An Econometrics Assessment of Food Security Estimation Using Fuzzy Logics: A Case in the Arid and Semi Arid Lands of Kenya

By Sulo Timothy & Chelangat Sharon

Moi University

Abstract - This paper takes into consideration the severe bottlenecks that have actually bedeviled econometric analysis and documentation of food security since time immemorial. It aims at modeling food security estimation using fuzzy logics. The paper shows econometrically how food security measurement drawbacks are overcome using residual diagnostic analysis by the effects of fuzzy logics on the leverage points of food security predictors. Further, the results indicate that the preliminary econometrics tests on the residual diagnostic analysis on the error variance, co linearity, multicollinearity and mahalanobis distances improved the estimation of food intake (the predicted criterion) because its predictors are stabilized upon data conversion into fuzzy membership functions. To a certain reasonable extent, it may be very safe to conclude that there is something quite positive in econometric research when fuzzy logics are applied in estimating food security, poverty among other similar subjective or qualitative variables.

Keywords : Food Security, Estimation, Fuzzy Logics, Residual Diagnostics, Econometrics of Food Security.

GJSFR-D Classification: FOR Code: 290104p

Strictly as per the compliance and regulations of :

An Econometrics Assessment of Food Security Estimation Using Fuzzy Logics: A Case in the Arid and Semi Arid Lands of Kenya

Sulo Timothy [α] & Chelangat Sharon [σ]

Abstract - This paper takes into consideration the severe bottlenecks that have actually bedeviled econometric analysis and documentation of food security since time immemorial. It aims at modeling food security estimation using fuzzy logics. The paper shows econometrically how food security measurement drawbacks are overcome using residual diagnostic analysis by the effects of fuzzy logics on the leverage points of food security predictors. Further, the results indicate that the preliminary econometrics tests on the residual diagnostic analysis on the error variance, co linearity, multicollinearity and mahalanobis distances improved the estimation of food intake (the predicted criterion) because its predictors are stabilized upon data conversion into fuzzy membership functions. To a certain reasonable extent, it may be very safe to conclude that there is something quite positive in econometric research when fuzzy logics are applied in estimating food security, poverty among other similar subjective or qualitative variables.

Keyword : Food Security, Estimation, Fuzzy Logics, Residual Diagnostics, Econometrics of Food Security.

I. Introduction

An econometric assessment of food security as socio economic variable is still quite elusive to many researchers. Information available on measurements of variables such as food security, hunger and poverty use absolute, alternative and subjective approach methods. These methods use single indicators such as income, headcount ratios, nutritional variables or household expenditures as proxies (Booth, 1996; Rowtree, 1969 And Orshanski, 1969). Using a single indicator or a few of them as proxies do not capture the real situation of the individual households. There are approximately 200 definitions and more than 450 indicators or explanatory variables of food security (Hoddinot, 1999). This study adopts the most commonly used definition that states "food security is widely defined as access by all people at all times to enough food for active healthy life." It is a condition in which a population has physical, social and economic access to sufficient safe and nutritious food secure population can meet its consumption needs during the given consumption period by using strategies that do not compromise future food security. Food security is therefore a very complex multidimensional phenomenon, which varies through a continuum of excessive stages as the conditions change. Food security is therefore a very complex multidimensional namely, vulnerability access sufficiency and sustainability. Therefore, food security status of a given population is very complex product in a farming system characterized by interdependency and interactions at varying levels between agents such as public sector entities markets NGOs and the community among many others. These interactions result into non-linear effects - population.

The difficulty in documenting an econometric model analyzing food security stems from the fact that 450 variables are really just too many for a model. Whereas food insecurity intervention is multi-criteria in decision making, econometric estimation of parameters in such a model may be impossible because of errors in variables, heteroscedasticity, multicollinearity and autocorrelation. Further, the situation becomes even much more complicated because these relationships arise from the larger systems of food security status. This tasks often difficult and the interaction of the pieces is uncertain and complexity with respect to food security, there is no limitations of uncertainty and complexity with respect to food security and there is no consensus to date among social scientists in how food security should be econometrically documented. This research explores an econometric model approach that can accommodate the imperfect, non linear and uncertain interactions, which can explain much more precisely, the socioeconomic and the ecological outcomes of a food security status in a given population. It is generally agreed that the transition from a state of complete food deprivation to a comfortable state happens rather gradually (Micelli, 1998). Likewise in Kenya, Agricultural growth, an important precursor of food security has been below potential due to poor past policies and this will take some time to remedy (Mbogoh, 2003). This piece of research attempts to model this concept of food security while taking into consideration the severe bottlenecks that have actually bedeviled econometric analysis of food security.

a) Problem Statement

In Kenya, food insecurity in the ASAL areas is quite prevalent with sporadic cases of acute food insecurity leading to malnutrition and deaths. The poor

Author α σ : Moi University, Department of Agricultural Economics.

account for 80% of the populations in the ASAL Kenya (Republic of Kenya, 1999). The current food policy in Kenya only aims at food production and availability at national level with very little tangible measures to translate to adequate household food security especially in some of the high-risk areas (Atieno, 1996). Nearly 3.3 million people, mostly pastoralists in the northern and eastern parts of the country, need emergency food assistance every year. Their survival is at stake because the farming systems cannot do well. About 83% of the land area in Kenya is ASAL and growth in food production has not kept pace with population growth such that the burden of annual formal food transfers and drought relief supplies is continually increasing (Republic Of Kenya, 1992). About ksh 6 million is used to import food annually so as to meet the recurrent deficits. The share of imports in cereal supply in Kenya is increasing whereas domestic production is decreasing. The volatile socio-cultural, ecological and geographic constraints in the ASAL have limited efforts by the government and other agents to address this problem.

