Fuzzy Approach to Semi-parametric of a Sample Selection Model

L. MUHAMAD SAFIIH [1]*, A.A.BASAH KAMIL[2], M. T. ABU OSMAN [3]
[1, 3]Mathematics Department
Faculty of Science and Technology, University Malaysia Terengganu
21030 Kuala Terengganu, Terengganu, MALAYSIA.
safiihmd@umt.edu.my, abuosman@umt.edu.my, http://www.umt.edu.my

[2]School of Distance Learning
Universiti Sains Malaysia
11800 USM Penang, MALAYSIA.
anton@usm.my; http://www.usm.my

Abstract: The sample selection model is studied in the context of semi-parametric methods. With the deficiency of the parametric model, such as inconsistent estimators etc, the semi-parametric estimation methods provide the best alternative to handle this deficiency. Semi-parametric of a sample selection model is an econometric model and has found interesting application in empirical studies. The issue of uncertainty and ambiguity still become are still major problem and are complicated in the modelling of a semi-parametric sample selection model as well as its parametric. This study, focuses on the context of fuzzy concept as a hybrid to the semi-parametric sample selection model. The best approach of accounting for uncertainty and ambiguity is to take advantage of the tools provided by the theory of fuzzy sets. It seems particularly appropriate for modelling vague concepts. Fuzzy sets theory and its properties, through the concept of fuzzy number, provide an ideal framework in order to solve the problem of uncertain data. In this paper, we introduce a fuzzy membership function for solving uncertain data of a semi-parametric sample selection model.

Key-Words:- uncertainty, semi-parametric sample selection model, crisp data, fuzzy number, membership function

1 Introduction

The sample selection model is studied in the context of semi-parametric methods. With the deficiency of the parametric model, such as inconsistent estimators etc, the semi-parametric estimation methods provide an alternative to handle this deficiency. The study of semi-parametric econometrics of the sample selection models has received considerable attention from statisticians as well as econometricians in the late 20th century (see Schafgans,1996). The termed "semi-parametric" is used as a hybrid model for the selection models, which do not involve parametric forms on error distributions. Hence, only the regression function part of the model of interest is used. Consideration is based on two perspectives: firstly, no restriction of estimation of the parameters of interest for the distribution function of the error terms; secondly, restricting the functional form of heteroskedasticity to lie in a finite-dimensional parametric family (Schafgans,1996).

Cosslett (1990) considered semi-parametric estimation of a two-stage method similar to Heckman (1976) for the bivariate normal case where the first stage consisted of semi-parametric estimation of binary selection model and the second stage consisted of estimating the regression equation. Ichimura and Lee (1990) proposed an extension of applicability of a semi-parametric approach. It was proven that all models can be represented in the context of multiple index frameworks (Stoker, 1986) and shown that it can be estimated by the semi-parametric least squares method if identification conditions are met (see also, Klein and Spady (1993), Gerfin (1996), Martins (2001), Khan and Powell (2001)). General speaking, the previous studies in this area concentrated on sample selection models using parametric, semi-parametric or nonparametric approaches. More specifically, none of these researchers put efforts into studies that analysed semi-parametric sample selection models in the context of fuzzy environments like fuzzy sets, fuzzy logic or fuzzy sets and systems (Muhamad Safiih , 2007).

Muhamad Safiih *et al.*(2006), introduced a concept of membership function to the sample selection model by using an S-curve membership function. In this paper, a real data set (income from Malaysian Family and Population Survey 1994) was used to get the upper and lower membership function.

However, the concept of fuzzy membership function have been widely used in other areas. The first article was by Tanaka, Hayashi and Asai (1982) based on a linear regression hybrid with fuzzy concept called fuzzy linear regression model. In this study, fuzzy dependent variable and crisp independent variable are formulated as a mathematical programming problem through the regression problem. Kao and Chyu (2002) used this concept on regression coefficients, which provide the best explanation for the relationship between the independent, and dependent variables.

