it is necessary.

3.2. DATA SOURCES

The principal data sources are the censuses, surveys and auxiliary data, on which Geographic Information Systems (GIS) is well used. A census is carried out on the whole population, but it takes too much time and it is very expensive if done every year. Other disadvantages are that many indicators are missing and it becomes outdated. This latter can be partly resolved by using postcensal estimates. But much information is needed. Surveys, however, are collected more often than the first one. They are focused on specific analyses like the income or consumption and give wider information for subpopulation. It uses a large sampling size in order to give reliable results. It is possible to extract more information by combining them together and generate the small area estimation.

The relevance of GIS in practice are listed below:
- to simplify the PSU stages;
- measure accessibility, like the public infrastructure and the distance towards to the public sanitary;
- to extract data on climatic vulnerability, which influence some sectors and the poverty itself.

It is also important which variables are selected to insert on the analysis, which depends on the well-being and poverty line definition chosen and whether to use the monetary or non-monetary variables. This point must be clear in order to collect the factors to the final computation of unit administrative levels.

3.3. GENERAL MODELS TO CONSTRUCT SAE

Before going ahead, there are three kinds of models that can be applied to estimate the small area:
- Design-based approach
- Model linking
- Linear mixed model

The design-based approach analyzes a representative sample area in order to describe a certain quantity and it holds to the whole population. It is based on inclusion probabilities, which is then used to compute to each of observations in the sample. This model leads to the Horvitz-Thompson estimator. Statistical properties such as the unbiasedness and consistency are satisfied. The former means that the estimator is correct, while the latter consist of that its variance tends to zero if the sample size increases. As the estimator, with its mean and variance, are determined, it is possible to construct the interval confidence, where the real value lies at a significant level, α. The model linking, on the other hand, is based on the stochastic relationship between the variable independent and the dependent one. Therefore, random effects are on the parameter and it considers an additional erratic term on it. It is useful in the practice because it can interpret which determinants are more significant, the factors that influence a relative analysis. By borrowing the strength to both of them, it is constructed the combined model. The purpose of this last model is to obtain an unbiased estimator, applying a model and, at the same time, using a design-based approach. This last model will be discussed in the subsequent paragraph.

Another classification can be the direct and the indirect estimators, either. Direct estimators are the censuses and the surveys. Because of the incompleteness in data sources and the unplanned surveys for any domains, it is necessary to use also the indirect estimators. However, some resolutions can lead towards an optimal direct estimator, relying on many factors. Most of the surveys and censuses are based on large area. Therefore, by providing many small strata, it is possible to draw sample on each of them. Another solution is minimizing the clustering effects because it reduces the effectiveness of the sample size. A different resolution is based also on the sample allocation, integration of surveys, dual frame surveys and repeated surveys (Rao, 2003).

The indirect estimator is more utilized in practice because of the higher reliability. This latter may be split into the implicit and explicit model. The implicit models are synthetic and composite estimators. Their minimum standard error is smaller than the direct estimators. Examples of explicit models are EBLUP, Empirical Bayes and Hierarchical Bayes.

3.4. LINEAR MIXED MODELS

Given a vector of p auxiliary variables, x_i, we are interested in estimating a parameter, θ_i.

$$\theta_i = x_i\beta + z_i v_i$$

It is assumed that the direct estimators are available at sub-domain levels and an unbiased frame sampling is defined as: $\hat{\theta}_i = \theta_i + e_i$

In this paragraph, it is presented the general linear mixed models with more details and its application, following up on:
- Unit-Level Population Model
- Area-Level Population Model
- Synthetic and Empirical BLUP
- Contextual Model

The properties of the BLUP estimator are:
- Best: minimizes the prediction error
- Linear: linear mixed model from Fay-Herriot
- Unbiased: the expected value of the estimator is zero

Ordinary Least Square:

$$\hat{\beta} = argmin_{\beta}(e^{*\prime}e^{*})$$

By the derivative respect to β, it is then obtained the solution of the estimates of $\hat{\beta}$.

$$\hat{\beta} = \frac{X'V^{-1}Y'}{X'VX}$$

Unit-Level Population Model

It computes the population values and it is based on the auxiliary variable, including fixed and random effects. Both of these latter are assumed to be identically and independently distributed and with a Normal distribution.

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{Zu} + \boldsymbol{e}$$

where Y is a vector of the unit population values considered on the whole small area, X is a vector of all the auxiliary variables, with the random effects, σ_u, and the random errors, σ_e. By the Maximum Likelihood and its log-likelihood function, it is possible to estimate β, given that σ_u and σ_e are known. Whether it is unknown, Fisher scoring algorithm is utilized to determine all the parameters: β, σ_u and σ_e.

An example of this model is the **nested-error model**, where y_{ij} is the value of target variable for unit j within area I and v_i is the random effect of area i:

$$y_{ij} = x_{ij}^T\beta + v_i + e_{ij}, \quad j = 1, \dots, N_i, \quad i = 1, \dots, m$$

$$v_i \sim iid\ N(0, \sigma_v^2), \quad e_{ij} \sim iid\ N(0, \sigma_e^2)$$

Area-Level Population Model

It is possible to obtain the area-level model by aggregating the unit-level model to area levels. Here it is presented the population area means:

$$\overline{\boldsymbol{Y}} = \overline{\boldsymbol{X}}\beta + \boldsymbol{u} + \overline{\boldsymbol{e}}$$

The notation is similar to the first model, apart from that it considers mean vectors of each area on it.

An application of this model is the **Fay and Herriot model**, which includes:

- Area effects, vi
- Sampling errors, ei

It is also characterized by these elements:

- Fixed effects, the expected value of the independent variable, y.
- Random effects, which influence the variance-covariance.

This model leads on Empirical Best and Linear Predictor. It is a combination of a model linking area relative to a parameter of interest $\overline{y_i}$ and a sampling model.

$$\theta_i = g(\bar{Y}_i) = x_i^T\beta + v_i, \quad i = 1, \dots, m$$
$$v_i \sim iid\ N(0, \sigma_v^2), \quad \sigma_v^2\ unknown$$

where v_i is the variation between areas.

$$\hat{\theta}_i^{DIR} = \theta_i + e_i, \quad i = 1, \dots, m$$
$$e_i | \theta_i \sim ind(0, \psi_i), \quad \psi_i\ unknown$$

where e_i is the random effects

Combining them together, we obtain:

$$\hat{\theta}_i^{DIR} = x_i^T\beta + v_i + e_i\,, \quad i = 1, \dots, m$$

Given this model:

$$\bar{Y}_d = x_d\beta_d + \mu_d + e_d$$

The BLUP estimator:

$$\hat{\bar{Y}}_d^{BLUP} = \tilde{\mu}_d = x_d\tilde{\beta} + \tilde{\mu}_d = x_d\tilde{\beta} + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_d^2}(\bar{Y}_d - x_d\tilde{\beta})$$

The EBLUP estimator, by substituting $\hat{\sigma}_d^2$ to σ_u^2:

$$\hat{\bar{Y}}_d^{BLUP} = \frac{\hat{\sigma}_d^2}{\hat{\sigma}_u^2 + \hat{\sigma}_d^2}\bar{Y}_d + \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_d^2}(\bar{Y}_d - x_d\tilde{\beta})$$

Random effect variance estimation of this model could be defined by different methods described below.

1. Method of moments

$$\hat{\sigma}_u^2 = \frac{1}{D-p}\left[\sum_{d=1}^{D} \tilde{u}_d^2 - \sum_{d=1}^{D} \hat{\sigma}_d^2 \left(1 - x\left(\sum_{d=1}^{D} x_d' x_d\right)^{-1} x_d'\right)\right]$$

2. Maximum likelihood method:

$$l(\sigma_u^2, \beta; y) = -\frac{D}{2}\ln 2\pi - \frac{1}{2}\ln|V| - \frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta)$$

3. REML method:

$$l_R(\sigma_u^2; y) = -\frac{D-p}{2}\log 2\pi + \frac{1}{2}\log|X'X| - \frac{1}{2}\log|V| - \frac{1}{2}X'V^{-1}X - \frac{1}{2}y'Py$$

Synthetic and Empirical BLUP (regression coefficients are available)
From the estimates of regression parameters, which are $\hat{\beta}^U$and $\hat{\beta}^A$, it is possible to introduce the synthetic techniques, which consist of linking the variable of interest to the auxiliary variables:

$$\widehat{\bar{\boldsymbol{Y}}}^{SU} = \bar{\boldsymbol{X}}_k'\hat{\beta}^U$$

$$\widehat{\bar{\boldsymbol{Y}}}^{SA} = \bar{\boldsymbol{X}}_k'\hat{\beta}^A$$

The derivation of EBLUP is discussed in the subsequent paragraph.

Contextual Model
This model is a combination of the covariates individual and area level. It is useful in practice as in some cases it is necessary to add the latter to the individual level model. It also allows analyzing the ecological fallacy’s effect, which is a divergent result of the regression coefficient estimates from the expectations.
It includes the area-level covariates, called as “contextual effects”:

$$Y_{ik} = \boldsymbol{X}_{ik}^{*\prime}\beta^* + u_k^* + e_{ik}^*$$

3.5. EMPIRICAL BEST LINEAR UNBIASED PREDICTOR (EBLUP)

After a general description of different models of small area estimation, in this paragraph is presented the EBLUP. Its properties are:
- Linear because it derives from mixed linear model
- Unbiased as the estimate is correct and it is equal to the expected value.
- Best because it has the smallest MSE within the overall linear unbiased estimators

BLUP ESTIMATOR
Given a general linear mixed model:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zv} + \boldsymbol{e} \qquad (3.5.1)$$

where **y** is the output and it is a vector of n x 1, while **X** and **Z** are full rank matrices.
It is then analyzed its linear combination:

$$\mu = 1^T\beta + m^T v \qquad (3.5.2)$$

Henderson had proposed the following BLUP estimator, for known δ:

$$\tilde{\mu}^H = t(\delta, y) = 1^T\tilde{\beta} + m^T\tilde{v} = 1^T\tilde{\beta} + m^T G z^T v^{-1}(y - X\tilde{\beta}) \qquad (3.5.3)$$

where:
$\tilde{\beta}$ is obtained using BLUE of β: $\tilde{\beta} = \tilde{\beta}(\delta)$; (3.5.4)

$$\tilde{v} = \tilde{v}(\delta) = G z^T v^{-1}(y - X\tilde{\beta}). \qquad (3.5.5)$$

BLUP ESTIMATOR OF μ_i
From the combined model:

$$\hat{\theta}_i^{DIR} = x_i^T\beta + b_i v_i + e_i, \quad i = 1, \dots, m \qquad (3.5.6)$$

and by substituting some parameters of (3.5.3) on it, we obtain:

$$\mu_i = \theta_i = x_i^T\beta + b_i v_i \qquad (3.5.7)$$

where $1_i = x_i$ and $m_i = b_i$. At the end, the BLUP estimator of θ_i is the following:

$$\tilde{\theta}_i^H = x_i^T\tilde{\beta} + \gamma_i(\hat{\theta}_i - x_i^T\tilde{\beta}) = \gamma_i\hat{\theta}_i + (1-\gamma_i)x_i^T\tilde{\beta} \qquad (3.5.8)$$

EBLUP ESTIMATOR
Here is considered the δ unknown. Therefore, it is necessary a two stage estimator because of the estimation of the variance parameter δ: $\hat{\delta} = \hat{\delta}(y)$, so that it becomes $\mu^H = t(\hat{\delta}, y)$.
It remains an estimator unbiased under certain condition (Rao, 2003).

3.5.1. BLUP under the Fay-Herriot model

To sum up, the combination of a direct estimator and a model linked gives us the Fay-Herriot mixed model. Then, it is presented the linear and unbiased estimator of θ_i:

$$\theta_i^{BLUP} = x_i^T\tilde{\beta} + \tilde{v}_i \qquad (3.5.9)$$

In some cases, micro data could not be obtained and here is the importance of this method. It has been shown before the derivation of EBLUP, which is a model that estimates small area for each θi. It represents the weighted mean of the direct estimator and the regression synthetic estimator $x_i^T\tilde{\beta}$, in which $\tilde{\beta}$ has properties of BLUE.

$$\hat{\theta}_i^{EBLUP} = \gamma_i\hat{\theta}_i^{DIR} + (1-\gamma_i)x_i^T\tilde{\beta} \qquad (3.5.10)$$

The relevance of weight is given to the former element, the direct estimator, in the case of small sampling variance. Otherwise, it weighted more the synthetic estimator. This latter is based on the coefficient β, using the linear regression. However, the synthetic estimator does not count the area-based random effects, so that this estimator cannot be precise alone. The larger is the sample size, the more reliable the direct estimator. Combining both of these components, we obtain a composite estimator and a more accurate indicator's result.
The other elements of the model are computed as:

$$\tilde{\beta} = \left(\sum_{i=1}^{m}\gamma_i\mathbf{x_i}\mathbf{x_i^T}\right)^{-1}\sum_{i=1}^{m}\gamma_i\mathbf{x_i}\hat{\theta}_i^{DIR} \qquad (3.5.11)$$

$$\tilde{v}_i = \gamma_i\left(\hat{\theta}_i^{DIR} - \mathbf{x_i^T}\tilde{\beta}\right) \qquad (3.5.12)$$

$$\gamma_i = \frac{\sigma_v^2}{\sigma_v^2+\psi_i} \qquad (3.5.13)$$

The gamma coefficient is called as shrinkage factor. It measures the relationship between the model and the total variance. The lesser is the shrinkage factor, the better is the estimate, which will gain more accuracy than the direct one.
This methodology is useful when there is the availability of auxiliary variables within domains, and not concerning to the individual or familiar level. The drawback of this analysis will be limited only to that relative disaggregation and cannot go further of this domain levels.