b) Measurement of food security

There are two major methods that are widely used to measure food security although both are subject to measurement problems. The first method involves estimation of gross household production and purchases over a period of time. Estimate growth and depletion of food stocks held over that period. The second method is to undertake a twenty-four hour recalls of food consumption for individual members of a household and analyze each type of food mentioned for its calorific content. This method gives data for one aspect of food security. Food consumption estimates. However, it has several drawbacks such as memory lapses, observer bias, respondent fatigue, a short and possibly unrepresentative recall period and high data collection costs. Further, the two methods do not capture important aspects of food security in relation to vulnerability, access and suitability. The two methods only capture elements of sufficiency none of them has been used to monitor food security include food balance sheets, rainfall and marketing data and even anthropometrical measurements (Maxwell & Frankenberger, 1992). They have noted a variety of indirect indicators that can be used as predictors for food security at the household level. These include; asset ownership, household size, and dependency ratio. Their discussion was based on single indictors and they suggest that combining the indicators could improve specificity. They however do not say how widely different indicators could be combined. The first food security measurement and research conference was held in 1994 in USA. The aim was to synthesize the direction of food security measurement and develop a consensus on the content. The second food security measurement

and research priorities. Consequently in the year 2000, a first attempt to measure household food security was done by the United States department of agriculture (USDA).They used a standard six-item subset of indicators to capture two thresholds of identifiable household food security. The measurement that they described was only concerned with food insecurity and hunger when household in the United States. Their measurement technique however cannot be used in developing country like Kenya because of several differences in the structure of wages, livelihood set-ups among others. Food insecurity indicators are also likely to vary very widely between a developed livelihood set-up like USA and in a developing country like Kenya.

A traditional approach to measuring food security has been limited by the fact that food consumption figures are only used or only one indicator or a proxy of food security is used. Hamilton et al 1997, USDA, for instance, uses the 18 item scale that basically captures 4 scenarios of a household's food security. These include food budget and supply inadequacy, inadequate quality, reduced intake and consequences of reduced intake. Blumberg et al 1999 proposes the six item scale that only captures the food eaten in household and whether the household can afford. However the scale developed by Blumberg et al 1999 cannot be used to measure the severe levels of food insecurity especially where hunger is experienced in high-risk groups of developing countries. Maxwell, 1995 measured food security using coping strategies only as an indicator. The above efforts to measure food security have been able account for all aspects of food security as outlined above and there is no consensus to date among social scientists on how food insecurity should be measured and whatever is used for a food security survey is just a matter of convention. Other lessons learned from measurements of other social phenomena such as poverty and childhood development are available. Official poverty line has been used to measure a constant level of purchasing power or constant living standard. Much more recent work on measurement of poverty using fuzzy sets in Switzerland shows that the use of several poverty indicators helps in giving a more complete picture of poverty than the sole use of a few indicators (Miceli, 1998). Williams et al 1999 developed national food insecurity prevalence. The specific scaling procedure that was used in a *Rasch* model, which is a form of non-linear factor analysis that fits within the general family item response theory models. The model is widespread in educational testing where the underlying premise is that the probability of affirming a question increases with the household underlying level of food insecurity and falls as the severity of the condition measured by the particular item goes up.

c) Drawbacks in Food Security Measurement

Whereas research findings in Kenya and other areas indicate that food security is strongly linked to many other factors such as markets and market policies (summer , 2000), infrastructure components such as roads, storage facilities, electricity, financial institutions, telephone, extension services and information technology (Atieno 1996) and income (Midmu 1992, Metzger & zyl 1992, Thimm 1993 & Atieno 1996) there scant literature today that documents measurement of food security with all the above factors taken into account particularly in high risk areas of Kenya. Using a few indicators or proxy to estimate food security does not capture the real situation of individual households. It is argued in this research that using all the possible crucial variables anyhow could actually capture and portray the true picture of food security in high risk food insecurity measurement technique in Kenya today. Argwings-Kodhek et al, 2002 notes that an important priority for food insecure areas should ensure functional departments and communities. However, an effective early warning, monitoring and program performance by these organizations will only be possible if there is more reliable and consistent econometric variable to monitor food insecurity. An earlier mentioned, poverty and food insecurity are really birds of a feather. The two variables have similar characteristics and most of the predictor variables are the same. Literature review on measurement of poverty was therefore found to be quite relevant and necessary for this research. In the past few decades, much contentment on the use of single indicator of resources to measure poverty has led to many authors advocating for use of alternative multivariate methods.

A major advantage of multivariate dimensional measure of poverty over the traditional methods of using one indicator is that it not only captures their general living conditions. In addition, Whenlan, 1993 notes that a global index of poverty based on set of deprivation indicators seems appropriate than indices based on only income or expenditure to assess a situation of permanent poverty. Such an index should ideally take account of the basic needs including food, clothing, housing that are mostly related to social life and sometime exerting some constraints on it. Working conditions leisure, health, education, environment, family and social activities are some examples of some of these kinds of variables. Some authors have tried to emphasize other aspects of poverty than just the monetary ones when measuring poverty. Townsend, 1997 selected sixty indicators that were supposed to summarize the common activities in society. Then he derived a deprivation index based on twelve of the items. There is another interesting approach that was proposed by Mack & Lansley, 1985. They developed and refined the theoretical and empirical work of Townsend and proposed a measure of poverty that is based on the social perceptions of needs, which means that those items classified as necessary by more than 50% of the population are defined as necessities. Hallerod, 1994 suggested a similar method expect that all items are retained as necessities to some extent in the poverty measure. Each item is given a weight based on the proportion of the population that regards it as a necessity. The current practice for measuring poverty by using a poverty line is strongly refuted by Cerioli et al 1990 due to the fact that there is no sharp division of the total population between poor and the non-poor. A poverty line at any rate is unrealistic and cannot be used sustainably.