The purpose this paper is to introduce a membership function of a sample selection model in which historical data contains some uncertainty. Examples of uncertainties are experience, income, etc,. Referring to Heckman sample selection model in (4) which has been widely used, the problem of biasness still occurred, if we choose only those who participate i.e. 1 and leave behind those who do not non participate or 0. By introducing the concept of membership function into the model, our study actually gives a degree not only for participation but for non-participation as well. With this concept, the problem of biasness will be minimised and, at the same time, the information as well as the reliability of data will be increased.

Through this, the fuzzy concept provides an ideal framework to deal with problems for which there does not exist a definite criterion for discovering what elements belong or do not belong to a given set (Miceli, 1998). Fuzzy set is defined by a fuzzy set in a universe of discourse U characterised by a membership function which denoted by the function μ_A maps all elements of U that take the values in the interval [0,1] that is $A: X \rightarrow [0,1]$ (Zadeh, 1965). The concept of fuzzy sets by Zadeh is extended from the crisp sets, that is the two-valued evaluation of 0 or 1, {0, 1}, to the infinite number of values from 0 to 1, [0, 1]. (see Terano *et.al*. 1994).

2 Representation of uncertainty

Generally, fuzzy number represents an approximation of some value which is in the intervals terms $[c^{(l)}, d^{(l)}], c^{(l)} \le d^{(l)}$ for l $0,1,..,n$, is given by the α- cuts at the α-levels μ_l with $\mu_l = \mu_{l-1} + \Delta\mu$, $\mu_0 = 0$ and $\mu_n = 1$, usually provide a better job set to compare the corresponding crisp values. As widely practiced, each α-cuts ${}^{\alpha}A$ of fuzzy set A are closed and related with interval of real numbers of fuzzy numbers for all $\alpha \in (0,1]$ and based on the coefficient $A(x)$: if ${}^{\alpha}A \ge \alpha$ then ${}^{\alpha}A = 1$ and if ${}^{\alpha}A < \alpha$ then ${}^{\alpha}A = 0$ which is the crisp set ${}^{\alpha}A$ depends on α.

Closely related with a fuzzy number is the concept of membership function. In this concept, the element of a real continuous number in the interval [0,1] or, in other words, representing partial belonging or degree of membership are used. The triangular membership function is used. These are represented as a special form as:

$$\mu_A(x) = \begin{cases} \dfrac{(x-c)}{(n-c)} & if\ x \in [c,n] \\ 1 & if\ x = n \\ \dfrac{(d-x)}{(d-n)} & if\ x \in [n,d] \\ 0 & otherwise \end{cases} \qquad (1)$$

From that function, the α-cuts of a triangular fuzzy number can be defined as a set of closed intervals as

$$[(n-c)\alpha + c, (n-d)\alpha + n], \forall \alpha \in (0,1] \qquad (2)$$

and the graph of a typical membership function is illustrated in Figure 1

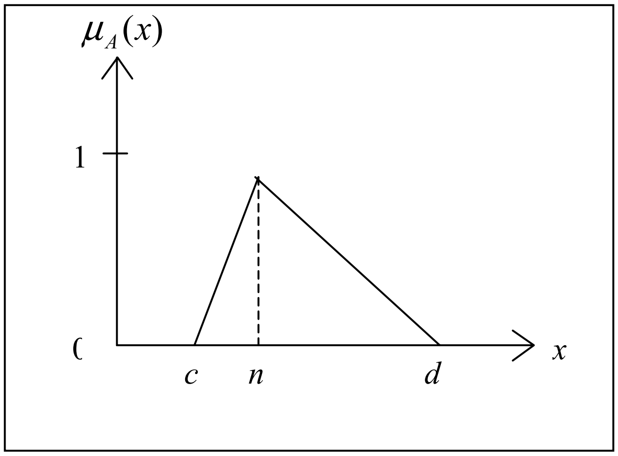

Figure 1: A triangular fuzzy number

For the membership function $\mu_A(x)$, the assumptions are as follows:

(i) monotonically increasing function for membership function $\mu_A(x)$ with $\mu_A(x) = 0$ and $\lim_{x\to\infty} \mu_A(x) = 1$ for $x \le n$

(ii) monotonically decreasing function for membership function $\mu_A(x)$ with $\mu_A(x) = 1$ and $\lim_{x\to\infty} \mu_A(x) = 0$ for $x \ge n$

3 The α-cuts representation of fuzzy number

Before going deeper into fuzzy modelling of PSSM, an overview some definitions are presented (Yen *et.at.* (1999), Chen and Wang (1999)) used in this study is related to the existence fuzzy set theory introduced by Zadeh's(1965). The definitions and properties are as follows:

Definition 1*:* the fuzzy function is defined by $f : X \times \widetilde{A} \to \widetilde{Y}; \widetilde{Y} = f(x, \widetilde{A})$, where

1) $x \in X$; X is a crisp set;
2) $\widetilde{A}$ is a fuzzy set, and
3) $\widetilde{Y}$ is the codomain of x associated with the fuzzy set $\widetilde{A}$

Definition 2: Let $A \in F(\Re)$ be called a fuzzy number if:

1) exist $x \in \Re$ such that $\mu_A(x) = 1$
2) for any $\alpha \in [0,1]$

$A_\alpha = [x, \mu_{A_\alpha}(x) \ge a]$, is a closed interval with $F(\Re)$ represents all fuzzy sets, $\Re$ is the set of real numbers.

Definition 3: define a fuzzy number A on $\Re$ to be a triangular fuzzy number if its membership function $\mu_A(x) : \Re \to [0,1]$ is equal to

$$\mu_A(x) = \begin{cases} \dfrac{(x-l)}{(m-l)} & if \ x \in [l, m] \\ 1 & if \ x = m \\ \dfrac{(u-x)}{(u-m)} & if \ x \in [m, u] \\ 0 & otherwise \end{cases}$$

where $l \le m \le u$, x is a model value with l and u be a lower and upper bound of the support of A respectively. Then the triangular fuzzy number denoted by (l, m, u). The support of A is the set elements $\{x \in \Re \mid l < m < u\}$. A non-fuzzy number by convention occurred when $l = m = u$.

Theorem 1*:* The values of estimator coefficients of the participation and structural equations for fuzzy data converge to the values of estimator coefficients of the participation and structural equations for non-fuzzy data respectively whenever the value of $\alpha -$cut tend to 1.

Proof. From the centroid method that followed to get the crisp value, the fuzzy number for all observation of w_i as

$$W_{ic} = \frac{1}{3}\left(Lb(w_i) + w_i + Ub(w_i)\right)$$

when there is no $\alpha - cut$. The lower bound and upper bound for each observation referred to by the definition 3 above.

Since we follow the triangular membership function, is followed see Figure 4.2, then $A = \left(Lb(w_{i(\alpha)}), \alpha\right)$ and $B = \left(Ub(w_{i(\alpha)}), \alpha\right)$
where