ESTIMATION OF MSE:

$$MSE(\hat{\theta}_i^{EBLUP}) \approx E(\hat{\theta}_i^{EBLUP} - \theta_i)^2$$

Approximation of MSE by Taylor linearization method:

$$MSE\left(\hat{\theta}_i^{EBLUP}\right) = g_1(\sigma_v^2) + g_2(\sigma_v^2) + g_3(\sigma_v^2)$$
$$mse\left(\hat{\theta}_i^{EBLUP}\right) = g_1(\hat{\sigma}_v^2) + g_2(\hat{\sigma}_v^2) + g_3(\hat{\sigma}_v^2)$$

- $g_1(\hat{\sigma}_v^2)$ is due to prediction of random effects:

$$g_1(\hat{\sigma}_v^2) = \gamma_i\psi_i$$

- $g_2(\hat{\sigma}_v^2)$ is due to estimation of β:

$$g_2(\hat{\sigma}_v^2) = \sigma_v^2(1-\gamma_i)^2\boldsymbol{x}_i^T\left(\sum_{i=1}^{m}\gamma_i\boldsymbol{x}_i\boldsymbol{x}_i^T\right)^{-1}\boldsymbol{x}_i$$

- $g_3(\hat{\sigma}_v^2)$ is due to estimation of $\hat{\sigma}_v^2$:

$$g_3(\hat{\sigma}_v^2) = (1-\gamma_i)^2\gamma_i\sigma_v^{-2}\bar{V}(\hat{\sigma}_v^2)$$

The $\hat{\theta}_i^{EBLUP}$ gains more efficiency when γ_i is small, meaning that the model error is small relative to the total variability. The expected value of this MSE is nearly unbiased, where a negligible value is added to the true MSE estimates:

$$E\{mse(\hat{\theta}_i^{EBLUP})\} = MSE(\hat{\theta}_i^{EBLUP}) + o(1/m)$$

PARAMETRIC BOOTSTRAP ESTIMATION OF MSE

Once the model fitting of $\hat{\sigma}_v^2$ and $\hat{\beta}$ are estimated, the next step is to generate bootstrap relative to the area effects and to the sampling errors, where they are independently to each other. Then, a bootstrap of direct estimators from the sampling model is generated. These information enable us to fit the model to new bootstrap data and calculate the bootstrap EB estimator. The previous steps before is repeated a large numbers, B. Therefore, it is possible to calculate the bootstrap MSE estimator:

$$mse(\hat{\theta}_d^{EB}) = \frac{1}{B}\sum_{i=1}^{B}(\hat{\theta}_i^{EB*(b)} - \theta_i^{*(b)})^2$$

The estimation of $\hat{\sigma}_v^2$ can be obtained by:

- Moment estimator: $\hat{\sigma}_{vs}^2$
- Maximum Likelihood estimation: $\hat{\sigma}_{vML}^2$
- Residual Maximum Likelihood: $\hat{\sigma}_{vRE}^2$

They are unbiased models, provided that v_i and e_i are symmetrically distrusted around zero and normally distributed. The relative efficiency of estimators of $\hat{\sigma}_v^2$ is:

$$\bar{V}(\hat{\sigma}_{vRE}^2) = \bar{V}(\hat{\sigma}_{vML}^2) \le \bar{V}(\hat{\sigma}_{vm}^2) \le \bar{V}(\hat{\sigma}_{vs}^2)$$

where s stands for simple moment estimator and m is for Fay-Herriot moment estimator.

3.5.2. NESTED-ERROR MODEL (BLUP and BP under a finite population)

$$\tilde{\bar{Y}}_i^{BLUP} \approx \gamma_i\{\bar{y}_i + (\bar{\boldsymbol{X}}_i - \bar{\boldsymbol{x}}_i)^T\tilde{\beta}\} + (1-\gamma_i)\bar{\boldsymbol{X}}_i\tilde{\beta}$$

Variance components: $w = (\sigma^2, \rho)'$

Estimation MSE:

$$MSE[\tilde{\theta}_d(\widehat{w})] = g_{1d}(w) + g_{2d}(w) + g_{3d}(w)$$
$$MSE[\tilde{\theta}_d(\widehat{w})] = g_{1d}(w) + g_{2d}(w) + g^{PR}{}_{3d}(w)$$

where:

$$\hat{g}_1(\sigma) = (1 - \hat{\gamma}_k)\hat{\sigma}_u^2$$
$$\hat{g}_2(\sigma) = (\boldsymbol{X}_k - \hat{\gamma}_k\bar{x}_k)'\left[\widehat{MSE}_\xi(\beta)\right](\boldsymbol{X}_k - \hat{\gamma}_k\bar{x}_k)$$
$$\hat{g}_3(\sigma) = \left(\frac{\hat{\sigma}_e^2}{n_k}\right)\left(\hat{\sigma}_u^2 + \frac{\hat{\sigma}_e^2}{n_k}\right) + \left[Var(\hat{\sigma}_u^2) + \frac{\hat{\sigma}_u^2}{\hat{\sigma}_e^2}Var(\hat{\sigma}_e^2) - 2\frac{\hat{\sigma}_u^2}{\hat{\sigma}_e^2}Cov(\hat{\sigma}_u^2, \hat{\sigma}_e^2)\right]$$

$$\hat{\gamma}_k = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \frac{\hat{\sigma}_e^2}{n_k}}$$

Parametric Bootstrap Estimation of MSE

Firstly, the model fitting of $\hat{\sigma}_v^2$, $\hat{\sigma}_e^2$ and $\hat{\beta}$ are estimated, using Maximum Likelihood, Residual Maximum Likelihood or Henderson method 3. The next step is to generate bootstrap relative to the area effects and to the sampling errors, where they are independently to each other. Then, a bootstrap of direct estimators from the sampling model is generated. It follows the calculation of the target quantities for the bootstrap population:

$$F_{\alpha i}^* = \frac{1}{N_i}\sum_{i=1}^{N_i} F_{\alpha ij}^*, \quad F_{\alpha i}^* = h_\alpha\left(y_{ij}^*\right), \quad i = 1, \ldots, m$$

From the bootstrap population, it must be taken only the indices sampled. After this latter step, it is fitted the model to bootstrap sampled and obtained the bootstrap EBP. The previous steps before is repeated for large numbers, B. Therefore, it is possible to calculate the bootstrap MSE estimator:

$$mse\left(\hat{F}_{\alpha i}^{EB}\right) = \frac{1}{B}\sum_{i=1}^{B}\left\{\hat{F}_{\alpha i}^{EB}(b) - F_{\alpha i}^*(b)\right\}^2$$

3.6. Empirical Best Prediction of non-linear domain parameters with unit-level models

Empirical Best (EB) is another methodology to obtain an empirical best domain both linear and non-linear estimators, using regression models with unit level. This method is used to estimate indicators such as HCR, Fuzzy monetary, Fuzzy supplementary. It is also useful for binary data, count data and linear mixed model.

The procedures are the following:

1. Obtain the posterior density of the small area parameter of interest, μ: f(μ|y,λ)

2. Estimate the model parameters, λ, from the marginal density: f(y|λ)

3. Use the estimated conditional density for making inferences about μ
Optimal estimator of θ_i is derived from the conditional expectation of θ_i:

$$E\left(\theta_i|\hat{\theta}_i,\beta,\sigma_v^2\right) = \hat{\theta}_i^B = \gamma_i\hat{\theta}_i + (1-\gamma)z_i^T\beta$$

where: $\gamma = \frac{b_i^2\sigma_v^2}{b_i^2\sigma_v^2+\psi_i}$

It is optimal because the MSE is the smallest among the other estimator of θ_i and it is also called as Best prediction (BP) estimator of θ_i because it is obtained from conditional distribution without the prior distribution.

EB or EBP estimator of $\boldsymbol{\theta_i}$ from $\boldsymbol{\hat{\theta}_i^B}$

$$\hat{\theta}_i^{EB} = \hat{\theta}_i^B\left(\hat{\beta},\sigma_v^2\right) = \hat{\gamma}_i\hat{\theta}_i + (1-\hat{\gamma})z_i^T\hat{\beta}$$

The EB estimator is equal to the EBLUP estimator, under normality.

EB or EBP is applied also to find the EB estimator of any function: $\phi_i = h(\theta_i)$
The EB estimator is obtained from the Bayes estimator: $\hat{\phi}_i^B = E(\phi_i|\,\hat{\theta}_i,\beta,\sigma_v^2)$
Its computation requires Monte Carlo approximation:

$$\hat{\phi}_i^{EB} \approx \frac{1}{R}\sum_{r=1}^{R} h(\theta_i^{(r)})$$

This latter can be simplified by:

$$\hat{\phi}_i^{EB} \approx \frac{1}{R}\sum_{r=1}^{R} h(\hat{\theta}_i^{EB} + z_i^{(r)}\sqrt{\hat{\gamma}_i\psi_i})$$

where r=1,…,R. The higher is r, the more accurate will be the approximation.

MSE estimation
Because of the equality between the EB and EBLUP, it is also applied to the mean squared error estimation:

$$MSE\left(\hat{\theta}_i^{EB}\right) = MSE\left(\hat{\theta}_i^{EBLUP}\right) = g_1(\sigma_v^2) + g_2(\sigma_v^2) + g_3(\sigma_v^2)$$

where each factor of the right-hand has been defined before.

Another method is the Jackknife method, proposed by Jiang, Lahiri and Wan.
Decomposition of MSE($\hat{\theta}_i^{EB}$):

$$\mathrm{MSE}\left(\hat{\theta}_i^{EB}\right) = M_{2i} + M_{1i}$$

Here are presented the steps to calculate it:

1. From the full data set {($\hat{\theta}_i$, z_i); i = 1,…,m}, an lth area of data set ($\hat{\theta}_l$, z_l) is deleted in order to calculate the m estimators of β and σ_v^2. The estimators of θ_i: $\hat{\theta}_{i,-l}^{EB} = k_i(\hat{\theta}_i, \hat{\beta}_{-l}, \hat{\sigma}_{v,-l}^2)$

2. The estimators of $\widehat{M}_{2i}$ is the variability for estimating the parameters:

$$\widehat{M}_{2i} = \frac{m-1}{m}\sum_{l=1}^{m}(\hat{\theta}_{i,-l}^{EB} - \hat{\theta}_i^{EB})^2$$

3. The estimators of $\widehat{M}_{1i}$ is the MSE when the model is known:

$$\widehat{M}_{1i} = g_{1i}(\hat{\sigma}_v^2) - \frac{m-1}{m}\sum_{l=1}^{m}[g_{1i}(\sigma_{v,-l}^2) - g_{1i}(\sigma_v^2)]$$

4. Jackknife estimator of $MSE(\hat{\theta}_i^{EB})$:

$$mse(\hat{\theta}_i^{EB}) = \widehat{M}_{1i} + \widehat{M}_{2i}$$

It is applicable to ML, REML, moment estimators and EB estimator of any function.
A more complex model is HB approach, which is also used to handle small area models and avoids underestimation of the mean standard of error.

Linear mixed model

The best predictor estimator of μ_i is given by the conditional expectation of μ_i, given the parameters:

$$\hat{\mu}_i^B = \hat{\mu}_i^B(\beta, \delta) = E(\mu_i|y_i, \beta, \delta) = 1_i^T\beta + m_i^T\hat{V}_i^B$$

The $\hat{\mu}_i^B$ depends on β and δ, which are estimated using ML or REML. After we obtain the estimators of $\hat{\beta}$ and $\hat{\delta}$, we determined the empirical BP estimator of μ_i:

$$\hat{\mu}_i^{EB} = \hat{\mu}_i^{EB}(\hat{\beta}, \hat{\delta}) = 1_i^T\hat{\beta} + m_i^T\hat{V}_i^B(\hat{\beta}, \hat{\delta})$$

Estimation of the MSE

$$mse(\hat{\mu}_i^B) = \widehat{M}_{1i} + \widehat{M}_{2i}$$

$$\widehat{M}_{2i} = \frac{m-1}{m}\sum_{l=1}^{m}(\hat{\mu}_{i,-l}^{EB} - \hat{\mu}_i^{EB})^2$$

$$\widehat{M}_{1i} = g_{1i}(\hat{\delta}) - \frac{m-1}{m}\sum_{l=1}^{m}[g_{1i}(\hat{\delta}_{-l}) - g_{1i}(\hat{\delta})]$$

where $g_{1i}(\hat{\delta})$ is from the MSE of BLUP estimator.

3.7. POVERTY MAPPING

"The Poverty Mapping is a methodology for providing a detailed description of the spatial distribution of poverty and inequality within a country. It combines individual and household survey data and population census data with the objective of estimating welfare indicators for specific geographic area as small as village or hamlet." (Wikipedia definition)

Another method to provide poverty's level among domains is the poverty mapping. The methodology described for this SAE's technique is contributed by Elbers, Lanjouw and Lanjouw (2003) and it is a combined model. The first step is the collection of data from Census at micro level and surveys. Then, comparison and harmonization of the data sources are necessary. This latter consist of examining the coherence on definitions, measurement, procedures and with reference periods close as possible to the poverty mapping's analysis. The Census will give the number of the population relative to a certain year, with detailed data concerning of the household average. The survey is conducted using Enumeration Areas (EAs) and Primary Selection Units (PSUs), on which will be taken the sample size, and give an estimation of the consumption or expenditure level. The next step is to define the variables to insert on the regression, which will be the main element to the computation of the poverty measurement at the end. Common variables will be gathered together from both of the sources. If there are some missing values on the variables selected, it is used the imputation procedure.
The level of disaggregation must be decided in order to compute poverty or inequalities within the country. This method allows a much deeper study on the sub-administrative level. However, the precision relies on many factors and will be discussed further the aggregating data.

3.7.1. Estimation of the model

The following stage is the prediction model for consumption, in which the covariates variables for whole the population are selected from the comparison of Census and survey. The dependent variable of the model is a monetary indicator, e.g. disposable income. Then it is added an error component, to a better estimation of the model due to the within-cluster correlation in the disturbances and the heterogeneity.

Therefore, here is presented a linear approximation to the conditional distribution of the logarithm consumption expenditure of household h in cluster c:

$$\ln y_{c,h} = E\left[y_{c,h} \mid x_{c,h}^{T}\right] + u_{c,h} = x_{c,h}^{T}\beta + u_{c,h} \tag{3.7.1.1}$$

where y is expenditure consumption, x are the main regressors that influence significantly the consumption and that may base on basic foods, utilities and other variables chosen before. The erratic term u is composed by the cluster level effects (η_c) and the household level effects ($\varepsilon_{c,h}$):

$$u_{c,h} = \eta_c + \varepsilon_{c,h} \tag{3.7.1.2}$$

Its estimation is the following:

$$\hat{u}_{c,h} = \hat{\eta}_c + e_{c,h} \tag{3.7.1.3}$$

The estimation of household specific variance:

$$\hat{\sigma}^2_{c,h} = \left[\frac{AB}{1+B}\right] + \frac{1}{2}\mathrm{var}(r)\left[\frac{AB(1-B)}{(1+B)^3}\right] \tag{3.7.1.4}$$

where:

$$A = (1.05) * \max(e_{c,h}) \tag{3.7.1.5}$$

$$\ln\left[\frac{e_{c,h}}{A-e_{c,h}}\right] = z'_{c,h}\alpha + r_{c,h} \tag{3.7.1.6}$$

$$B = z'_{c,h}\alpha \tag{3.7.1.7}$$

3.7.2. Simulation

The distribution of this model is important to generate expenditure for each household, using the Census' data. It is taken a considerable number of simulations. For each of them, the parameter $\tilde{\beta}$ of the regression model are drawn, whereas the disturbance term ($\widetilde{u_{c,h}}$) are determined by bootstrap procedure.

$$\hat{y}_{c,h} = \exp\left(x^T_{c,h}\tilde{\beta} + \widetilde{u_{c,h}}\right) \tag{3.7.2.1}$$

where $\hat{y}_{c,h}$ is the simulated values of the consumption expenditure of each household.

3.7.3. Level of aggregation of the census data: the loss in precision

In this paragraph the loss in precision using census aggregate data is analyzed. The result is that there will be always an underestimation (overestimation) whenever the poverty's rate is below (above) 50%. The further is the aggregation of census data, the wider the trade-off between the household data to the other domain selected. In this case, the EAs (figure 3.1) and the provincial level (figure 3.2) are presented. Special cases are presented when poverty is equal to 50%, to 0% or to 100%, where they have the same results at different aggregate census data. This means that it is better to use household level census data. In practice, however, it may not possible to obtain this information so that it is important to concern about the loss of the precision. For instance, errors in poverty's estimates in EAs census level is lower than the provincial or regional aggregating census data.