d) Measurement of poverty using fuzzy sets

Several recent studies have proposed a multidimensional measure of poverty based on the theory of fuzzy sets. Ceriol & Zani 1990 used this method to evaluate living conditions in an Italian county. Others who extended some theoretical aspects have followed their work. The applications that have been followed so far concern mainly Italy and Poland. CERIOLI & ZANI 1990 used fuzzy sets to asses living conditions in Switzerland in 1990. Measurement of poverty using fuzzy sets involves a multidimensional analysis presenting both qualitative and quantitative variables each of them presenting a certain degree of privation. It is assumed that all the modalities can be ranked by increasing risk of poverty. An example is given by a variable showing individual subjective evaluation of their own situation such that the possible values could be very good, fairly good, average, fairly bad and very bad. Given minimum and maximum scores corresponding to those poverty limits, the membership functions as proposed by9 CERIOLI & ZANI 1990, can be expressed as.

$$\Psi_e = \frac{\Psi_{ij} - \Psi^{min}}{\Psi^{max} - \Psi^{min}}$$

Where Ψ_{ij} is the score of individual ₁ ensuing from indicator j. With this specification, the membership function increases linearly as the risk of poverty rises. Continuous variables such as income and expenditure are also found among living conditions indicators. In literature some authors provide an alternative to the problem of setting a unique clear-cut poverty line. For instance, KAKWANI 1995 proposes a method that takes into account the uncertainty about the exact value of the poverty threshold. On the other hand, ATKINSON, 1987 AND FOSTER & SHORROCKS 1988 suggests an ordinal approach related to stochastic dominance. All these methods suggest an interval supposed to contain the poverty line instead of setting of two limits. The first one is minimum value and the second one is a maximum value of the chosen indicator beyond which an individual can be regarded as out of poverty. For

those values of the variable included can be regarded as out of poverty. For those values of the variable included between the two limits, the membership function must take its values in the interval (0, 1). Further, a natural requirement for this function is that it be continuous and decreasing at least for those indicators for which an increase in value means an improvement of well being. CERIOLI & ZANI 1990 define the following membership function.

$$\Psi_e = \frac{\sum^{max} - \sum_{ij}}{\sum^{max} - \sum^{min}}$$

where $\sum_{ij}$ is the score of individual I ensuing from indicator j.

In conclusion, multidimensional measurement of poverty using fuzzy sets shows that the use of several not only helps in giving a more complete picture of living conditions, but also gives an image of poverty that is closer to what is perceived   head count ratios, discriminant   analysis and even to second order stochastic characteristic such that in many writings of poverty food insecurity is mirrored of fuzzy sets when documenting,   estimating   the   phenomenon   of programme (PRSP)in Kenya an econometric analysis of food insecurity could be long overdue and handy at any rate.

II.   Methodology

a)   Theoretical framework

The oxford English dictionary defines the word fuzzy as of something not clear in shape or sound. It implies vagueness in discerning its nature or group belonging. ZADEH 1965 first developed fuzzy concept in 1965. He mentions that some classes of objects encountered do not have a precisely defined criterion of membership. They do not constitute classes or sets in the usual way. Fuzzy concept is an aspect of mathematics and engineering and recent advances show that the same concept can be used in modeling some very important socioeconomic variables that are not clear in nature or generally vague in their description. Humans bring forth the objects of reality as linguistically labeled concepts within the requisites referential domains are chosen and adjusted for the purpose of contextual communication or negotiation. The degrees of fuzziness of linguistic labels are therefore context dependent and implied by the referential choice Zeleney 1991.

b)   Definition of fuzzy

Let X be a set and x be some element of X. A fuzzy subset A and X is defined as; A=(x, ʋA(x) for all x Ex, where ʋA,is called a membership function and is an application from X in (0,1). This means the function associates a real number in degree of belonging of x to A. because the concept of food security is not sharply

defined and multidimensional, the same concept can be used to define the fuzzy set of food security. If A is a fuzzy set, its membership to food security can only take the values between 1 and 0. In that case, $Y(x) = (1)$, $Y(x)= (0)$ or $0 < Ya(x) < (1)$. The membership function represents the degree of membership to the fuzzy subset. For the case of multidimensional analysis, increasing order of subjective evaluations can rank qualitative variables. An example is given by values attached to, for instance, excellent, extremely good, very good, good, fairly good, average, fairly bad, very bad and worse.

This study adopted a strategy that captured the concept of fuzzy logics while at the same time very adequately represent the degree of membership of the fuzzy subset to the food security indicators in question.

Fortunately, other subjective or qualitative explanatory variables of food security can easily lend themselves for analysis using fuzzy logics. The model adopted in this research can analyze both parametric and nonparametric variables in one econometric formulation and different kinds of data can be compared easily with respect to food insecurity. An empirical relationship function that was initially proposed in this research is represented here below as:

$$Xo = \beta^1 x_1 + \beta^2 x_2 + \beta^3 x_3 + \ldots\ldots\ldots + \beta^{41} x_{41} + \beta^{42} x_{42} + \beta^{43} x_{43}$$

Where, x_0- food intake $0 < x_0 < 20$, x_1- x43 represent the predictors.

However this relationship was subjected to PCA.

Food intake (x0) was used in this research as the explained variable and a proxy of food security. Food intake was chosen to reflect mainly because it is the end result of so many factors (predictors). HAMILTON et al. 1997 method was adopted when estimating this variable (x0).