$$Lb(w_{i(\alpha)}) = Lb(w_i) + \alpha\left(w_i - Lb(w_i)\right)$$

and

$$Ub(w_{i(\alpha)}) = Ub(w_i) + \alpha\left(w_i - Ub(w_i)\right)$$

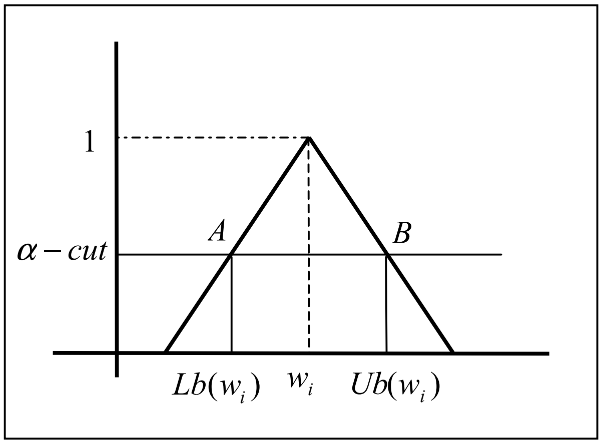

Figure 2: Membership function and α -cut

Applying the $\alpha - cut$ into the triangular membership function, the fuzzy number is obtained that depends on the given value of the $\alpha - cut$ over the range 0 and 1 is as follow:

$$W_{ic(\alpha)} = \frac{Lb(w_i) + \alpha\left(w_i - Lb(w_i)\right) + w_i + Ub(w_i) + \alpha\left(w_i - Ub(w_i)\right)}{3}$$

$$= \frac{Lb(w_{i(\alpha)}) + w_i + Ub(w_{i(\alpha)})}{3}.$$

When α approaches 1 from below then $Lb(w_{i(\alpha)}) \rightarrow w_i$ and $Ub(w_{i(\alpha)}) \rightarrow w_i$. Further obtained is $W_{ic(\alpha)} \rightarrow \frac{w_i + w_i + w_i}{3} = w_i$,

$$W_{ic(\alpha)} \rightarrow w_i.$$

The last equation states that when α approaches 1 from below then $W_{ic(\alpha)} \rightarrow w_i$. Similarly, for all observations x_i and z_i, $X_{ic(\alpha)} \rightarrow x_i$ and $Z_{ic(\alpha)} \rightarrow z_i$ respectively, as α tends to 1 from below. This implies that the values of estimator coefficients of the participation and structural equations for fuzzy data converge to the values of estimator coefficients of the participation and structural equations for non-fuzzy data respectively whenever the value of $\alpha -$cut tend to 1.

4 Development of Fuzzy Semi-parametric of Sample Selection Models

Before constructing a fuzzy SPSSM, first, the sample selection model proposed by Heckman (1976) is considered. In SPSSM, it is assumed that the distributional assumption of (ε_i, u_i) is weaker than the distributional assumption of the parametric of sample selection model. Then, the sample selection model is now called a semi-parametric of sample selection model (SPSSM).

In the development of SPSSM modelling using fuzzy concept, as a development of fuzzy PSSM, the basic configuration of fuzzy modelling i.e. involved fuzzification, fuzzy environment and defuzzification (see Muhamad Safiih 2007). For fuzzification stage, an element of real-valued input variables is converted in the universe of discourse into value of membership fuzzy set. In this approach, a triangular fuzzy number is used over all observations. The α - cut method with increment value of 0.2 started with 0 up to 0.8. This is then applied to the triangular membership function to get a lower and upper bound for each observations (x_i, w_i and z_i^*) which is defined as

$$\tilde{w}_{isp} = (w_{il}, w_{im}, w_{iu}), \tilde{x}_{isp} = (x_{il}, x_{im}, x_{iu})$$

and

$$\tilde{z}^*_{i\ sp} = (z_{il}, z_{im}, z_{iu}) \quad (3)$$

In order to solve the model in which occurs uncertainties, fuzzy environments such as fuzzy sets and fuzzy number are more suitable, as the processing of the fuzzified input parameters. Since it is assumed that some original data contains uncertainty, under the vagueness of the original data, the data will then be considered as fuzzy data. That means, each observation considered has variation values. The upper bound and lower bound of the observation are commonly chosen depending on each data structure and the experience of the researchers. For a large size of observation, the upper bound and lower bound of each observation are quite difficult to obtain.