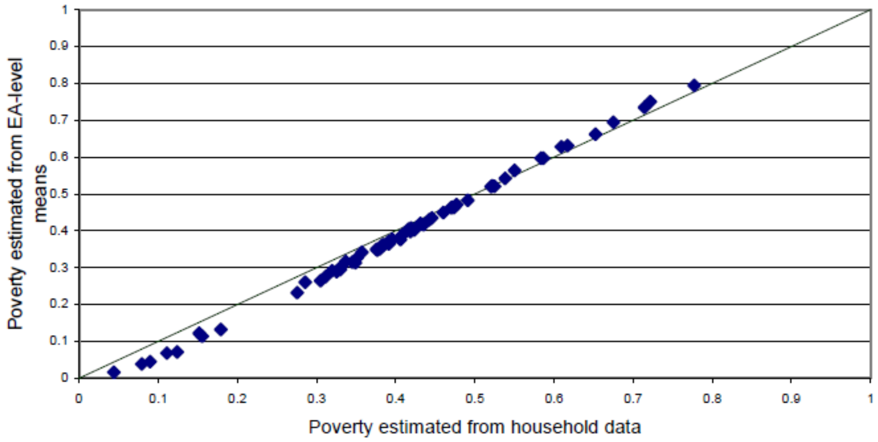


Figure 3.1: Comparison of provincial poverty estimates using household-level census data and using enumeration area means
Source: International Food Policy Research Institute

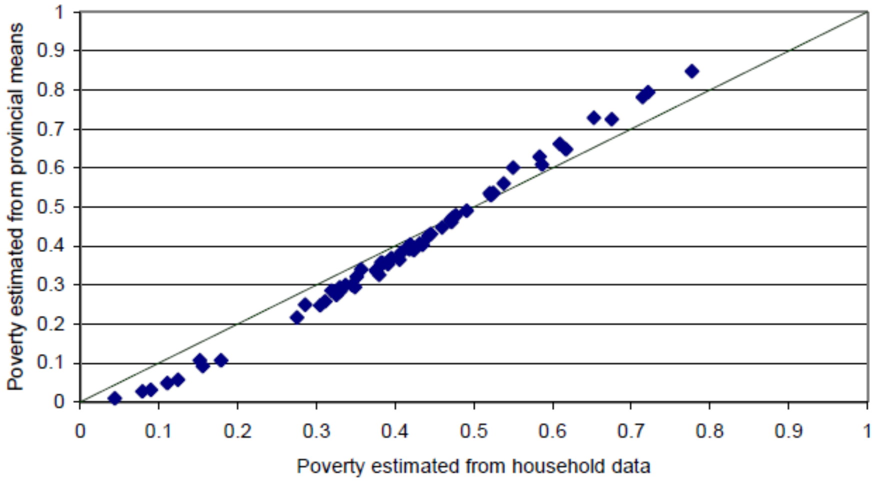


Figure 3.2: Comparison of provincial poverty estimates using household-level census data and using provincial means
Source: International Food Policy Research Institute

3.7.4. Poverty Measurement

Once the stage of simulation is done, it is possible to compute the poverty indexes. The principal indicators are headcount ratio, poverty gap and severity poverty. To summarize all these formulae, Foster-Greer-Thorbecke had proposed:

$$\mathrm{FGT} = \frac{1}{n}\sum_{i=1}^{q}\left(\frac{y_q - y_i}{y_q}\right)^{\alpha}$$

where y_q is the poverty line and α is the elasticity of poverty's adversity.

Within inequalities, it can be calculated the Gini coefficient:

$$\text{Gini} = \frac{2}{n^2}\sum_{i=1}^{n}\left(\frac{y_i - \bar{y}}{\bar{y}}\right)$$

Their interpretation is explained in the previous chapter. These indexes are then drawn in each sub-administrative level and it shows the poverty's differences among domains, giving a better comprehension of a targeted strategy to deprive poverty.

4. GENERAL POVERTY IN THE PHILIPPINES

4.1. General poverty

The general economy in the Philippines are showing an increase of Gross National Product (GNP) in these last years. However, it does not exist a truly correlation between this increase and the poverty reduction. Therefore, in this chapter, some causes of poverty will be analyzed further. If we compare this nation to the Association of Southeast Asian Nations (Asean) member, the trend of improvements is very little, even though it was initially better than them. The main products of the Philippines are rice, corn, sugar, coconut and other crops. But the income gained from this sector did not contribute a lot to the general growth for three reasons:

- No access to irrigation water
- New technologies are less effective
- A little percent of land can be irrigated

Because of this limits, the income gained in this sector cannot sustain the growth of the general economy and, thus, to reduce poverty.

4.2. Targeting by government

To eradicate poverty, the government has experienced its intervention by poverty targeting. Targeting measures can be classified as:

- Targeting by activity, in which the benefits are distributed progressively based on the activity that can receive the benefits;
- Targeting by indicator is based on income or other factors that identify who is poor;
- Targeting by location sets the benefits relying on the area of residence;
- Targeting by self-selection or self-targeting, finally, is designed only for the poor.

In this argument, it should better to include also two forms of error, which are consisted by undercoverage and leakage. The former is verified when the benefits do not reach the target group, while the latter occurs when it is addressed to somebody who does not belong to that relative group.

The benefits that poor should receive is determined by the poverty line, z, in order to bring their level to the boundary: z-y is the total amount that should be transferred.

The more targeting ratio is close to unity, the lesser efficient is the poverty targeting, because it means that the benefits are being received to non-target group (B), who does not have the necessity as those who are in the target group:

$$targeting\ ratio = \frac{B}{(C+y)}$$

where,

y - target group receiving benefit

B - non-target group receiving benefit

C - target group not receiving benefit

However, other factors that causes difficulty to the poverty targeting can be also the lack of information, the costs of individual participation, the incentive effects, high administrative costs and unavailability of benefit terms in the end.

Targeting in the Philippines can be categorized into the broad targeting, the type of interventions, and narrow targeting, which makes the separation between the inclusion and the exclusion area in order to qualified the beneficiaries. Some on government interventions are the following:

- Aquino administration: Tulong sa Tao – a provision of subsidized credit;
- Ramos administration: Comprehensive and Integrated Delivery of Social Services (CIDSS) – a project dependind on 33 indicators to meet the minimum basic needs, to maintain security and to enable the fundamental environment such as basic education and people's participation;
- Estrada administration: Care for Poor identifies the 100 poorest families in each province and city. It has also the purpose to meet the basic needs and it provides benefits like food supports and medical assistance;
- Macapagal-Arroyo administration: KALAHI (Kapi-Bisig Laban sa Kahirapan) – it is combined to the CIDSS and it is based in two stages. The first stage is based on ranking the top 40 poverty's indicator and the second one consists on selecting the poorest quarter of municipalities;
- National Food Authority (NFA) – rice redistribution of the government to give subsidy to the farmers as it is bought at higher price than in the market place and subsidy to the consumer because it is sold for sales;
- Aquino's government has also introduced Conditional Cash Transfer (CCT) program. NSCB argued that this program did not improve the income inequality because only ¼ of budget is accounted from the total amount, which is not enough to fight poverty;
- Pantawid Pamilyang Pilipino Program 4Ps ha the purposes to invest to children's future, keep them in school and heathier.

From a simulation study of Balicasan and Pernia (2003) of longitudinal provincial data for the 1980s and 1990s, it is analyzed the average per capita expenditure per quantile.
The explanatory variables are:

- Initial condition variables – human capital endowment, farm and land characteristics, social capital, geographic attributes, political economy characters.
- Time-varying variables – relative price incentives, road access, electricity, agrarian reform, overall average per capita income.

The simulation's results are that the irrigation and local dynasty is significant from 1st to 3rd quantile. The typhoon causes negative effects to the all income percentile. Farm size benefits only the richest quantile. Among the time-varying variables, per capita income and the terms of trade are significant. The roads are also substantial to the regression, but it increase more when it is interacted to the schooling.

4.3. Main causes of Poverty and Vulnerability

Poverty is when a household or family are defined poor for a certain period of time. Vulnerability happens when there is high probability of poverty even they do not experience it every year. There are two main classification of vulnerability: the chronic poverty is a household that experienced once or twice poor, while the transient one fell into poverty more than two times.

An analysis of vulnerability is presented by Bayudan-Dacuycuy Coonie and Lim Joseph Anthony (2014). They utilized panel data, in which is determined how many times the household fall into poverty. By merging datasets from Annual Poverty Indicator Survey (APIS) and Family Income Expenditure Survey (FIES) from 2003 to 2008, it is obtained the information necessary to make the analysis concerning the poverty's frequency and who is vulnerable to poverty. The method utilized are Multinomial Logit Regressions to identify the reasons of poverty's frequency and Probit regressions for panel data. The measure of vulnerability is calculated in order to obtain the transient and the chronic poor, and which variables have impacts to this analysis.

At national level 59% of household has experienced the poverty per capita expenditure against poverty threshold, while per capita food expenditure against food threshold is 75%. In both per capita expenditure against poverty threshold and per capita food expenditure against food threshold the poverty has more impact on the rural than the urban. At level regions, the poverty's frequency is found in Cordillera Administrative Region (CAR), CARAGA and ARMM while, in Metro Manila it is less than the other regions. Another consideration to take into account is that per capita food expenditure against food threshold has a more incidence to be at least once poor than per capita expenditure against poverty threshold does.

The main significant socioeconomic variables to be among "never poor households" are:

- Higher education
- Older people and for the age of the spouse of the household head
- Service makers
- Mean family size
- Lowest mean number young members 0-15 years old
- Higher mean number household members aged over 25

Marginal effects based on the estimates of the Multinomial Logit regressions per capita expenditure against poverty threshold shows that a household headed by a higher education contributes to be less effective to the poverty's frequency. Between household younger (0-15 years old) and elder people (16-25 years old), the first case has a more incidence to poverty than the second case. Positive shocks are less likely to be always poor, while negative shocks have more effects on the vulnerability of poverty.
Using Probit regressions, a comparison between the transient and chronic poor to the never poor per capita expenditure against poverty threshold has been examined, the results are:

- As the education reached the higher level, it has effects to be less transient or chronically poor;
- Household with older heads are less likely to be poor;
- On occupation, those that belonged to primary sector are more likely to experience both type of vulnerability. This consequence is the similar to who works in trade, but less than

the first category. Professionals, technician are less likely to be transient poor;
- Age level of the members of the household, in which the older than 25 years old is less likely to the vulnerability, which is more pronounce as the age are getting smaller than them;
- Positive shocks and negative shocks have similar effects as the previous results, but relative to be likely transient and chronic poor.
- Armed conflict regions is also one of the weakness causing more poverty.

Using Probit regressions, a comparison between chronic poor and the transient poor are studied. Education with higher level, age members in the household and region with armed conflict are more likely to be chronic poverty and chronic food poor in the same manner to the descriptions above. Among all the occupations, farmers, forestry workers, fishermen, laborers and unskilled workers are likely to experience both poverty. Positive shocks, like new job with higher salary, are determinant to be less chronic and chronic food poor. Negative shocks such as natural disaster are more likely to be chronic food poor.

Policy makers should introduce some interventions direct in specific sectors in order to diminish chronic poverty and the vulnerability of households:
- Education attainment, which will lead to a better job qualification and an adequate wages;
- Development of infrastructures, very important to improve the education and sanitary level;
- Employment stability and market promotion;
- Better health care and allowing health insurance;
- Intervention to different level of poverty defined by the multidimensional approach, even the government has introduced some programs to mitigate poverty, such as CCT;
- And finally, a program in case of natural disasters.

4.4. Sample size determination

Basing on the CENSUS of 2000, the sub domains of the Philippines are divided in 16 regions, 83 provinces, 1,623 municipalities and 41,926 barangay. The next CENSUS (2010) has made a little change, moving into 17 regions and 85 provinces, considering also the other alteration due to new positions of municipalities and barangays.
The construction of the poverty line is based on cost-basic-needs approach, where it is the sum of food expenditure and non-food allowance. Non food expenditure- clothing, housing, transportation, health, and education expenses, and others - is determined by the ratio between food expenditure and total basic expenditure. Household is the unit utilized to estimate the poverty indicators.
The administrative units are the barangay, which the total number is 41,952. In the formation of PSUs, least accessible barangays (LABs), barangays with peace and order problem (POPs) and the number of households are considered on it. Because LABs represents a little proportion of the whole barangay, it is excluded from the frame. On the other hand, POPs has a significant number and it is assumed as a temporary situation, therefore it is included with the non-POPs barangay. Lastly, if the administrative unit is more than 500, then it is certainly be part of PSU; otherwise, it is necessary to merge more than a barangay to obtain 500 households.
Because of the presence of 350 LABs, therefore only 41,592 barangays are counted. The total number of PSUs formed is 16,582 to master sample frame (MS). The total sample size are determined from the poverty incidence by region. This sub-domain level allow us a little coefficient of variation (5%) and therefore more precise estimates, apart from NCR, which has 10%. The result

is counted as 44,000 households.
The sample allocation has two different methods. It can be proportional to the total number of households in the domain when precise estimates at national level is preferred; otherwise, the total sample size is allocated equally across domains to obtain precise estimates at domain level. The Kish allocation scheme is utilized in order to calculate the sample size per region.
Once the MS PSUs are determined, they are selected by systematic sampling with probabilities proportional to measure of sizes based on their number of households in the 2000 CPH. Large probabilities are excluded because they can be counted more than once. Then, the number of PSUs to be sampled per region is computed by the ratio within the sample allocated sample size and the desired subsample size per PSU.

4.5. Poverty trend of the Philippines

The poverty indicators computed and analyzed in this section are:

- Poverty incidence among families
- Income gap
- Poverty gap
- Severity of poverty

In the small area estimation, the poverty mapping or ELL method is applied, one of the method discussed on the second chapter. Using survey and household CENSUS together, it will be estimated the poverty of sub-population. The auxiliary variables utilized can be either income or expenditure. After the identification of auxiliary variables, X, the first stage regressions is implemented, in which consist of the problem of multicollinearity and the search for significant relationship. Then, in order to estimate the poverty sub-population, survey and household CENSUS are merged. We would like that the auxiliary variables have a large variability for each domain because it cause also a reliable estimates, in which the erratic terms are very small. Another factor to be considered is the number of sample size at each level in order to minimize the standard error. By the log transformation of expenditure/income, the averaged for each small area to produce a point estimate is found. Then, the bootstrapping simulation is applied. In the end, there is the production of final estimates, where the predicted values is transformed in exponential. By doing so, the predicted expenditure/income at sub-domain level is computed. Once we have chosen the poverty indicator, it is calculated using these values that have just found. We can compute the point estimates by taking the mean and standard deviation from the 100 bootstrap estimates.