An alternative total fuzzy approach in defining a membership function for qualitative variables that does not require setting any limit was also explored. CHELI et al 1994 and CHELI & LEMMI 1995 proposed the following membership function for a qualitative polytomic variable.

$$\upsilon Ej(j) = (0) \text{ if } \xi_{ij} = \xi_j^1 \text{ or } \upsilon Ej(j) = \xi_j^{max} - \xi_{ij}/\xi_j^{max} - \xi_j^{min} \text{ if } \xi 1$$

Fuzzy dynamic programming was run using Microsoft excel solver 9.00 package so as to obtain crisp optimal levels of each variable a given target group needs to be food secure. Further, the variables were analyzed using a factor analysis technique in SPSS statistical package version 11.50. Specifically principal components technique was used. The main objective of using principal components method was to identify variable dimensions that give the most explanation for the dynamics of food insecurity in the ASAL of Kenya. Principal components method is data reduction

technique that proved to be handy in this analysis. The results were documented and compare in the light of fuzzy sensitivity analysis results.

c) Use of cumulative food security index

Recent research findings reveal that the use of cumulative indices in food security analysis has become popular. MAXWEL 1999 used indices while measuring coping strategies as a food security indicator. He developed a simple scale of 1-4 for the frequency of each individual strategy and multiplied by the weighting factor based on each strategy. Unfortunately, one weakness this method has was that the results could not be compared with other food security data because they were in different units and therefore were measures of totally different phenomenon.

The resulting fuzzified data in this research are more or less indices. The food intake data was also based on an eighteen question simple scale of 1-3 and 1-2 scores, which were also fuzzified to obtain indices. The eighteen question scale of fuzzy data was further aggregated to get a single indicator of food intake (x0). This research adopted Chiappero-Martinnetti 1994 method of using generalized weighted average as the aggregation operator while evaluating the degree of membership of each individual to the fuzzy subset of food intake was therefore obtained from the eighteen question scale of fuzzified data using generalized weighted average as the aggregation operator. The main reason for having all the data fuzzified in this research was to ensure that all the different predictor (explanatory)variables with different units could be decomposed comfortably and compared statistically such that all econometric problems that arise in the process of estimation of parameters may be eliminated.

Justification for using fuzzy theoretical concept.

Application of fuzzy logics in the model has several very crucial advantages if applied in econometrics. Firstly, it attempts to standardize all the variables that have different predictor (explanatory) variables with different units of measurement. Secondly, a preliminary analysis of data reveals that the resulting fuzzified data (standardized variables) offers a solution to the problem of multicollinearity in variables. Farrar – Glauber test 1976 of multicollinearity in the raw data (not fuzzified) was found to be 3.968E-09 whereas the standardized determinant for the fuzzified data was found to be 7.671E-09. The observed chi-square also reduces after conversion into fuzzy membership functions. This implies that multicollinearity reduces. Further, a preliminary analysis on co linearity statistics indicate that the tolerance values obtained from the raw data were lower than tolerance values obtained from the fuzzified data. Further, fuzzified data were found to have lower values of variance inflation factors (VIF). This is indicative of reduced collinearity in the variables once

subjected to conversion into fuzzy membership functions.

Thirdly, conversion into fuzzy membership functions of the original data also transforms the model into newly adjusted variables. This kind of adjustment provides a solution for heteroscedastic disturbances in econometrics. A preliminary analysis on heteroscedacity also reveals that plotting studentized residuals against food intake (predicted criterion) for the raw data shows non-linearity and heteroscedasticity. Upon data conversion into fuzzy membership functions, results of plotting the studentized residuals against food intake shows a prototype plot where the residuals approach a uniform spread, though not perfectly, along the zero mean value. These results further revealed that there is a relatively more uniform distribution around the zero mean for the fuzzified data than the raw data. The residual, standardized residuals, studentized residuals, deleted residuals and the deleted studentized residuals all reduced when the observed data was subjected to conversion into fuzzy membership functions. Standard deviation in this case reflects the degree of deviation from the mean 'zero' value. The econometric implication here is that heteroscedastic disturbances during the estimation of food intake (food security) as a dependent variable are minimized though not eliminated. Fourthly, conversion into fuzzy membership functions of the original data also minimizes errors in variables. Errors in variables in econometric analysis arise due to errors in observation, wrongly published data, and omitted data, use of dummy variables and indices. Due to the multidimensional nature of food security in this research, several variables were omitted. This was already an error in the raw data that needed a solution. Further, there were several dummies and indices used in this research. Conversion into fuzzy membership functions of the data could therefore play a very important role in arresting error problems that might have resulted from using raw data in the estimation analysis. The most crucial problem due to errors in variables is that the estimates of the coefficients become both biased and inconsistent.

Fifthly, a preliminary residual diagnostic analysis on the effects of fuzzy logics on the leverage points of food security predictors revealed that more predictors had values larger than the computed point of leverage value. These imply that observations carry a disproportionate weight in determining its predicted dependent variable value, thus minimizing its residual (ROUSSEEUW 1987. These imply that more observations increase increase their degree of influence on the regression results when they are subjected to data conversion into fuzzy membership functions. That is, more of the fuzzified food security data begin to fall in the general pattern of the remaining observations and hence closer to the regression line. The implication of these results is that estimation of food intake (food

security) as a dependent variable is reasonably improved.