Based on the fuzzy number, a fuzzy SPSSM is built with the form as:

$$\begin{aligned}
&\tilde{z}^*_{i_{sp}} = \tilde{w}'_{i_{sp}}\gamma + \tilde{\varepsilon}_{i_{sp}} \qquad i = 1,\dots,N \\
&d_i = 1 \ \ if \ \ d_i^* = \tilde{x}'_{i_{sp}}\beta + \tilde{u}_{i_{sp}} > 0, \\
&d_i = 0 \ otherwise \ i = 1,\dots,N \\
&z_i = z^*_{ic_{sp}} d_i
\end{aligned} \tag{4}$$

The terms $\tilde{w}_{i_{sp}}$, $\tilde{x}_{i_{sp}}$, $\tilde{z}^*_{i_{sp}}$, $\tilde{\varepsilon}_{i_{sp}}$ and $\tilde{u}_{i_{sp}}$ are fuzzy numbers with the membership functions $\mu_{\tilde{W}_{i_{sp}}}$, $\mu_{\tilde{X}_{i_{sp}}}$, $\mu_{\tilde{Z}_{i_{sp}}}$, $\mu_{\tilde{\varepsilon}_{i_{sp}}}$ and $\mu_{\tilde{u}_{i_{sp}}}$, respectively. Since the distributional assumption for the SPSSM is weak, then, for the analysis of the fuzzy SPSSM, it is assumed that the distributional assumption is weak.

To find an estimate for γ and β of the fuzzy parametric of sample selection model, one idea is to defuzzify the fuzzy observations $\tilde{W}_{i_{sp}}$, $\tilde{X}_{i_{sp}}$ and $\tilde{Z}^*_{i_{sp}}$. That means, converting this triangular fuzzy membership real-value into a single (crisp) value (or a vector of values) that, in the same sense, is the best representative of the fuzzy sets that will actually be applied. Centroid method or the center of gravity method is used i.e. it computes the outputs of the crisp value as the center of area under the curve. Let $W_{ic_{sp}}$, X_{icsp} and $Z^*_{ic_{sp}}$ be the defuzzified values of $\tilde{W}_{i_{sp}}$, $\tilde{X}_{i_{sp}}$ and $\tilde{Z}^*_{i_{sp}}$ respectively. The calculation of the centroid method for $W_{ic_{sp}}$, $X_{ic_{sp}}$ and $Z^*_{ic_{sp}}$ respectively is via the following formula:

$$\begin{aligned}
W_{ic_{sp}} &= \frac{\int_{-\infty}^{\infty} w\mu_{\tilde{w}_i}(w)dw}{\int_{-\infty}^{\infty} \mu_{\tilde{w}_i}(w)dw} = \frac{1}{3}(W_{i_l} + W_{i_m} + W_{i_u}) \\
X_{ic_{sp}} &= \frac{\int_{-\infty}^{\infty} x\mu_{\tilde{x}_i}(x)dx}{\int_{-\infty}^{\infty} \mu_{\tilde{x}_i}(x)dx} = \frac{1}{3}(X_{i_l} + X_{i_m} + X_{i_u}) \\
Z^*_{ic_{sp}} &= \frac{\int_{-\infty}^{\infty} z\mu_{\tilde{z}_i}(z)dz}{\int_{-\infty}^{\infty} \mu_{\tilde{z}_i}(z)dz} = \frac{1}{3}(Z_{i_l} + Z_{i_m} + Z_{i_u})
\end{aligned} \tag{5}$$

Then the crisp values for the fuzzy observation are calculated following the centroid formula as stated above. To estimate γ_{sp} and β_{sp} of SPSSM approach, applying the procedure as in Powell, then the parameter is estimated for the fuzzy semi-parametric sample selection model (fuzzy SPSSM). Before getting a real value for the fuzzy SPSSM coefficient estimate, first the coefficient estimate values of γ and β are used as a shadow of reflection to the real one. The value of $\hat{\gamma}$ and $\hat{\beta}$ are then applied to the parameters of the parametric model to get a real value for the fuzzy SPSSM coefficient estimate of $\gamma_{sp}, \beta_{sp}, \sigma_{\varepsilon_{i_{sp}}}, u_{i_{sp}}$. The Powell SPSSM procedure is then executed using the XploRe software.