Annual Per Capita Poverty Threshold and Poverty Incidence among Families, by Region 1991, 2006, 2009 and 2012

	Per Capita Poverty Threshold (PhP)				Poverty Incidence among Families (%)			
	1991	2006	2009	2012	1991	2006	2009	2012
PHILIPPINES	5,949	13,357	16,871	18,935	29.7	21.0	20.5	19.7
NCR	7,373	15,699	19,227	20,344	5.3	2.9	2.4	2.6
CAR	6,706	14,107	17,243	19,483	36.7	21.1	19.2	17.5
Region I	6,371	14,107	17,595	18,373	30.6	19.9	16.8	14.0
Region II	6,525	13,944	17,330	19,125	37.3	21.7	20.2	17.0
Region III	6,635	14,422	18,188	20,071	18.1	10.3	10.7	10.1
Region IV-A	6,409	13,241	17,033	19,137	19.1	7.8	8.8	8.3
Region IV-B	5,753	12,645	15,613	17,292	36.6	32.4	27.2	23.6
Region V	5,807	13,240	16,888	18,257	48.0	35.4	35.3	32.3
Region VI	5,580	12,684	15,971	18,029	32.3	22.7	23.6	22.8
Region VII	5,670	13,963	16,662	18,767	38.2	30.7	26.0	25.7
Region VIII	5,507	12,520	16,278	18,076	42.3	33.7	34.5	37.4
Region IX	5,330	12,743	16,260	18,054	36.4	40.0	39.5	33.7
Region X	5,529	12,917	16,878	19,335	42.6	32.1	33.3	32.8
Region XI	5,653	13,389	17,120	19,967	34.1	25.4	25.5	25.0
Region XII	6,272	13,319	16,405	18,737	47.4	31.2	30.8	37.1
Caraga	6,099	14,324	18,309	19,629	48.5	41.7	46.0	31.9
ARMM	5,201	12,647	16,683	20,517	26.9	40.5	39.9	48.7

Source: Philippine Statistics Authority

In all the poverty indicators, both region XII and ARMM has increased their poverty, while Caraga is showing a better off change. In the poverty inceidence among families, the situation in ARMM has worsen by 8.2%, following the Region XII by 5.9%. On the other hand, there are also some improvements in Region I (5.9%), in Region IVB (8.8%), Region VII (5%), Region IX (6.3%) and Region XIII (9.8%).

By making the analysis in the income gap, generally, there has been an improvement, in which the better results to the whole region are region XI (6%) and Caraga (4.5%). In area of Region III, Region XII and ARMM, there have had a worse situation than of three years ago and they are 0.5%, 3.3% and 3% respectively.

The of region VIII has increased of 0.5% relative to the poverty gap, whereas region IX has reduced by 4.2%. Similar trends follows also here the region XII, Caraga and ARMM in both of the poverty gap and severity of poverty.

Poverty incidence among families

Region	Poverty Incidence among families (%)			Increase/Decrease		
	2006	2009	2012	20006-2009	2009-2012	2006-2012
PHILIPPINES	21,0	20,5	19,7	-0,5	-0,8	**-1,3**
NCR	2,9	2,4	2,6	-0,5	0,2	-0,3
CAR	21,1	19,2	17,5	-1,9	-1,7	-3,6
Region I	19,9	16,8	14,0	-3,1	-2,8	-5,9
Region II	21,7	20,2	17,0	-1,5	-3,2	-4,7
Region III	10,3	10,7	10,1	0,4	-0,6	-0,2
Region IV-A	7,8	8,8	8,3	1,0	-0,5	0,5
Region IV-B	32,4	27,2	23,6	-5,2	-3,6	-8,8
Region V	35,4	35,3	32,3	-0,1	-3,0	-3,1
Region VI	22,7	23,6	22,8	0,9	-0,8	0,1
Region VII	30,7	26,0	25,7	-4,7	-0,3	-5,0
Region VIII	33,7	34,5	37,4	0,8	2,9	3,7
Region IX	40,0	39,5	33,7	-0,5	-5,8	-6,3
Region X	32,1	33,3	32,8	1,2	-0,5	0,7
Region XI	25,4	25,5	25,0	0,1	-0,5	-0,4
Region XII	31,2	30,8	37,1	-0,4	6,3	5,9
Caraga	41,7	46,0	31,9	4,3	-14,1	-9,8
ARMM	40,5	39,9	48,7	-0,6	8,8	8,2

Source: National Statistical Coordination Board

Income gap

Region	Income Gap			Increase/Decrease		
	2006	2009	2012	20006-2009	2009-2012	2006-2012
PHILIPPINES	***27,5***	***26,2***	***26,2***	***-1,3***	***0,0***	***-1,3***
NCR	17,9	16,6	17,2	-1,3	0,5	-0,7
CAR	29,2	28,6	26,9	-0,7	-1,7	-2,3
Region I	23,3	22,8	21,9	-0,5	-0,9	-1,4
Region II	23,1	23,4	21,5	0,3	-1,9	-1,6
Region III	20,9	21,6	21,4	0,7	-0,2	0,5
Region IV-A	22,7	19,4	22,1	-3,3	2,7	-0,6
Region IV-B	28,5	25,3	26,0	-3,2	0,7	-2,5
Region V	28,0	25,1	25,0	-2,9	-0,1	-3,0
Region VI	24,4	23,7	23,7	-0,6	-0,1	-0,7
Region VII	30,8	27,7	28,1	-3,1	0,4	-2,7
Region VIII	28,7	27,7	27,2	-1,0	-0,5	-1,5
Region IX	34,4	32,7	28,4	-1,7	-4,3	-6,0
Region X	31,0	30,6	30,3	-0,4	-0,2	-0,7
Region XI	28,4	27,9	27,1	-0,5	-0,8	-1,3
Region XII	27,8	27,5	31,1	-0,3	3,6	3,3
Caraga	31,9	32,4	27,4	0,5	-5,0	-4,5
ARMM	23,9	22,1	26,9	-1,8	4,9	3,0

Source: National Statistical Coordination Board

Poverty gap

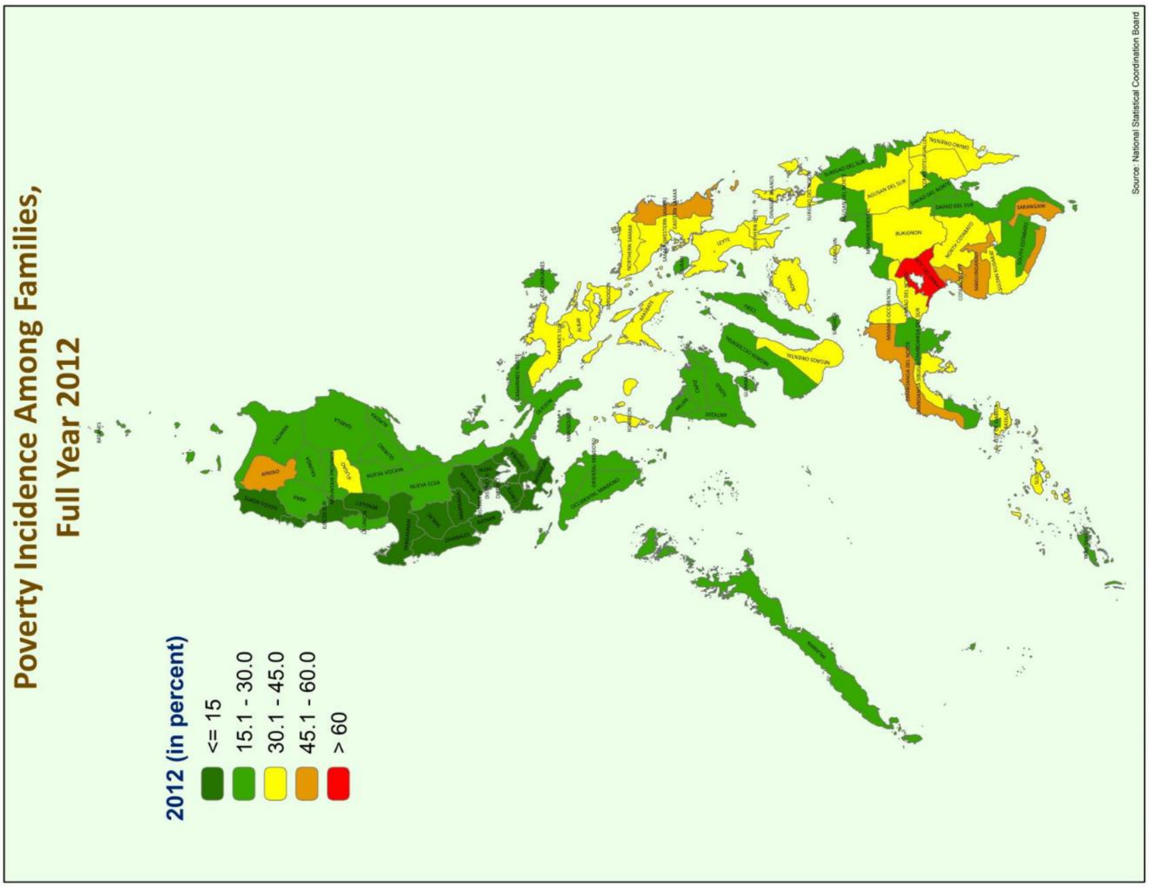
Region	Poverty Gap			Increase/Decrease		
	2006	2009	2012	20006-2009	2009-2012	2006-2012
PHILIPPINES	***5,8***	***5,4***	***5,1***	***-0,4***	***-0,2***	***-0,7***
NCR	0,5	0,4	0,5	-0,1	0,1	0,0
CAR	6,2	5,5	4,7	-0,7	-0,8	-1,5
Region I	4,6	3,8	3,1	-0,8	-0,8	-1,5
Region II	5,0	4,7	3,6	-0,3	-1,1	-1,4
Region III	2,2	2,3	2,2	0,1	-0,1	0,0
Region IV-A	1,8	1,7	1,8	-0,1	0,1	0,0
Region IV-B	9,3	6,9	6,1	-2,4	-0,8	-3,2
Region V	9,9	8,9	8,1	-1,0	-0,8	-1,8
Region VI	5,5	5,6	5,4	0,1	-0,2	-0,1
Region VII	9,4	7,2	7,2	-2,2	0,0	-2,2
Region VIII	9,7	9,5	10,2	-0,1	0,6	0,5
Region IX	13,8	12,9	9,6	-0,8	-3,4	-4,2
Region X	9,9	10,2	9,9	0,2	-0,2	0,0
Region XI	7,2	7,1	6,8	-0,1	-0,4	-0,4
Region XII	8,7	8,5	11,5	-0,2	3,0	2,8
Caraga	13,3	14,9	8,7	1,6	-6,2	-4,6
ARMM	9,7	8,8	13,1	-0,9	4,3	3,4

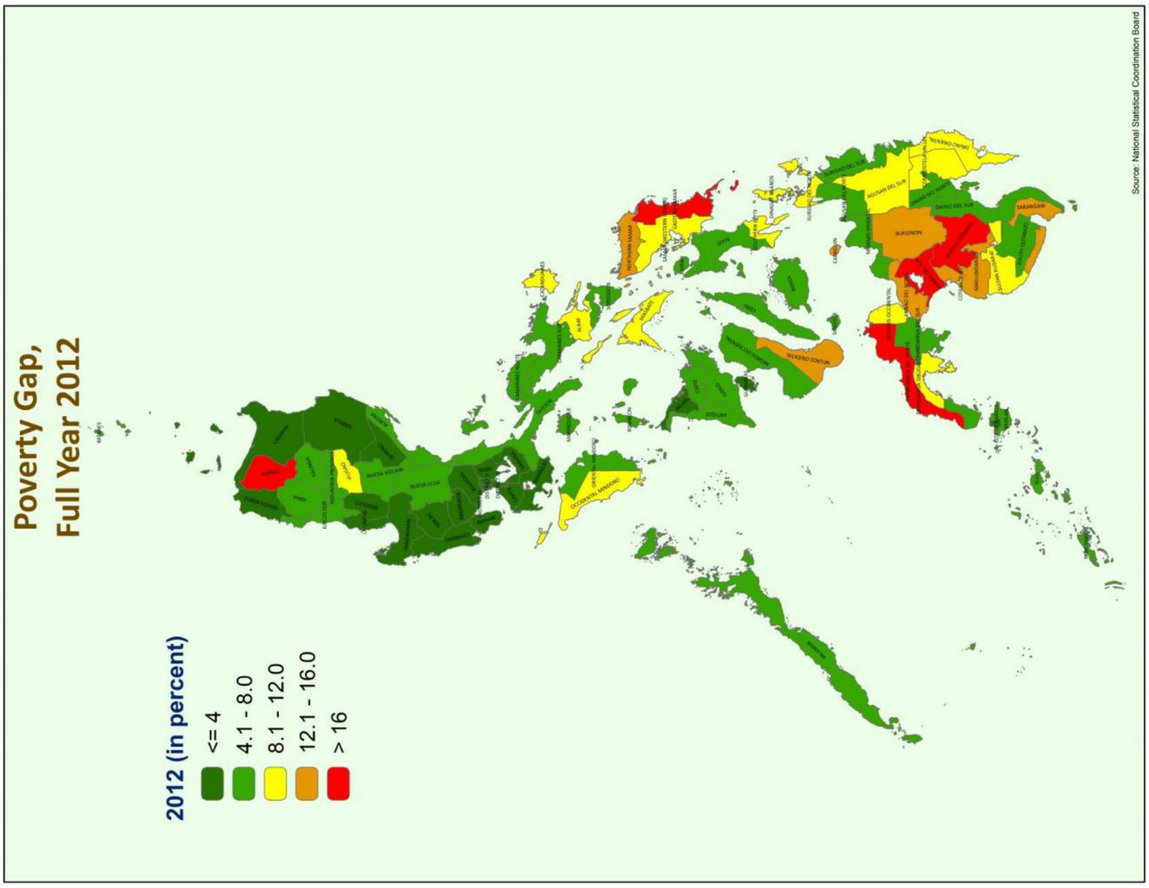
Source: National Statistical Coordination Board