Further, the impact of fuzzy logics on the mahalanobis distances reveals that the mean and the standard deviation of the mahalanobis distances for the fuzzified data are 36.690 and 25.304 respectively whereas the mean and standard deviation of the mahalanobis distances for the raw data are reduced such that the observation from the mean centre of all other observations decreases. That, is outlier characteristics in the pool of food insecurity data are reduced such that the data are much more uniform than before. The implication is that estimation of food intake or food security as a dependent variable is improved for our case. Mahalanobis distances are essentially a measure of the number of outliers among predictors variables (BELSEY 1980. Further, the studentized residuals analyses show that the range of studentized residuals for the raw data is 2.936 and 3. 071 whereas studentized residuals for the raw data is 1.477 and 2.591. This indicates that for the fuzzified data, more of the studentized residuals tend to fall in the range + 2 and -2 than for the raw data. These imply that the number of outliers tend to reduce such that the mean centre of all of observations decrease. The econometric implications are that the food security predictors are such that they are more uniform than before consequently; the estimation of food intake (predicted criterion) is improved in the regression model under consideration in this research. Loadings such that for n > 50, a loading was regarded as statistically significant at 99% level of significance. The first method was not adopted in this research because it is regarded as rather crude and has little statistical justification.

III. Research Design

A farming systems research approach was used to describe the prevailing food security systems, resource endowments, identify farm household objectives in attaining food security, constraints, the available coping or adaptive strategies and instruments.

a) Sampling procedures

The population was divided into three random sampling units in six villages of West Pokot, Uasin Gishu and Baringo districts of Kenya. The random selection was adopted to ensure that the study population got equal chances of being represented in the study population got equal chances of being represented in the study. Questionnaires were used for eliciting data from the various rural farm households. They include structured questionnaires and a checklist targeting information such as detailed above.

b) Secondary data and sources

Secondary data that was invaluable to this study included world development reports on food security and poverty measurements and livelihoods in the developed and in the developing countries, integrated food development programs in the semi arid lands of Kenya, existing development interventions by the government and non-governmental organizations.

c) Primary data and sources

Primary data such as food consumption data, access and availability, sources of income, income and income generating activities, farm output, yield per unit land area, population, food transfers spatially and temporally, household expenditure, food expenditure, poverty alleviation programmes and inter-sectoral collaboration, human resource development factors such as roads, education, health (number of diseases, distance to hospital), safe water and sanitation, nutritional health care among others were collected. Food intake was obtained using a recall method of eighteen questions each scaled 1-3 and 1-2 indicating the extent to which a household was deprived of food intake. A food intake index was obtained from these 18 questions and was used as the explained (predicted) variable in the econometric for this research.

d) Data analysis

Data was first fuzzified and subjected to various preliminary econometric tests of multicollinearity, heteroscedasticity and autocorrelation. Econometric criterion measures such as the statistical significance of the parameters, t-statistic among others were used to test the validity of the data. To test for multicollinearity, this research adopted Farrar-Glauber method, (1967) of analysis to test both raw and fuzzified data. The Farrar-Glauber test for multicollinearity actually involves the chi-square, F-test and the t-test. A plot of residuals against the fitted values (the predicted criterion) was used to detect whether the variance of the error term is constant or not (Heteroscedasticity). The fuzzified data was further run using Microsoft excel solver 9.0 so as to obtain sensitivity analysis reports of some variables with respect to food intake. This analysis aimed at identifying variables that show the greatest movement of units in the process of optimizing of food intake. Excel solver 9.0 can run a linear programming package of purpose of optimization. Data was also analyzed using factor analysis (principal components methods). This was aimed at arriving at the variables that give the most explanation to food insecurity. The principal components method is a data reduction technique. The principal component matrix was analyzed so as to identify variables that give the highest amount of loadings to the new reduced factors. Using seven different econometric models, a further analysis of specific variables based on principal components output was done so as to identify how they affect the level of household food intake assuming the other factors are held constant forecasting of complex societal

variables such as food insecurity and poverty because the fit to the regression model is crucial.

The econometric implications of fuzzy logics in food security estimation

This sub-section briefly aims at showing how fuzzy logics affect factors affecting food intake(predictors) and especially its impact on the accuracy of estimation of the predicted variable (household food intake level) in a regression model. As mentioned earlier, an important problem facing estimation of poverty, food multicollinearity and heteroscedastic disturbances in the usual kind of data used. The impact of fuzzy logics on residual statistics, serial correlation, predictor multicollinearity and heteroscedastic disturbances were critically analyzed using SPSS version 11.5 and the result are shown here below.

The impact of fuzzy logics on residual statistics in food security estimation

Presence of unequal variance of error is one of the most common violations of assumptions if food security estimation data. Diagnosis in this research was made with residual plots. Studentized residuals were plotted against the predicted plots. A pattern of plots showing the behavior of variance was closely observed for the raw data and for the data converted into fuzzy membership functions. The household food intake data that was used in making the plots were from the three districts in the study areas. Results of plotting studentized residuals against household food intake scores (the predicted criterion) for the raw data shows non-linearity and heteroscedasticity in figure 4.1. The results in figure 4.1 show that the residuals for household food intake scores range between - 4 and 4. Upon conversion of data into fuzzy membership functions, results of plotting the studentized residuals against household food intake in figure 4.2 here below