The following procedure is to determine the fuzzy data for the observations that assumed involve uncertainty.

Step 1: The original observations $x_1, x_2, \cdots, x_n$ is ordered in terms of non-decreasing values as $x_1 \le x_2 \le \cdots \le x_n$.

Step 2: For some observations that have the same value, the same values are written as one observation. The observation is rearranged as $y_1 < y_2 < \cdots < y_m$, where $m \le n$.

The observation is rearranged to facilitate stating the lower bound and upper bound of all observations.

Step 3: Choose an appropriate lower bound for y_i. The $Lb(y_i)$ is given and chosen in order to construct triangular membership function of y_i. The value of $Lb(y_i)$ is less than and very far from the value of y_i.

Step 4: Determine the lower bound for y_i using formula

$$Lb(y_i) = y_i - \frac{y_i - y_{i-1}}{2}, \text{ for } i = 2, 3, \cdots, m.$$

For a quite large size of observations, it is difficult to choose every one of the lower bounds of observation. So, the lower bound of y_i is defined as the middle value of the observations y_{i-1} and y_i for $i = 2, 3, \cdots, m$.

Step 5: Choosing an appropriate upper bound for y_m. The $Ub(y_m)$ is given and chosen in order to construct triangular membership function of y_m. The value of $Ub(y_m)$ is greater than and not quite far from the value of y_m.

Step 6: Determine the upper bound for y_i using formula

$$Ub(y_i) = y_i + \frac{y_{i+1} - y_i}{2}, \text{for } i = 1, 2, \cdots, m-1.$$

For a quite large number of observations, it is difficult to choose every one of the upper bounds of the observation. So, the upper bound of y_i is defined as the middle value of the observations y_i and y_{i+1} for $i = 1, 2, \cdots, m-1$.

Step 7: Based on the observation y_i together with its lower bound and upper bound, a triangular fuzzy number is constructed as $\widetilde{Y}_i = \left(Lb(y_i), y_i, Ub(y_i)\right)$.

Step 8: For all observations together with the associated lower bound and upper bound, a membership function is defined as

$$\mu_{\widetilde{Y}_i}(y) = \begin{cases} \dfrac{y - Lb(y_i)}{y_{im} - Lb(y_i)} & if\ y \in [Lb(y_i), y_{im}] \\ 1 & if\ y = y_{im} \\ \dfrac{Ub(y_i) - y}{Ub(y_i) - y_{im}} & if\ y \in [y_{im}, Ub(y_i)] \\ 0 & otherwise \end{cases}$$

The triangular membership function may or may not be symmetric. It depends on the values of the lower bound and upper bound of the observation y_i.

Step 9: Following the centroid method, calculate the defuzzified value (crisp value) of $\widetilde{Y}_i$, i.e. Y_{ic} via formula as:

$$Y_{ic} = \frac{\int_{-\infty}^{\infty} y \mu_{\widetilde{Y}_i}(y) dy}{\int_{-\infty}^{\infty} \mu_{\widetilde{Y}_i}(y) dy} = \frac{1}{3}\left(Lb(y_i) + y_i + Up(y_i)\right).$$

From the centroid method, a crisp set value of Y_{ic} is obtained, for $i = 1, 2, \cdots, m$, which is associated with the set of crisp value y_i, for $i = 1, 2, \cdots, m$. Further, a crisp set value is obtained for the observation x_i, for $i = 1, 2, \cdots, n$.

Many defuzzification methods have been proposed and discussed in the literature. That the centroid method is the most prevalent and physically-appealing one as it weights all the values with different possibilities to form a single value. In other words, all possible values have been accounted for.

Step 10: Apply α-cut method into the triangular membership function. Use α-cut by increments of 0.2 starting with 0 up to 0.8.