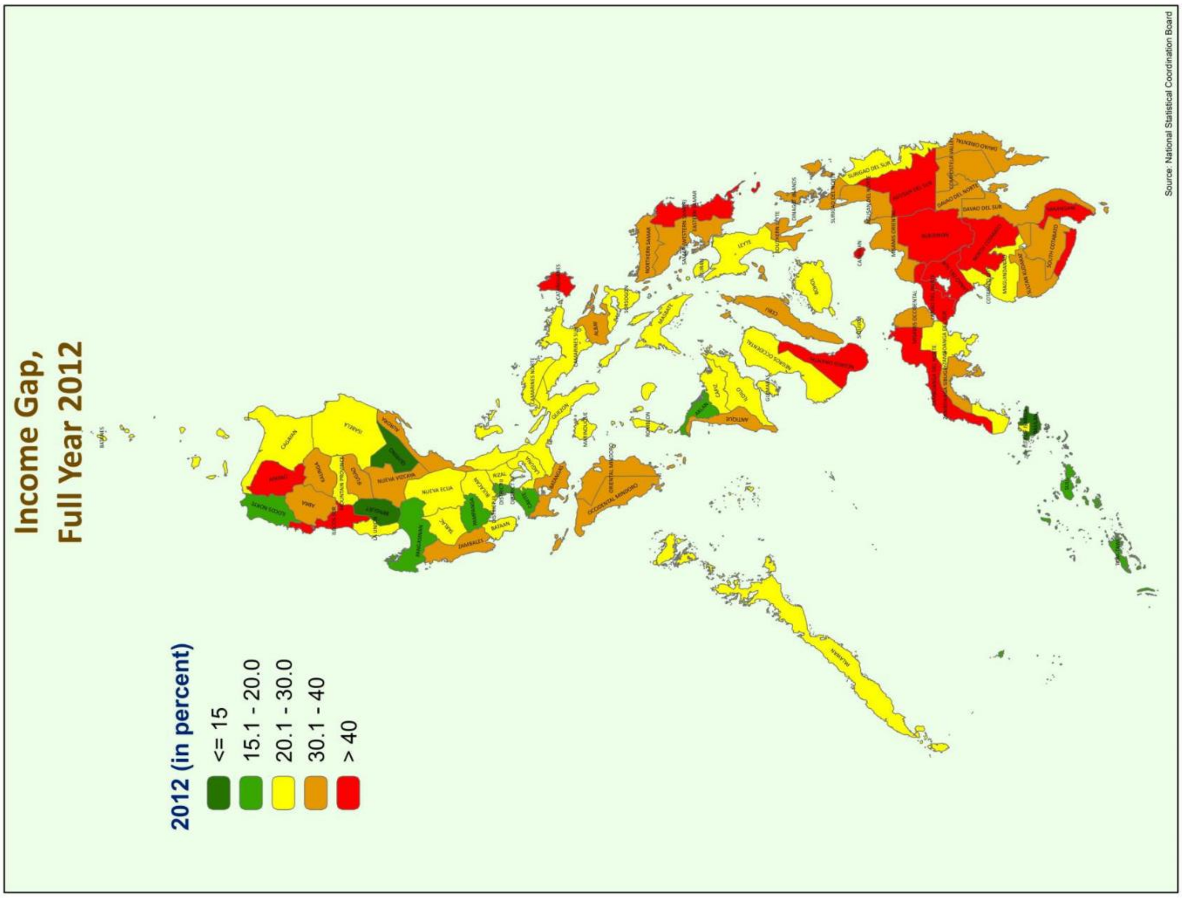
Severity of poverty

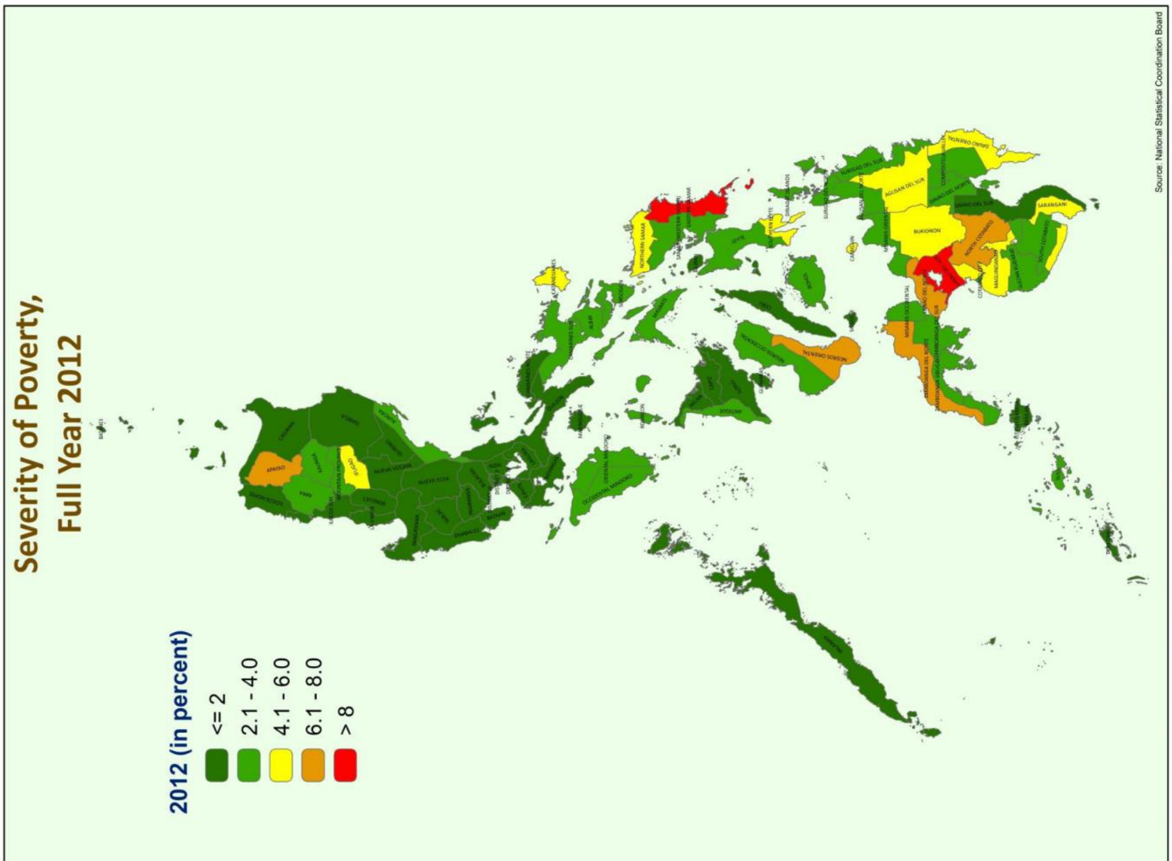
Region	Severity of Poverty			Increase/Decrease		
	2006	2009	2012	20006-2009	2009-2012	2006-2012
PHILIPPINES	***2,2***	***2,0***	***1,9***	***-0,2***	***-0,1***	***-0,3***
NCR	0,2	0,1	0,1	-0,1	0,0	-0,1
CAR	2,5	2,2	1,8	-0,3	-0,4	-0,7
Region I	1,6	1,3	1,0	-0,3	-0,3	-0,6
Region II	1,7	1,6	1,2	-0,1	-0,4	-0,5
Region III	0,7	0,8	0,7	0,1	-0,1	0,0
Region IV-A	0,6	0,5	0,6	-0,1	0,1	0,0
Region IV-B	3,7	2,6	2,3	-1,1	-0,3	-1,4
Region V	3,8	3,2	2,9	-0,6	-0,3	-0,9
Region VI	1,9	1,9	1,9	0,0	0,0	0,0
Region VII	4,0	2,8	2,9	-1,1	0,0	-1,1
Region VIII	3,9	3,7	3,9	-0,3	0,2	0,0
Region IX	6,5	5,7	3,8	-0,7	-1,9	-2,7
Region X	4,1	4,3	4,1	0,2	-0,3	0,0
Region XI	2,8	2,9	2,6	0,1	-0,3	-0,2
Region XII	3,3	3,2	4,9	-0,1	1,7	1,6
Caraga	5,8	6,5	3,4	0,7	-3,1	-2,4
ARMM	3,2	2,8	5,0	-0,4	2,2	1,8

Source: National Statistical Coordination Board


Poverty Incidence Among Families,
Full Year 2012
2012 (in percent)
<= 15
15.1 - 30.0
30.1 - 45.0
45.1 - 60.0
> 60
Source: National Statistical Coordination Board


Poverty Gap,
Full Year 2012
2012 (in percent)
<= 4
4.1 - 8.0
8.1 - 12.0
12.1 - 16.0
> 16
Source: National Statistical Coordination Board


Income Gap,
Full Year 2012
2012 (in percent)
<= 15
15.1 - 20.0
20.1 - 30.0
30.1 - 40
> 40
Source: National Statistical Coordination Board


Severity of Poverty,
Full Year 2012
2012 (in percent)
<= 2
2.1 - 4.0
4.1 - 6.0
6.1 - 8.0
> 8
Source: National Statistical Coordination Board

5. SMALL AREA ESTIMATION OF POVERTY AND INEQUALITY IN THE PHILIPPINES

This chapter is relative to the analysis of small area estimation. Firstly, we take the main direct measures concerning the main poverty indicators. They are the head count ratio, poverty gap, severity gap and income gap. The food threshold, income and expenditure are added to this analysis. Then, the EBLUP model is used to obtain estimates that are more efficient. In order to do that, it is necessary to collect covariates of macroeconomic indexes. Most of these variables are carried out from National Statistical Office (NSO), Census of Population and Housing (CPH), while the others rely on the nature of their data.
This study is presented into subdivisions of 17 regions in the Philippines and they are introduced to the table 5.1. National Capital Region (NCR) has an important role to the overall economic because it is the commerce and industry center of the country. For more information, all the regions should be divided between urban and rural area because there is a significant change in poverty measure and economic development. The EBLUP estimates are calculated by region, after that the coefficient gamma has been computed. It goes on the calculation of their standards errors, the ratio of the estimates and the ratio of their mean squares error. Finally, some figures will be shown to give the outcome's interpretation.

5.1. COVARIATES

The covariates used in this analysis are in totally nine. Each source of them is listed below.

- GDP at current prices is conducted by National Statistical Coordination Board in 2012 and it is expressed in thousand Pesos, by Region.

- Activity Rate (AR) and Unemployment Rate (UR): National Statistics Office, October 2012 Labor Force Survey.

$$\text{AR} = \frac{\text{Labor force}}{\text{Working age population}} * 100$$
$$\text{UR} = \frac{\text{Unemployed persons}}{\text{Labor Force}} * 100$$

where:

- working age population are the household members who have at least 15 years old.
- labor force is somebody who is economically active and it includes employed and unemployed people.
- unemployed persons are characterized by these three factors: somebody who has any job, looking for it and available to start to work.

Activity rate is the proportion of the population in the active labor force over the working age people, whereas unemployment rate represents the proportion of unemployed persons over the labor force.

- Urbanization: Census of Population and Housing, 2010.

- Population density: National Statistics Office, 2000 and 2010 Census of Population and Housing; Land Management Bureau, 2010 and 2007 Masterlist Land Area of the Philippines.

- Infant Mortality Rate (IMR): Philippine Health Statistics (PHS), number and rate/1000 live births, 2012

$$IMR = \frac{\text{Total no. of deaths under 1 year of age registered in a given calendar year}}{\text{Total no. of registered live births of same calendar year}} \times 1{,}000$$

- HH size: Based on the 2010 Census of Population and Housing

- Youths and Elderly people: NSO, Census of Population and Housing, 2010
  The age of youths is within 0-14 years old, while the elder is over than 64 years old.

5.2. Sub-domain of the Philippines by Region

Philippine Standard Geographic Codes of 17 regions are the following:

1	National Capital Region
2	Cordillera Administrative Region
3	Region I – Ilocos
4	Region II - Cagayan Valley
5	Region III- Central Luzon
6	Region IVA - CALABARZON
7	Region IVB- MIMAROPA
8	Region V – Bicol
9	Region VI - Western Visayas
10	Region VII - Central Visayas
11	Region VIII - Eastern Visayas
12	Region IX - Zamboanga Peninsula
13	Region X - Northern Mindanao
14	Region XI - Davao
15	Region XII - SOCCSKSARGEN
16	Region XIII - CARAGA
17	ARMM

Table 5.1: Region classification of the Philippines

There have been changes nowadays because it becomes 18 regions, including in 2015 Negros Is land Region (NIR).

5.3. DIRECT ESTIMATORS

Poverty Incidence among Families is the proportion of the overall family's poverty, based on the absolute poverty's definition and poverty line defined by region.
Income Gap indicates how much percentage of poor's income is below to the poverty line. Its

dominator is the number of poor families or individuals.
Poverty Gap is the proportion of poor's income respect to the total population.
Severity Poverty expresses the mean distance from the poverty line and it is calculated to the total population.
Average annual income is based on the revenue of all kind of jobs.
Average annual expenditure is the amount in order to meet the basic needs. From this information, it is possible to make an analysis of which goods are bought and see whether a person is well off or not.
Food threshold is the basic food needs to have enough nutrition.

5.4. EBLUP ESTIMATOR

Estimator EBLUP is the new estimates from the combination of the covariates defined before and the direct one in order to have more efficient estimates and to lessen the original MSE.
The next step is to verify whether there is some collinearity among covariates and to precede to the calculation of the coefficient gamma, which is needed to the EBLUP estimation. This latter coefficient varies between 0 and 1. The smaller it is, the more efficient is the EBLUP estimator.
Ratio estimator is computed by the Estimator of EBLUP over the direct estimator
Ratio MSE is calculated by the s.e. of Estimator EBLUP over the s.e. of direct estimator

5.5. Analysis of the direct and EBLUP estimates

Poverty Incidence among Families

Region	Poverty Incidence a mong Families	Standard Error	Estimator Eblup	s.e.	gamma	Ratio estimator	Ratio MSE
1	2.6	0.315	2.61	0.314	0.997	1.002	1.000
2	17.5	2.188	18.40	2.111	0.876	1.051	0.965
3	14	1.218	14.06	1.209	0.958	1.004	0.993
4	17	1.598	17.89	1.579	0.929	1.052	0.988
5	10.1	0.859	10.12	0.855	0.979	1.002	0.996
6	8.3	0.730	8.35	0.728	0.984	1.005	0.997
7	23.6	2.478	24.05	2.356	0.846	1.019	0.951
8	32.3	1.809	32.29	1.766	0.911	1.000	0.976
9	22.8	1.459	22.70	1.434	0.941	0.995	0.982
10	25.7	1.619	25.42	1.583	0.928	0.989	0.978
11	37.4	1.907	36.34	1.851	0.902	0.972	0.971
12	33.7	2.123	33.55	2.049	0.882	0.995	0.965
13	32.8	2.427	32.31	2.354	0.851	0.985	0.970
14	25	2.300	24.59	2.219	0.864	0.984	0.965
15	37.1	2.412	35.73	2.305	0.853	0.963	0.956
16	31.9	1.850	31.85	1.798	0.908	0.998	0.972
17	48.7	3.166	49.57	3.192	0.771	1.018	1.008

Table 5.2: Poverty Incidence among Families and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

In the poverty incidence among families, only the Region IVB (MIMAROPA) has gained in terms of MSE of 4.9%. The other regions have maintained nearly the same of the direct value and it is also confirmed by the subsequent graphics below (figure 5.1 and 5.2). The first graphic shows the distance between the direct estimates and the EBLUP one, where they are very similar. The second graphic consists of the efficiency of the estimates and there are very little adjustments from the original data. The area, which is divided by region, is sorted by decreasing number of households. Given this information, we would like to compare all the poverty indicators, without using the sample size. Because this latter data is proportional to the increase of households, it is supposed that the MSE will be decreasing when it goes up. Consequently, the second figure shows this trend and it decreases on the initial areas where the number of households is high.