shows prototype plot where the residuals tend to approach a uniform spread, along the zero mean value. The results in figure 4.2 show that the residuals for household food intake membership functions range between -3 and 3. All the residuals reduced. These results therefore suggest that there is a relatively more uniform distribution around and much closer to the zero mean value for the fuzzy membership functions data than the raw data (household food intake scores observed). The residual statistics in tables 1 and 2 further confirm and revealed that the residual standard deviations for residuals, standardized residuals, studentized residuals, deleted residuals and the deleted studentized residuals all reduced when the observed data was subjected to conversion into fuzzy membership's functions. Standard deviation in this research reflects the degree of deviation from the mean 'zero' value. The standard deviations of various residuals for the raw data (household food intake scores observed) in table 4.2 respectively are 0.563, 0.847, 1.169, 1.48 and 1.197. Whereas the standard deviations of same residuals for the fuzzy membership functions data in table 4.1 respectively are 0.153, 829, 1.04, 0.36 and 1.058. These results suggest that the adverse effects of heteroscedasticity in model estimation of food security are alleviated to a reasonable extent when data is converted into fuzzy membership functions. The results of this particular analysis suggest that the standard deviations for all forms and types of residuals computed are reduced. The most important econometric implications here is that heteroscedastic disturbances (stochastic factors) are minimized though not eliminated completely. To all intents and purposes, these results further imply that the estimation of food intake and purposes, these results further imply that the estimation of food intake per household(food security) as a dependent variable or the predicted criterion is improved at any rate in a given regression model.
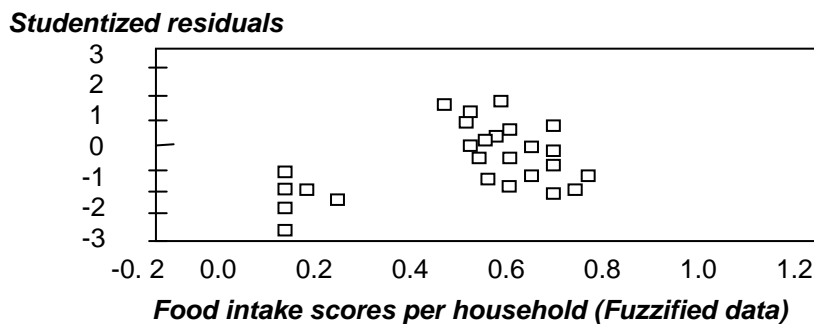
Figure 4.1 : Scatter plot of studentized residuals against food intake scores per household in west Pokot and Baringo districts, 2010 (raw data).
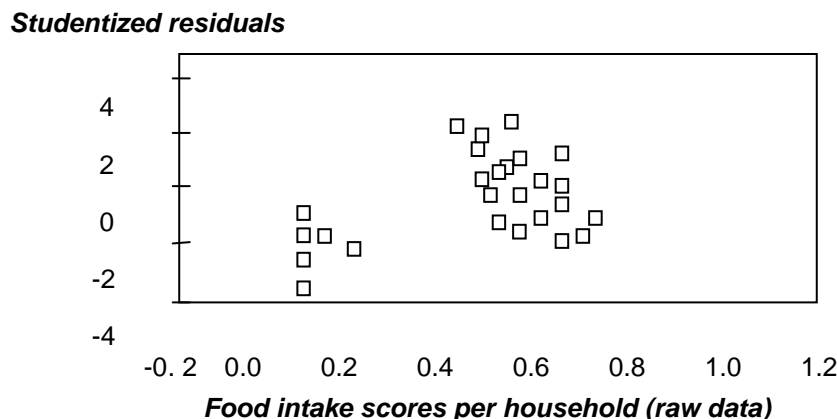
Source : Authors own Compilation, 2012

Figure 4.2 : Scatter plot of studentized residuals against food intake levels per household in west Pokot and Baringo districts, 2010 (fuzzified data).

Source : Authors own Compilation, 2012

The impact of fuzzy logics on predictor leverage points

Residuals were also used to identify those observations that are from the remaining observations on one or more independent variables of food security. The computed point of leverage value for the observed data that was used in this research was 0.6068. Values larger than the average value imply that these observations carry value, thus minimizing its residual {Rousseeuw, 1987}. The leverage values for the raw data and the data converted to fuzzy membership functions were computed and compared so as to measure the degree of influence and how each of the observation and computed (fuzzified) data had on the predicted dependent variable (household food intake level). The residual statistics results in table 1 and 2 here below show that when data is converted into fuzzy membership functions, the mean and standards deviation of the centered leverage value is 0.310 and 0.205 respectively. When the data is raw (not presented as fuzzy membership functions), the mean and standard deviation of the centered leverage value were 0.281 and 0.220 respectively. These results imply that more observations increase their degree of influence on the regression results when they are subjected to conversion into fuzzy membership functions. That is, more of the fuzzified food security data begin to fall in the general pattern of the remaining observations and hence closer to the regression line. The implication of these results is that estimation of food intake level (a proxy for food security) as a dependent variable is improved to some reasonable extent.

Table 1 : Residuals Statistics for the Impact of Rural infrastructure on household food intake in West Pokot and Baringo Districts, 2010.

	Minimum	maximum	mean	standard deviation	N
Predicted value	-1145	.7745	.3840	.12426	129
Standardized Predicted value	-2.169	3.143	.000	1.000	129
Standard error of Predicted	-1.9382 1.3281	.3512	.31536	127	
Residual	-.4796	.2733	.0000	.15346	129
Standardized residual	-2.591	1.477	.000	.829	129
Studentized residual	-2.904	2.650	.016	1.044	127
Delet	-7839	2.4970	.0328	.36862	127
Studentized deleted Residual	-3.037	2.747	.013	1.058	127
Mahalanobis distance	.000	126.834	36.690	25.304	129
Cooks distance .000	4.433	.070	.413	127	
Centered leverage value.000	.991	.310	.205	129	

Note : data in fuzzy membership functions.

Source : Authors own Compilation, 2012

Table 2 : Residual Statistics for the Impact of Rural infrastructure On Food Intake Level per Household in West Pokot and Baringo Districts, 2010.

	Minimum	maximum	mean	std. deviation	n
Predicted value .2763	2.9279	1.4696	.46138	146	
Standard predicted value -2.586	3.161	.000	1.000	146	
Standard error of predicted value .05503	.66398	.32630	.14451	146	
adjusted predicted value	-6.6350 8.1666	1.4063	1.33421	127	
residual	-2.0421 1.9527	.0000	.56318	146	
standardized residual	-3.071	2.936	.000	.847	146
studentized residual	-3.659	3.610	.028	1.169	127
deleted residual	-6.0555 8.4684	.0633	1.48222	127	
Studentized deleted residual -3.901	3.841	.026	1.197	127	
mahalanobis distance	.000	143.569	43.719	34.725	146
cooks distance	.000	3.816	.085	.402	127
centered leverage value .000	.990	.281	.220	.146	

Note: Raw data

Source: Authors own Compilation, 2012

The impact of fuzzy logics on serial correlation.