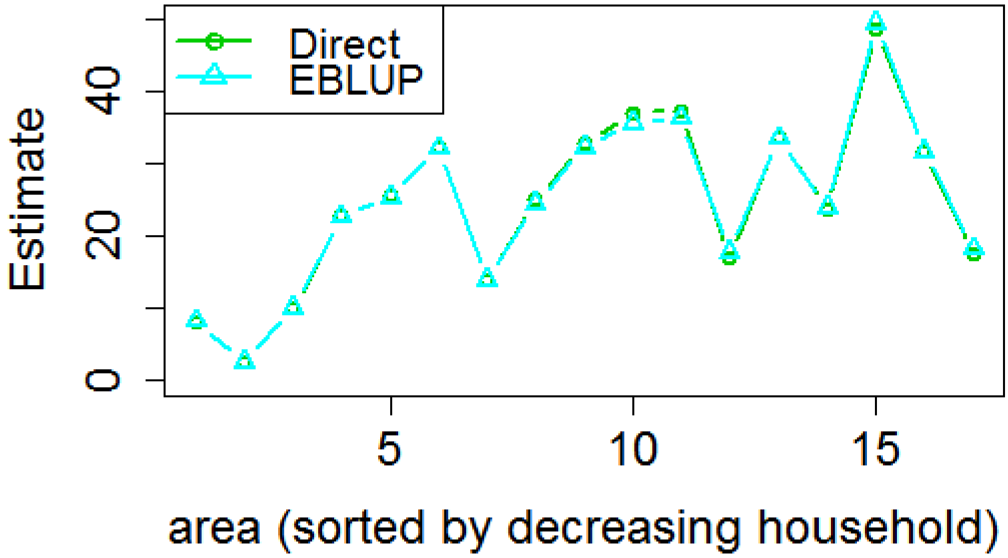


Figure 5.1: The trend of the direct and EBLUP estimates in the poverty incidence among families
Source: Personal elaboration

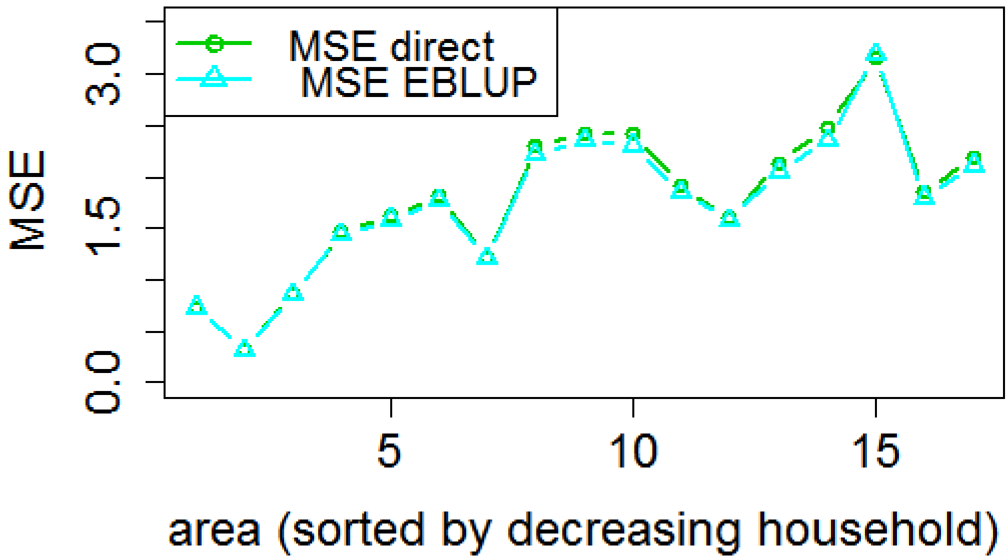


Figure 5.2: The trend of direct and EBLUP MSE in the poverty incidence among families
Source: Personal elaboration

Income Gap

Region	Income Gap	Standard Error	Estimator Eblup	s.e.	gamma	Ratio estimator	Ratio MSE
1	17.2	1.5	18.25	1.439	0.716	1.061	0.959
2	26.9	1.6	26.95	1.502	0.689	1.002	0.939
3	21.9	1.2	21.89	1.160	0.798	1.000	0.967
4	21.5	1.6	23.50	1.488	0.689	1.093	0.930
5	21.4	0.8	21.34	0.788	0.899	0.997	0.985
6	22.1	1.2	22.16	1.153	0.798	1.003	0.961
7	26	1.3	26.27	1.216	0.771	1.010	0.935
8	25	0.8	25.24	0.787	0.899	1.010	0.984
9	23.7	0.9	23.71	0.883	0.875	1.000	0.981
10	28.1	1	27.59	0.964	0.850	0.982	0.964
11	27.2	1	26.91	0.960	0.850	0.989	0.960
12	28.4	1.1	28.33	1.049	0.824	0.998	0.954
13	30.3	1.1	29.98	1.071	0.824	0.990	0.974
14	27.1	1.2	26.68	1.148	0.798	0.985	0.957
15	31.1	1.4	30.25	1.293	0.743	0.973	0.924
16	27.4	1	27.37	0.962	0.850	0.999	0.962
17	26.9	1.7	27.23	1.725	0.663	1.012	1.014

Table 5.3: Income Gap and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

The gain in terms of MSE are located in the table shown below.

Region	Ratio MSE
National Capital Region	4.1%
Cordillera Administrative Region	6.15%
Region IX - Zamboanga Peninsula	6.98%
Region XI - Davao	6.45%
Region XII - SOCCSKSARGEN	7.64%

As illustrated at the first result of poverty indicator, here the gain of efficiency is a little higher than the former. Nonetheless, it remains little adjustments from the initial data. The trend in figure 5.3 shows how these two estimates are quite near to each other and so the results to the trend of ratio MSE.

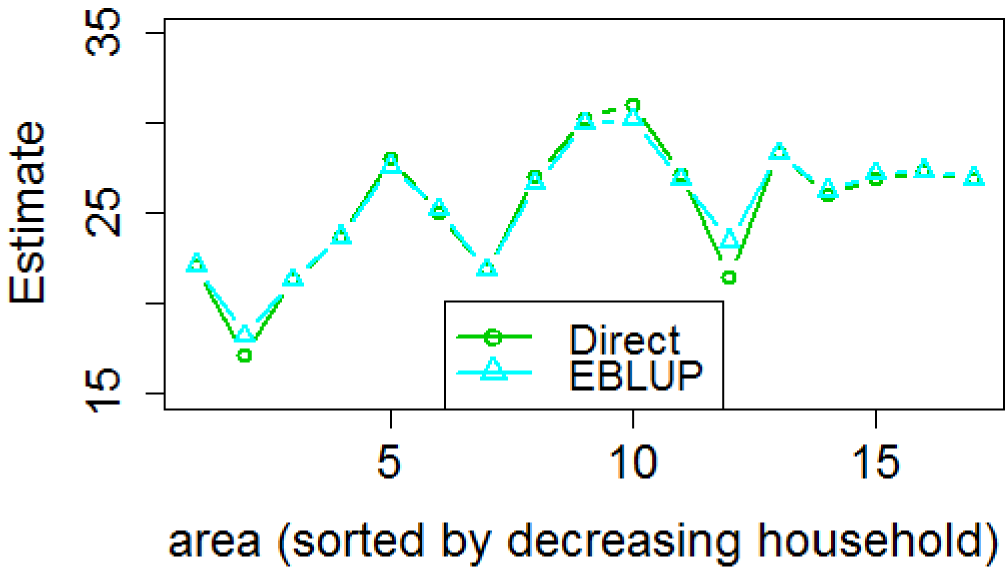


Figure 5.3: The trend of the direct and EBLUP estimates in the income gap
Source: Personal elaboration

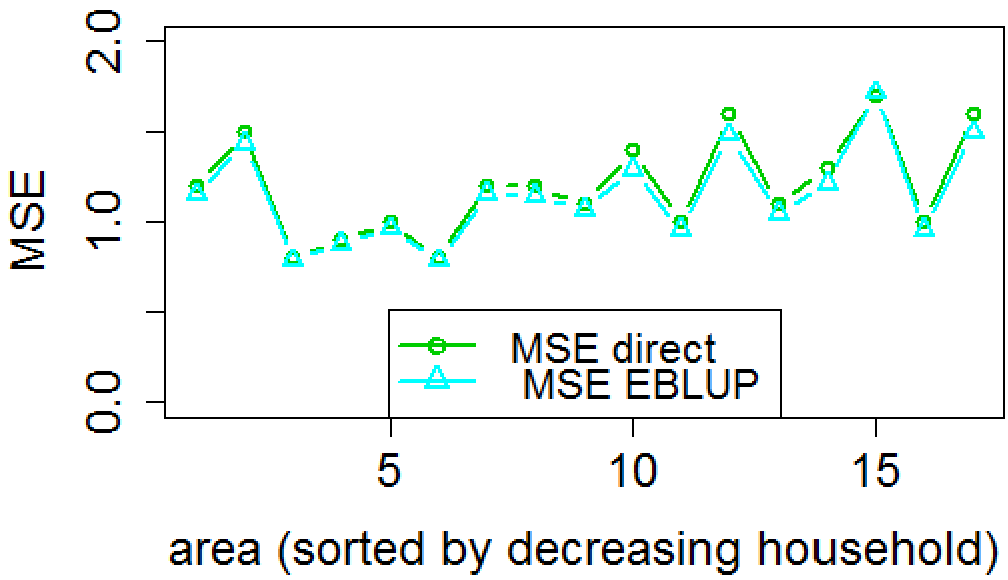


Figure 5.4: The trend of direct and EBLUP MSE in the income gap
Source: Personal elaboration

Poverty Gap

Region	Poverty Gap	Standard Error	Estimator Eblup	s.e.	gamma	Ratio estimator	Ratio MSE
1	0.5	0.1	0.50	0.100	0.998	1.000	1.000
2	4.7	0.7	4.91	0.697	0.907	1.045	0.995
3	3.1	0.4	3.11	0.399	0.968	1.004	0.997
4	3.6	0.5	3.81	0.498	0.950	1.058	0.996
5	2.2	0.2	2.20	0.200	0.992	1.001	0.999
6	1.8	0.2	1.81	0.200	0.992	1.006	0.999
7	6.1	0.8	6.22	0.776	0.882	1.020	0.970
8	8.1	0.5	8.15	0.496	0.950	1.006	0.992
9	5.4	0.5	5.40	0.496	0.950	1.001	0.992
10	7.2	0.6	7.11	0.590	0.930	0.987	0.983
11	10.2	0.8	9.81	0.773	0.882	0.961	0.967
12	9.6	0.8	9.51	0.775	0.882	0.990	0.969
13	9.9	0.9	9.71	0.878	0.855	0.981	0.976
14	6.8	0.8	6.64	0.780	0.882	0.977	0.975
15	11.5	1	10.83	0.950	0.827	0.942	0.950
16	8.7	0.7	8.68	0.683	0.907	0.997	0.976
17	13.1	1.5	13.41	1.515	0.680	1.024	1.010

Table 5.4: Poverty Gap and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

Among all regions, only the Region XII (SOCCSKSARGEN) has gained efficiency of 5% while the others are similar to the direct estimates. These results are shown in the figure 5.5. Again, the trend of MSE is very similar between these two estimates.

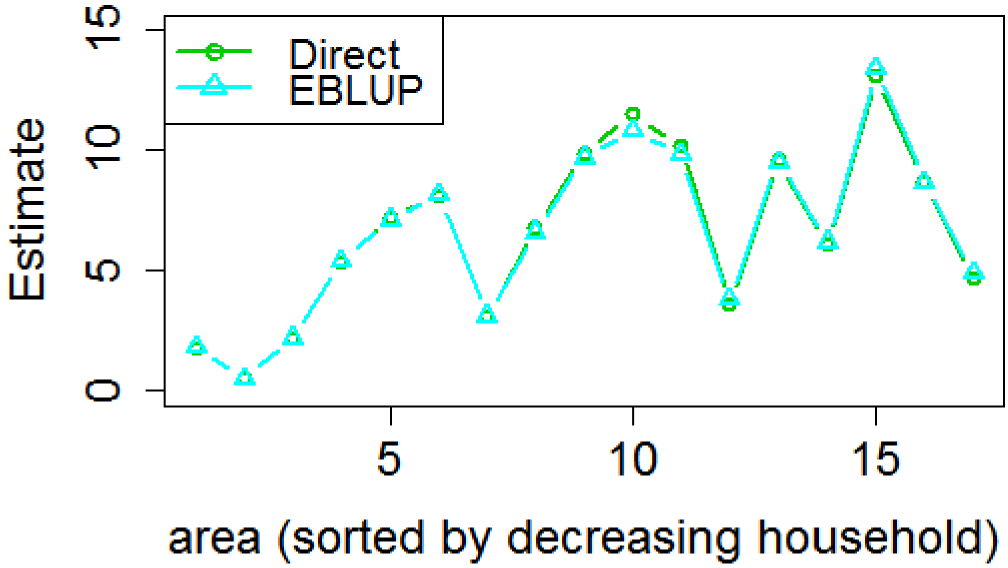


Figure 5.5: The trend of the direct and EBLUP estimates in the poverty gap
Source: Personal elaboration

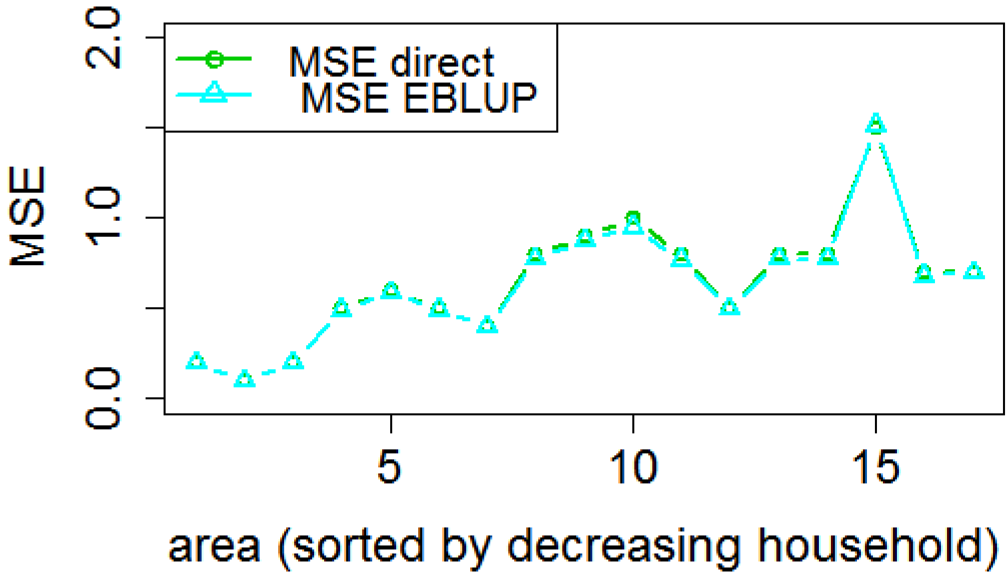


Figure 5.6: The trend of direct and EBLUP MSE in the poverty gap
Source: Personal elaboration

Severity Poverty

Region	Severity Poverty	Standard Error	Estimator Eblup	s.e.	gamma	Ratio estimator	Ratio MSE
1	0.1	0	0.10	0.000	1.000	1.000	-
2	1.8	0.3	1.90	0.293	0.895	1.055	0.976
3	1	0.2	1.01	0.199	0.951	1.008	0.993
4	1.2	0.3	1.39	0.296	0.895	1.156	0.986
5	0.7	0.1	0.70	0.100	0.987	1.002	0.999
6	0.6	0.1	0.61	0.100	0.987	1.011	0.999
7	2.3	0.3	2.35	0.292	0.895	1.021	0.972
8	2.9	0.2	2.93	0.198	0.951	1.010	0.989
9	1.9	0.2	1.91	0.197	0.951	1.003	0.987
10	2.9	0.3	2.83	0.291	0.895	0.975	0.970
11	3.9	0.4	3.67	0.379	0.828	0.941	0.947
12	3.8	0.4	3.73	0.380	0.828	0.982	0.950
13	4.1	0.5	3.93	0.473	0.755	0.959	0.946
14	2.6	0.4	2.50	0.383	0.828	0.961	0.957
15	4.9	0.6	4.28	0.531	0.681	0.874	0.886
16	3.4	0.3	3.38	0.291	0.895	0.995	0.971
17	5	0.8	5.18	0.793	0.546	1.037	0.991

Table 5.5: Severity of poverty and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

In the severity of poverty, similar low increase results are obtained: Region VIII (Eastern Visayas) has gained of 5.30%, Region IX (Zamboanga Peninsula) of 5% and Region X (Northern Mindanao) of 5.39%. Interestingly, the Region XI (Davao) has gained of 11.43%. They are shown in figure 5.7, while the subsequent figure illustrates the trend of their MSE, which are also very close between the two estimates.

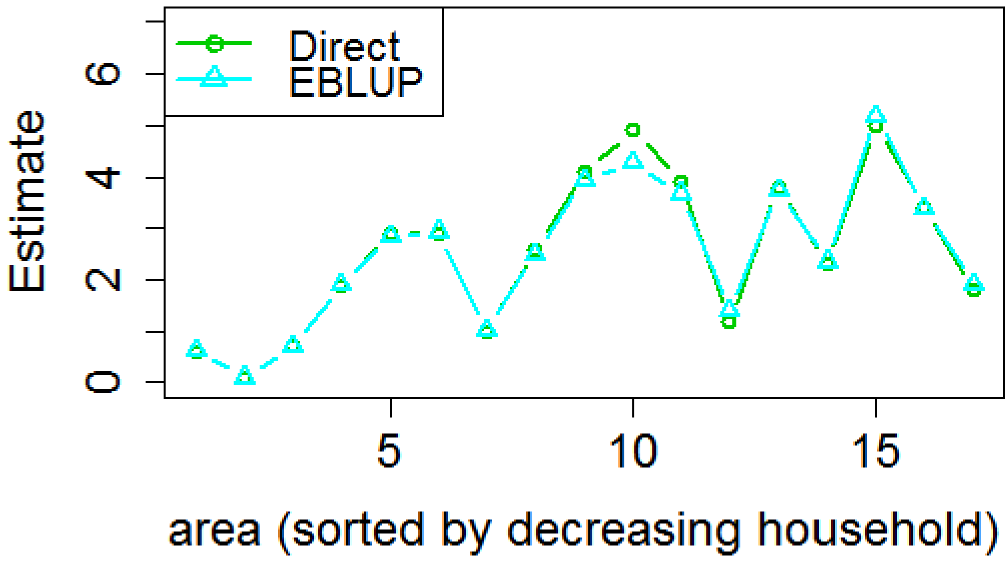


Figure 5.7: The trend of the direct and EBLUP estimates in the severity of poverty
Source: Personal elaboration

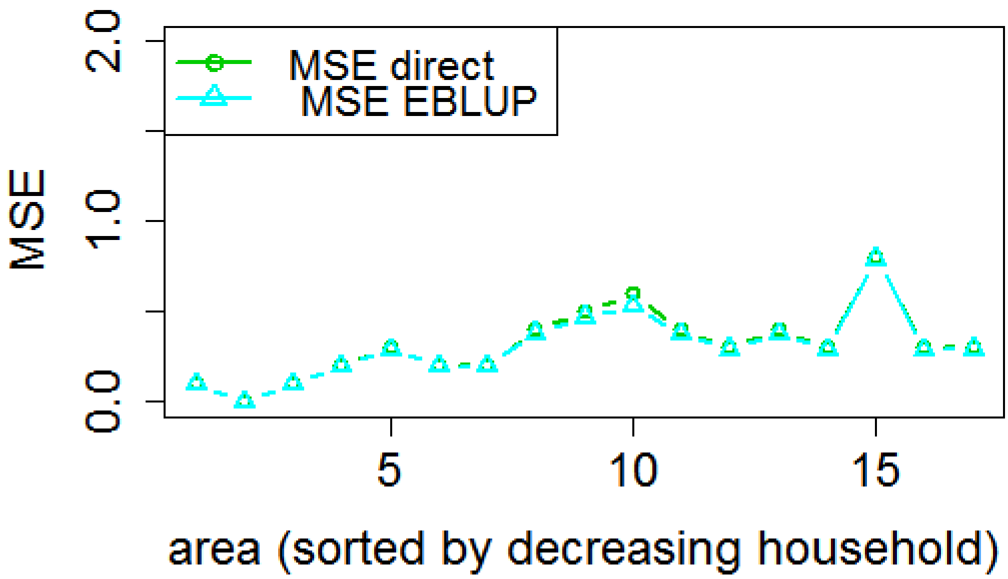


Figure 5.8: The trend of direct and EBLUP MSE in the severity of poverty
Source: Personal elaboration

Average annual income

Region	Average annual income	Standard Error	Estimator Eblup	s.e.	gamma	Ratio estimator	Ratio MSE
1	379,195	9,107	374,625.6	9,008.498	0.887	0.988	0.989
2	256,893	15,029	253,852.7	14,206.868	0.742	0.988	0.945
3	203,707	70,28	206,010.6	6,958.577	0.929	1.011	0.990
4	194,650	5,802	191,882.2	5,763.706	0.951	0.986	0.993
5	259,400	7,909	260,239.4	7,775.197	0.912	1.003	0.983
6	284,189	15,402	282,018.1	14,291.068	0.733	0.992	0.928
7	178,847	8,528	180,367.8	8,304.904	0.899	1.009	0.974
8	162,376	5,643	163,038.7	5,595.980	0.953	1.004	0.992
9	201,978	7,344	201,684.2	7,248.056	0.923	0.999	0.987
10	208,846	7,255	209,092.3	7,121.614	0.925	1.001	0.982
11	165,754	8,320	167,493.8	8,091.716	0.904	1.010	0.973
12	161,861	6,205	161,728.6	6,117.712	0.944	0.999	0.986
13	189,525	9,307	192,114.7	9,125.251	0.882	1.014	0.980
14	194,298	8,010	194,195.4	7,862.300	0.910	0.999	0.982
15	162,632	8,367	164,914.4	8,155.292	0.903	1.014	0.975
16	180,115	7,084	180,199.1	6,949.333	0.928	1.000	0.981
17	129,576	4,661	129,025.1	4,681.731	0.968	0.996	1.004

Table 5.6: Average annual income and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

In the average annual income, the gain in terms of MSE is obtained in the Cordillera Administrative Region of 5.47% and in the Region IVA (CALABARZON) of 7.21%. Figure 5.9 illustrates the estimates of the direct and EBLUP while figure 5.10 shows the trend of their MSE. The result is that the direct estimates are very close to EBLUP and their MSE is similar to the initial data.

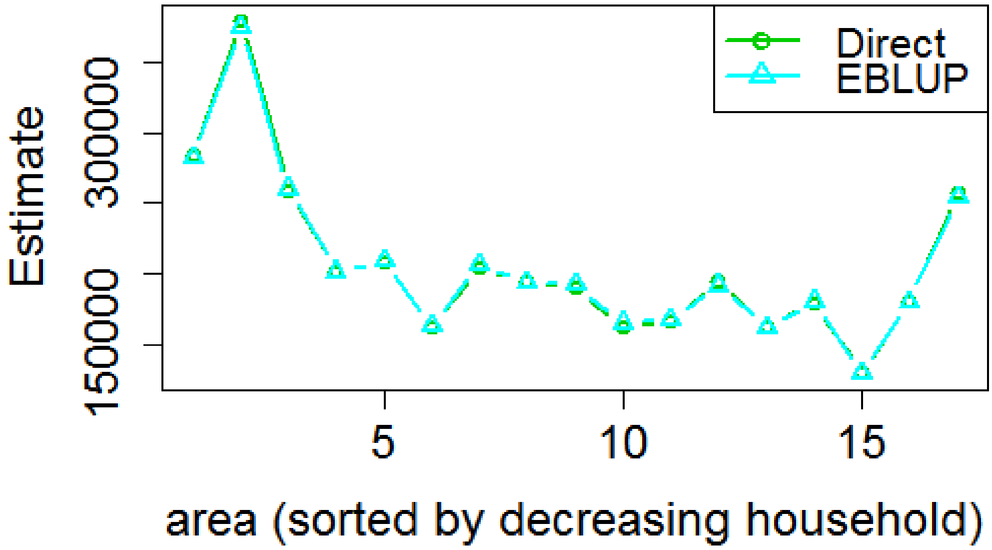


Figure 5.9: The trend of the direct and EBLUP estimates in the average annual income
Source: Personal elaboration

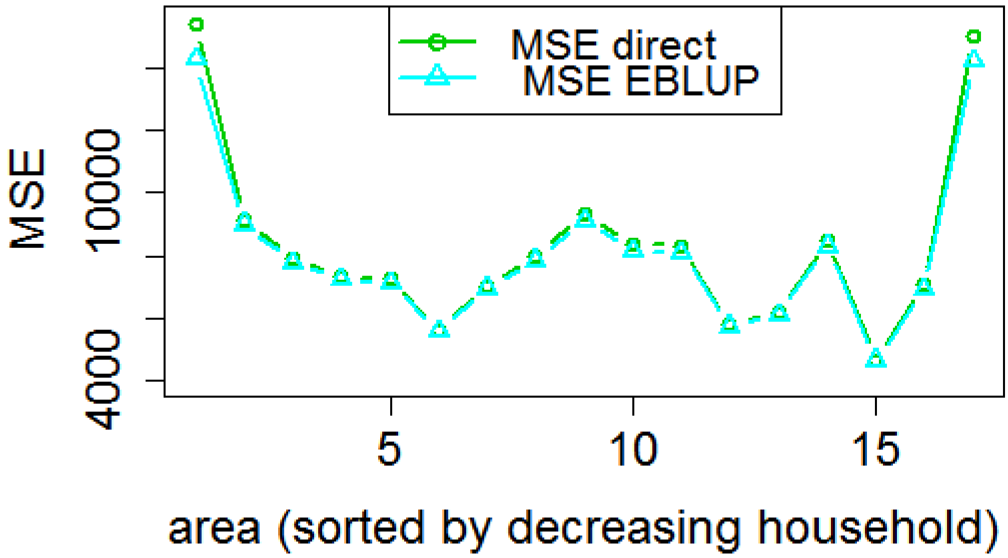


Figure 5.10: The trend of direct and EBLUP MSE in the average annual income
Source: Personal elaboration

Average annual expenditure

Region	Average annual expenditure	Standard Error	Estimator Eblup	s.e.	gamma	Ratio estimator	Ratio MSE
1	324,837	8,180	318,616.6	8,006.497	0.880	0.981	0.979
2	188,323	8,091	186,689.4	7,966.465	0.882	0.991	0.985
3	158,603	4,499	159,535.7	4,472.220	0.960	1.006	0.994
4	139,914	3,752	138,972.8	3,729.079	0.972	0.993	0.994
5	211,242	5,108	212,079.8	5,052.984	0.949	1.004	0.989
6	242,704	9,246	242,719.4	9,029.486	0.851	1.000	0.977
7	138,252	5,920	138,532.5	5,805.089	0.933	1.002	0.981
8	143,554	4,494	144,002.8	4,445.880	0.960	1.003	0.989
9	162,672	5,090	163,402.2	5,012.629	0.950	1.004	0.985
10	163,972	5,392	165,020.3	5,298.367	0.944	1.006	0.983
11	131,942	5,900	132,929.5	5,779.150	0.934	1.007	0.980
12	121,592	4,341	121,677.2	4,297.587	0.963	1.001	0.990
13	142,750	6,451	144,789.6	6,324.335	0.922	1.014	0.980
14	155,556	5,720	154,833.1	5,607.792	0.937	0.995	0.980
15	139,772	5,450	139,884.9	5,366.499	0.943	1.001	0.985
16	141,658	4,901	141,351.7	4,837.114	0.953	0.998	0.987
17	113,980	3,163	113,590.6	3,161.547	0.980	0.997	1.000

Table 5.7: Average annual expenditure and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

In the average annual expenditure, the gain in MSE is very little. All the region have reached less than 4%. Consequently, the trend of the direct estimates and the EBLUP one is very close to each other and it is shown in figure 5.11. Their MSE tendency is illustrated in figure 5.12, where a little gain of efficiency is obtained.

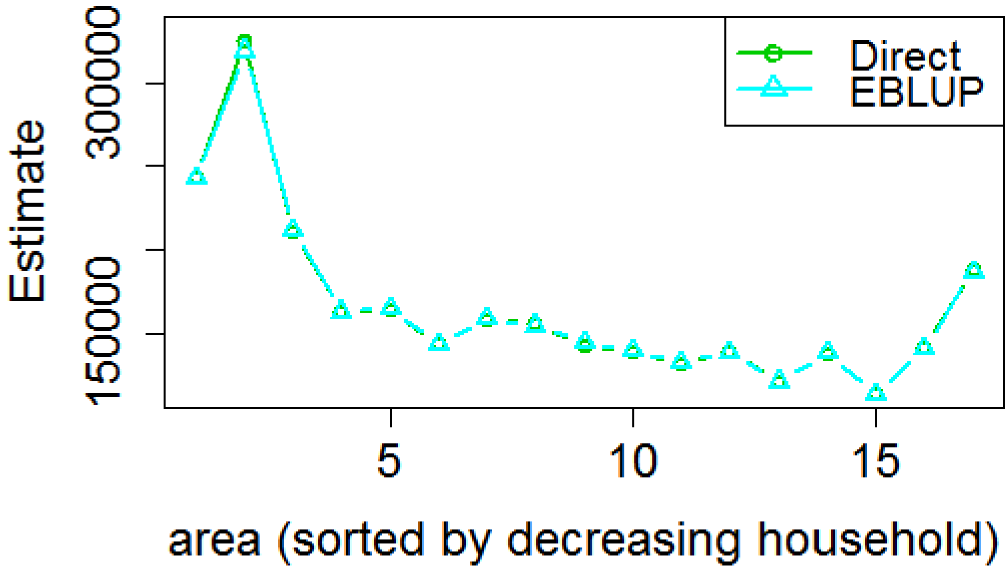


Figure 5.11: The trend of the direct and EBLUP estimates in the average annual expenditure
Source: Personal elaboration

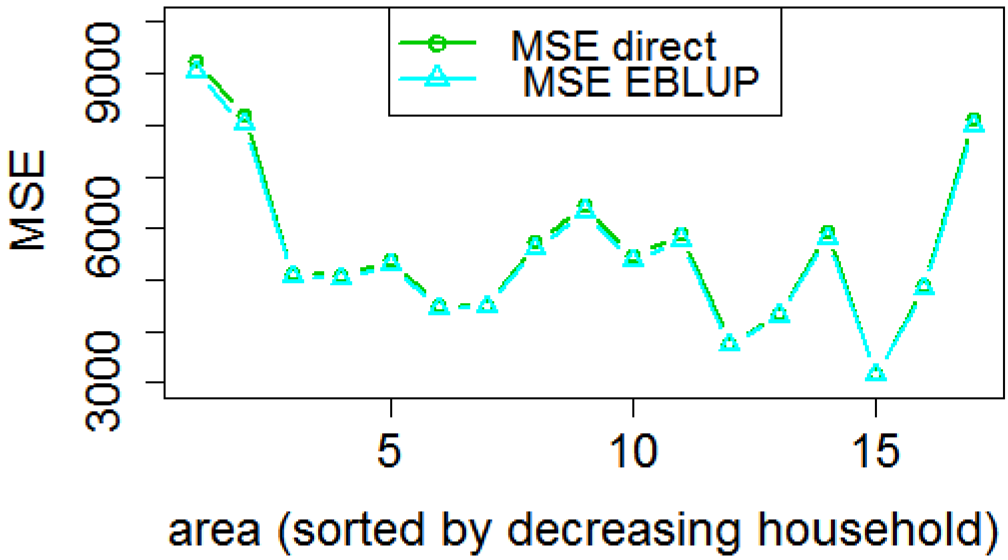


Figure 5.12: The trend of direct and EBLUP MSE in the average annual expenditure
Source: Personal elaboration

Food threshold

Region	Food	Standard Error	Estimator Eblup	s.e.	gamma	Ratio esti mator	Ratio MSE
1	36.5	0.57	36.49	0.576	0.773	0.999	1.010
2	40.9	0.89	41.62	0.838	0.583	1.019	0.941
3	45.3	0.65	45.35	0.642	0.724	1.001	0.988
4	49.4	0.56	49.34	0.560	0.779	0.998	1.001
5	43.7	0.55	43.44	0.544	0.785	0.994	0.989
6	40.7	0.87	40.40	0.801	0.594	0.994	0.921
7	46.5	1.01	46.21	0.852	0.521	0.995	0.843
8	49.0	0.94	48.54	0.841	0.556	0.990	0.894
9	44.4	0.78	45.16	0.738	0.645	1.017	0.946
10	44.5	0.85	44.74	0.762	0.605	1.006	0.897
11	45.8	1.14	47.15	0.892	0.460	1.030	0.782
12	49.0	1.03	48.90	0.855	0.511	0.998	0.830
13	43.8	1.09	43.92	0.946	0.482	1.002	0.868
14	47.6	0.96	47.24	0.847	0.546	0.992	0.882
15	48.2	1.07	47.07	0.880	0.492	0.976	0.822
16	47.5	1.02	47.90	0.849	0.516	1.008	0.832
17	59.8	1.08	59.40	1.104	0.487	0.994	1.022

Table 5.8: Food threshold and EBLUP estimates
Source: The first and second column are derived from direct estimates, while the others has been computed personally.