The kind of data that was used in this research was cross-sectional data. Consequently, the data referred to a given point in time such that temporal dependence was automatically ruled out by the nature itself of cross section random samples. However the impact of fuzzy logics can be observed once time series data is available. The behavior of residuals in this research was found to be near normal. Results in this research indicate that the residuals from the fuzzified data approach the normal curve than the residuals resulting from raw data from the field.

The Farrar – Glauber test for multicollinearity

As earlier noted, multicollinearity is an important predicament in the estimation of household food intake levels because most of the independent variables move together rendering the estimation of parameters indeterminate. This research therefore adopted the Farrar-Glauber test (1976) to test for multicollinearity in the raw data and also in the resulting fuzzified data. To test for the existence and severity of multicollinearity, chi-square test was used. Farrar-Glauber test found out that the quantity x* can be used to test for the severity of multicollinearity from the observed sample and the computed data (fuzzified data). This is given by the formulae:

x*= -(n-1-1/6(2k+5) log$_e$(value of the standardized determinant).

Where x* was the observed or the computed value x, n is the sample size, k is the number of explanatory variables. This quantity, x* has a *Chi-* distribution with v=1/2k (k-1) degrees of freedom. A standardized determinant is the determinant of the self-correlation matrix resulting from the explanatory variables of the model. The standardized determinant for the raw data (not fuzzified) was found to be 3.968E-09. Using the above formulae developed by Farrar-Glauber test the observed (computed) x was 2204.40. On the other hand, standardized determinant for the fuzzified data was found to be 7.671E-09. Using the above formulae developed by Farrar and Glauber, the observed (computed) x for the fuzzified variables was 2041.09. According to Farrar-Glauber test, the higher the observed x, the more severe the multicollinearity. It is observed above that the value of the computed *Chi-* square reduces after conversion into fuzzy membership functions. The value of any standardized determinant lies between Zero and unity. The closer the value to zero the stronger the degree of multicollinearity and vice versa. The two most important measures for testing the impact of collinearity are:

- Tolerance levels and variance inflation factors
- Using condition indices and decomposing the regression coefficient variance (BELSLEY 1980 COHEN 1983)

A condition index is a measure of the relative amount of variance associated with an eigen value so that a large condition index indicates a high degree of collinearity. The tolerance value is defined as one minus the degree of collinearity. The tolerance value is defined as one minus the proportion of variables variance explained by the other predictors. Tolerance of a variable is 1-R^2 where R^2 is the coefficient of determination for the prediction for the variable by the

other predictor variables. Tolerance values approaching zero indicate that the variable is highly predicted with the other predicted variables. Thus a high tolerance value indicates little multicollinearity and tolerance values approaching zero the variable is almost totally accounted for by the other variables. Variance inflation factor (VIF) is a measure of the effect of the other predictor variables. Tolerance values approaching zero indicate that the variable is highly predicted with the other predictor variables. Thus a high tolerance value indicates little multicollinearity and tolerance values approaching zero the variable is almost totally accounted for by the other variables. Variance inflation factor (VIF) is a measure of the effect of the other predictor variables on a regression coefficient. Variance

inflation factor (VIF) is the reciprocal of tolerance values. Small values of VIF are therefore indicative of little multicollineraity. Tolerance values obtained from the raw data is tables 3 here below were found to be lower than tolerance values obtained from the fuzzified data. further, fuzzified data were found to have lower values of variance once subjected to conversion into fuzzy membership functions. When multicollienarity does not have a substantial impact on the regression variates (variables affecting household food security), it does have an impact on the accuracy in estimating the predicted variable (household food intake) positively for our case. Consequently, data conversion into fuzzy membership functions improves the estimation process of food intake (the predicted criterion in this research).

Table 3 : Coefficient Estimates For The Impact Of Rural Infrastructure On Food Intake Level Per Household In West Pokot and Baringo Districts, 2010.

	Standardized coefficients	Standard error	significance	Tolerance values	Variance inflation factors
Constant					
Distance to tarmac road	0.167	0.011	0.39	0.792	1.262
Distance to hospital	- 0.117	0.063	0.564	0.722	1.386
Distance to rail	- 0.367	0.017	0.181	0.403	2.482
Distance to school	- 0.092	0.058	0.637	0.776	1.288
Distance to financial institution	0.032	0.014	0.882	0.637	1.570
Distance to electricity supply	0.164	0.019	0.444	0.646	1.548
Distance to market	0.350	0.012	0.121	0.601	1.665
Distance to telephone	-0.043	0.014	0.839	0.646	1.547

Source: Authors own Compilation, 2012
Note: Fuzzified Data

Therefore data conversion to fuzzy membership functions reduced the extent to which variables are multicollinear. It is worth noting that changing data to fuzzy membership functions do not make the explanatory variables perfectly orthogonal. That is, multicollinearity is not completely eradicated from the model because the theoretical chi-square (x) in the model is 834, which is far below the observed values for both the fuzzified data and raw data. The implication of these results is that estimation of food intake or food security as a dependent variable is just improved and not made perfect.

The impact of fuzzy logics on the Mahalanobis Distances.