The last indicator, food threshold, has more improvements than the others do. In the table below, the gain of MSE is shown below:

Region	Ratio MSE
Cordillera Administrative Region	5.88%
Region IVA - CALABARZON	7.91%
Region IVB- MIMAROPA	15.66%
Region V – Bicol	10.58%
Region VI - Western Visayas	5.44%
Region VII - Central Visayas	10.33%
Region VIII - Eastern Visayas	21.75%
Region IX - Zamboanga Peninsula	17.03%
Region X - Northern Mindanao	13.21%
Region XI - Davao	11.76%
Region XII - SOCCSKSARGEN	17.79%
Region XIII - CARAGA	16.81%

Many regions are more efficient than the direct estimates. Even the trend of their estimates are similar to each other, which is shown in figure 5.13, the MSE tendency shows better results of efficiency. In fact, the line of the EBLUP MSE on the figure 5.14 is almost below to the direct MSE.

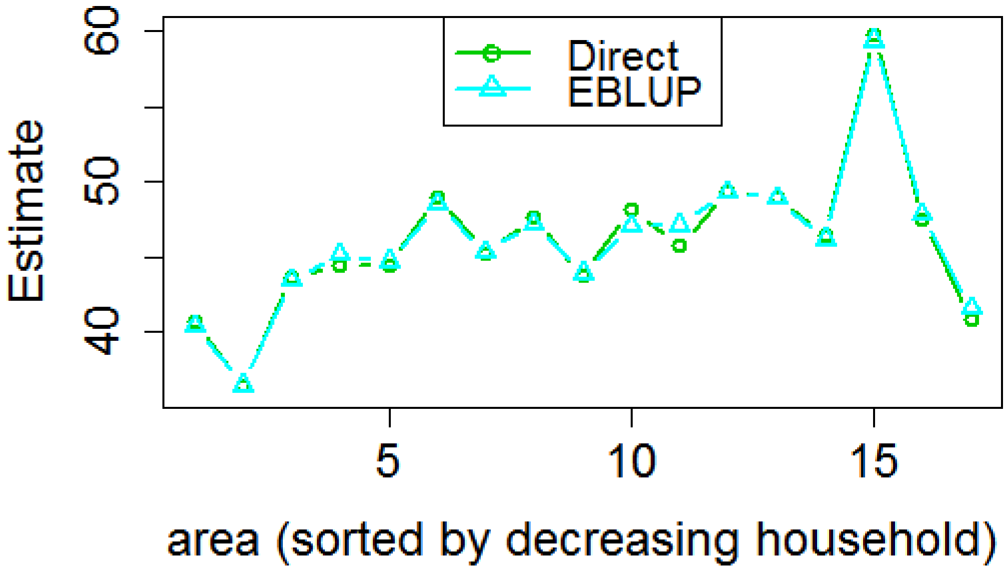


Figure 5.13: The trend of the direct and EBLUP estimates in the Food threshold
Source: Personal elaboration

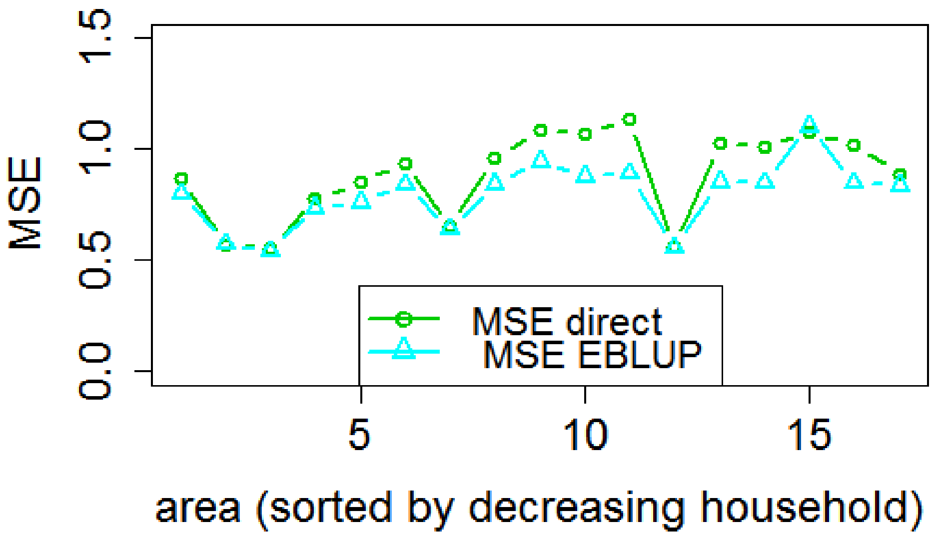


Figure 5.14: The trend of direct and EBLUP MSE in the Food threshold
Source: Personal elaboration

The gamma coefficient is generally less than one so that it produces some adjustments to the direct estimates, even it is a little variation. The Food estimates have some improvements, while the others are very little. It must be also considered that both expenditure and income estimates have initially a big standard error so that any kind of estimation will be not sufficient to reduce it. Nonetheless, it is possible to notice that where the EBLUP tendency in MSE is below to the direct MSE estimates, it shows that the former estimates are always better than the latter one.

6. CONCLUSION

After the analysis has been taken place, we can notice that the EBLUP estimates are very close to that of the direct. The reason could rely on larger sample size utilized on each region, even though it is not calculated in this analysis. Gamma coefficients are smaller than unity so that the EBLUP estimates is showing a good mode. However, the MSE gain is very small, only some regions in particular can be excluded but their interpretation change in which poverty indicator is utilized. Only the Food threshold had some adjustment using the EBLUP model. Because of the importance of this index, we can presume that approximately half of the population has difficulty to meet the basic food needs. Instead of using the sample size, the decreasing size of household is utilized in this analysis in order to see the MSE trend. As a rule of thumb, it decreases as the households are increasing, apart from the average annual expenditure and income indicators. In some cases it also remains constant: Income gap and Food threshold.

To obtain much gain in the MSE, the EBLUP model works much better if the sub-domains are extended to provincial and municipal level. However, it could not be implemented in this analysis as some information are not available.

REFERENCES

▪ Amartya Sen. (1976, March). Poverty: An ordinal approach to measurement, *Econometrica*, Vol. 44, No. 2.

▪ Balicasan Arnesio & Edillon Rosemarie. (2005). *Poverty targeting in the Philippines*, Asian Development Bank Institute, EEP Edition, USA, pp. 219-246.

▪ Balicasan Arnesio. (2001, June). Poverty in the Philippines: An update and reexamination, *Philippine Review of Economics*, Vol. XXXVIII, No. I. Retrieved from http://pre.econ.upd.edu.ph/

▪ Barcenas Marissa L. (2004, October). *The Development of the 2003 Master Sample (MS) for Philippines Household Surveys*. Retrieved from http://nap.psa.gov.ph/

▪ Bayudan-Dacuycuy Coonie & Lim Joseph Anthony. (2014).Chronic and transient poverty and vulnerability to poverty in the Philippines: Evidence using a simple spells approach, Economic Department, *MPRA*, Ateneo de Manila University. Retrieved from https://mpra.ub.uni-muenchen.de/

▪ Bellù Lorenzo Giovanni& Liberati Paolo. (2005, November). Social Welfare Analysis of Income Distributions, Ranking Income Distributions with Generalised Lorenz Curves, *EASYPOL,* Resources for policy making, Module 002, Analytical Tools. Retrieved from http://www.fao.org/

▪ Betti Gianni. (2014, October). *Analisi Statistica del Reddito e delle Condizioni di Vita* [Statistical Analysis of Income and Life's Condition], University of Siena.

▪ Bureau of Labor and Employment Statistics. (2011a). *Learning Statistics the Easy Way, Session on Labor Force Survey.* Retrieved from http://labstat.psa.gov.ph/

▪ Bureau of Labor and Employment Statistics. (2011b). *Session on Poverty Statistics.* Retrieved from http://labstat.psa.gov.ph/

▪ Chaudhury Nazmul, Friedman Jed & Onishi Junko. (2013, January). *Philippines Conditional Cash Transfer Program, Impact Evaluation 2012*, Report Number 75533-PH, Washington, DC: World Bank Group. Retrieved from http://documents.worldbank.org/

▪ Chiappero Martinetti Enrica. (2009). Povertà e benessere multidimensionale: un raffronto fra concetti, misure e politiche pubbliche [Poverty and multidimensional welfare: a comparison among concepts, measurements and public policy], *Quaderni CIES N.2*, Lavoro di ricerca commissionato allo IUSS e preparato per la Commissione di Indagine sull'Esclusione Sociale.

▪ Davis Benjamin. (2003). *Choosing a method for poverty mapping*, FAO, Rome, pp. 5-28. Retrieved from http://www.fao.org

▪ Deichmann Uwe. (1999, January). *Geographic aspects of inequality and poverty*, Working Paper No. 76510, Vol. No. 1, Washington DC: World Bank. Retrieved from http://siteresources.worldbank.org/

▪ Délice Williams. (2013, June). *What are causes of Global Poverty*. retrieved from http://borgenproject.org/

▪ Demombynes Gabriel, Elbers Chris, Lanjoux Jean O., Lanjoux Peter, Mistiaien Johan & Özler Berk. (2002, April). *Producing an Improved Geographic Profile of Poverty*, WIDER, Discussion Paper No. 2002/39, United Nations University. Retrieved from https://www.wider.unu.edu

▪ Grassi Donatella & Pannuzi Nicoletta. (2009). *Misura della povertà assoluta,* [Measurement of the absolute poverty], Metodi e Norme n. 39, ISTAT. Retrieved from http://www3.istat.it/

▪ Haslett Stephen & Jones Geoffrey. (2005, June). *Local Estimation of Poverty in the Philippines*, Manila. Retrieved from http://siteresources.worldbank.org/

▪ ISTAT, *La misura della povertà* [The measurement of poverty], 2009

▪ Lee Katheen. (2014, October). *Poverty in the Philippines*. Retrieved from http://borgenproject.org/

▪ Marichu Fajardo-Duka. (1992, December). Poverty measurement and decomposition of aggregate poverty change in the Philippines, *Philippines Review of Economics and Business*, Volume XXIX, No 2, U. P. School of Economics. Retrieved from http://pre.econ.upd.edu.ph/

▪ Maxwell Simon. (1999, February). The Meaning and Measurement of Poverty, *ODI Poverty Briefing 3.* Retrieved from https://www.odi.org

▪ Minot Nicolas & Baulch Bob. (2002, November). *Poverty mapping with aggregate census data: what is the loss in precision?*, MSSD Discussion Paper No. 49, Washington.

▪ Molina Isabel & Marhuenda Yolanda. (2015, March). *Basic direct and indirect estimators in sae package*.

▪ Mowafi Mona. (n. d.). *The Meaning and Measurement of Poverty: A Look into the Global Debate*.

▪ National Statistical Coordination Board (2012, November), *Small Area Estimates: For Evidence-based Policy-making and Better Targeted/Focused Programs*, 2nd Annual M&E Forum, Manila. Retrieved from http://devplan.neda.gov.ph/

▪ National Statistical Coordination Board. (2013, December). *2012 Full Year Official Poverty Statistics*.

▪ National Statistics Office. (2003). *Master Sample Brochure*.

▪ National Statistics Office. (2012). *Journal of Philippines Statistics - A Quarterly Issue*, Volume 63 number 1, First quarter 2012.

▪ Nazami-Rad Mohammad-Reza & Steel David. (2011). What Level of Statistical of Model Should We Use in Small Domain Estimation?, *CSSM*, Working Paper No. 02-11, University of Wollongong, 2011. Retrieved from http://ro.uow.edu.au/

▪ Neri Laura, Ballini Francesca & Betti Gianni. (2005, January). Poverty and Inequality Mapping in transition countries, *Statistics in Transition*, Working paper n. 52.

▪ Petrucci Alessandra & Pratesi Monica. (2014, March). Spatial Models in Small Area Estimation in the context of official statistics, *Statistica Applicata - Italian Journal of Applied Statistics,* Vol. 24 (1). Retrieved from http://sa-ijas.stat.unipd.it/

▪ Pratesi Monica, Tzadivis Nicola et al. (2010). Final Small Area Estimation Developments and Simulation Results, *SAMPLE*, DELIVERABLES 12 and 16, Working Paper No. 2, pp. 3-68. Retrieved from http://www.sample-project.eu/

▪ Pratesi Monica. (2012, August). *Recent Development in Small Area Estimation Methodology*. Retrieved from http://home.lu.lv/

▪ Rao J. N. K. & Molina Isabel. (n. d.). *Small Area Estimation Methods, Applications and Practical Demonstration*. Retrieved from http://halweb.uc3m.es/

▪ Rao J. N. K. (2003). *Small Area Estimation*, WILEY, Cartelon University.

▪ Ravallion Martin. (1998). Poverty lines in Theory and Practice, *LSMS,* Working Paper No. 133, Washington. Retrieved from http://documents.worldbank.org/

▪ Rosling Hans et al. (2014). *Global Health.*

▪ SAE package developers. (2007, January). *Introduction to Small Area Estimation*. Retrieved from http://www.bias-project.org.uk/

▪ Timbang Theresa D., Sinson Fe A. & Rebanal Lea Mylene R. (n. d.). *The 2012 Philippine Health Statistics*, Department of Health.

▪ Tzavidis Nikos. (2010, May). *What is Poverty Mapping*. Retrieved from http://www.methods.manchester.ac.uk/

▪ Tzavidis, N., Salvati, N., Pratesi, M. & Chambers, R. L. (2008). *M-quantile models with application to poverty mapping. Statistical Methods and Applications*, 17 (3), pp. 393-411.

▪ Vindigni Gabriella, Peri Iuri & Prosperi Paolo. (2011). *Aperte problematiche nell'analisi di povertà questioni di misura e progressi nel raggiungimento degli Obiettivi del Millennio*, [Open problems in analysis of poverty regarding measurement and development of Millenium's goals] Economia & Diritto Agroalimentare XVI: pp. 427-446.

▪ Weiss John. (2005). *Experiences with poverty targeting in Asia: an overview*, Asian Development Bank Institute, EEP Edition, USA, pp. 1-33.

▪ World Bank Institute. (2005, August). *Introduction to Poverty Analysis*, Poverty Manual, All, JH Revision. pp. 69-82; 95-105. Retrieved from http://siteresources.worldbank.org/

▪ Philippines Statistics Authority: https://psa.gov.ph/; http://www.nscb.gov.ph/stattables/