As noted earlier, factors affecting food security are many and widely different either in their measurement characteristics. Therefore, outlier characteristics in a pool of food insecurity data are a very much expected even after discarding observations that are obviously outliers in sets of observation with more uniform characteristics. This is an important predicament that ought to be handled so ass to for the same analysis is the mahalanobis distances which essentially measures the number of outliers among predictor variables (BELSLEY 1980). Mahalanobis distance is a measure of the impact of a single case based on differences between case value and the mean value for all other cases across the independent variables. The source of influence on regression results is for the case to be quite different on one or more predictor variables thus causing a shift of the entire regression equation. Though this method is not widely used, the residual statistics in table 2 above show that

the mean and the standard deviation of the mahalanobis distances for the fuzzified data are 36.690 and 25. 304 respectively whereas the mean and standard deviation of the mahalanobis distances for the raw data in table 3 were 43.719 and 34.725 respectively. These results definitely indicate that when the observed data is subjected to conversion into fuzzy membership functions before analysis the distances of the observation from the mean centre of all other observations decrease. That is, outlier characteristics in a pool of food insecurity data are reduced such that the data are much more uniform than before. The implication of these results is that estimation of food intake or food security as dependent variable is improved for our case.

Further, the studentized residual is another primary indicator of observations that are outliers on the dependent variables (Barnet, 1984). With a sample size more than fifty, the rule of thumb is that residuals outside the range of -2 and +2 are significant. Residuals falling outside this range are actually considered outlier. Results in table 2 show that the studentized residuals for the fuzzified data range between 1.477 and -2.591 whereas studentized residuals for the raw data in table 3 are 2.936 and -3.071. These results indicate that for the fuzzified data, more of the studentized residuals tend to fall in the range +2 and -2 than for the raw data. These results suggest that the number of outliers tend to reduce such that the mean centre of all of observations decreases. The econometric implications of these results are that the food security predictors are much more uniform than before. Consequently, the estimation of household food intake (the predicted criterion) is the regression model under consideration in this research.

IV.   Conclusions

Based on the preliminary econometrics tests on the residual diagnostic analysis on the error variance, co linearity and multicollinearity, mahalanobis distances and the predictor leverage points, the estimation of food intake (the predicted criterion) is improved because the predictors of food security are stabilized upon data conversion into fuzzy membership functions. To a certain extent, it may be very safe to conclude that there is something quite positive in econometric research when fuzzy logics are applied in establishing food security.

V.   Recommendations

Based on the results of this research report, it could be very safe to assert that the use of fuzzy logics in econometric research is a step forward in the improvement of parameter estimation, monitoring and forecasting of important societal variables such as poverty, food insecurity and others subjective variables that have important ramifications to the community well being. Preliminary analysis results shows that fuzzy logics positively change the behavior of residual statistics in food security as the predicted variable. It is therefore recommended to econometricians that the use of fuzzy logics in estimation, monitoring and forecasting of phenomena such as poverty and food insecurity in Kenya be given a priority.

References Références Referencias

1. Atieno R. [1996], Role of rural infrastructure in attaining household food security: research issues and policy options with reference to Kenya. Proceedings of the 14th international symposium on sustainable farming systems, Colombo Srilanka.
2. Booth C. [1996], Life and labor of people in London. A.M. Kelly, New York.
3. Chiappero-Martinnetti {1994}, A new approach to evaluation of well being and poverty by Fuzzy sets theory. Giornale Degli. Economisti e annalidi Economia , Volume 53,Nos 7-9 , pages 367-388.
4. Ceriol & Zani [1990], A fuzzy approach to the measurement of poverty. In Dagum C. and Zenga, Income and Wealth Distribution, inequality and poverty. Studies in contemporary Economics, Springer Verlag, Berlin pages 272-284.
5. Hallerod, B. [1994], A New Approach To The Direct Consensual Measurement of Poverty.
6. Hoddinot J. [1999], Operationalizing Household Food security in Development Projects: An Introduction. International Food Policy Research Institute [IFPRI]. Washington DC, 20006, 1999.
7. Micelli D. [1998], Measuring Poverty Using Fuzzy Sets, Discussion Paper 38. December 1998.National centre for social and Economic Modeling, Faculty of Management, Canberra.
8. Mbogoh S.G. [2003], An Overview of Major Policy Documents. Department of Agricultural Economics, University of Nairobi, Kenya. [2003].
9. Republic of Kenya [1999], National Poverty Eradication Plan, 1999-2015, Department Of Development Coordination, Kenya.
10. Republic of Kenya [1992], Development Policy for ASAL areas. 1992 Department Of Development Coordination, Kenya.
11. Frankenberger T. and Mccston M. [1998], The Household livelihood security Concept. Poverty Briefing Lists, ISSN 1465-2617.
12. Maxwell S. [1999], Meaning and Measurement of Poverty. Overseas Development Institute Poverty Briefing Lists, ISSN 1465-2617.
13. Hamilton, J. [1997], Household Food security in the United states in 1995. Summary Report Of The Food Security Project. Household Food Security In The United States, 2003/FANNRR Economics Research Service/ USDA.

14. Micelli, D. [1998], Measuring poverty using fuzzy sets. Discussion paper number 38, December 1998. National centre for social and economic modeling. Faculty of Management, University Of Canberra.
15. Kakwani N.C. [1997], Measuring poverty with uncertain poverty lines: journal of income Distribution. Volume 5, number 2,pages 202-210.
16. Rousseeuw, P. J. and Leroy, A.M.[1987], Robust Regression and outlier Detection. New York: Wiley.
17. Zadeh, L.A. Fuzzy sets: Information and Control. Volume 8, pages338-353
18. Zeleney M. [1991], Fuzziness, Knowledge and Optimization. New Optimality Concepts. In Delgado, M.et al [1994], Fuzzy optimization. Recent Advances. Physica –verlag: A Springer Verlag company.