The α-cut method is applied in order to get another crisp set value for the observation x_i. For

α-cut equals 1, the crisp set value for the observation x_i is the same with the classical model.

Step 11: For each α-cut, obtain another lower bound and upper bound for each observation using formula as

$$Lb(y_{i(\alpha)}) = Lb(y_i) + \alpha\left(y_i - Lb(y_i)\right) \text{ and}$$

$$Up(y_{i(\alpha)}) = Ub(y_i) + \alpha\left(y_i - Ub(y_i)\right)$$

Step 12: Following Step 7, 8, and 9, another crisp values Y_{ic} are obtained.

By following the α-cut method, five sets of crisp value for the observations x_i are obtained.

The effect of the fuzzy data and α-cut on the estimate of the parameters of the fuzzy parametric of sample selection model is stated in 4. When there is no α-cut, the lower bound and upper bound of all the observations are quite far from the observation. It implies that the fuzzy data which is produced via centroid method is also quite far from the middle of the triangular fuzzy number. As the value of α-cut increased from zero to one, the fuzzy data also approached the middle value of the triangular fuzzy number.

Executing the Powell (Powell, 1987) procedure by XploRe takes the data as input from the outcome equation (x and y, where x may not contain a vector of ones). The vector *id* containing the estimate for the first-step index $x_{i_{sp}}^{'}\hat{\beta}$, and the bandwidth vector *h* where *h* is the threshold parameter *k* that is used for estimating the intercept coefficient from the first element. The bandwidth *h* from the second element is used for estimating the slope coefficients. Fuzzy PSSM follows the above procedure then another set of crisp values $W_{ic_{sp}}$, $X_{ic_{sp}}$ and $Z_{ic_{sp}}$ is obtained. Apply the α - *cut* values on the triangular membership function of the fuzzy observations $\widetilde{W}_{i_{sp}}$, $\widetilde{X}_{i_{sp}}$ and $\widetilde{Z}_{i_{sp}}$ with the original observation, fuzzy data without α - *cut* and fuzzy data with α - *cut* to estimate the parameters of the fuzzy SPSSM. The parameters of the fuzzy SPSSM are thus estimated.

5 Conclusion

Normally, modelling plays an important part in estimating the parameters of economic problems. Different from other system designs, the model itself is generated by a mathematical function. In this paper, a description of the development of the FSPSSM has been presented. For handling the uncertainty that is involved in the original data, fuzzy number together with membership function takes an important part and they are derived from expert knowledge.

References:

[1] Chen, T. and Wang. M.J.J. (1999). Forecasting methods using fuzzy concepts, Fuzzy sets and systems.105. p. 339-352.

[2] Cosslett, S. (1991). Semiparametirc estimation of a regression models with sample selectivity, in W.A. Barnett, J. Powell, and G.E. Tauchen (eds), Nonparametric and semiparametric estimation methods in Econometrics and Statistics, Cambridge University Press, p 175-198.

[3] Gerfin, M. (1996). Parametric and semiparametric estimation of the binary response model of labor market participation. Journal of Applied Econometrics. 11. p. 321-339.

[4] Heckman, J.J. (1976). The common structure of statistical models of truncation, sample selection, and limited dependent variables, and a simple estimation for such models, Annals of Economic and Social Measurement, 5, p. 475-492.

[5] Ichimura, H and Lee L..F. (1990). Semiparametirc least square estimation of multiple index models: Single equation estimation, in W.A. Barnett, J. Powell, and G.E. Tauchen (eds), Nonparametric and semiparametric estimation methods in Econometrics and Statistics, Cambridge University Press.

[6] Kao, C and Chyu, C-L. (2002). A Fuzzy regression model with better explanatory power. Fuzzy Sets and Systems 126. p. 401-409.

[7] Khan, S. and Powell, J. L. (2001). Two-step estimation of semiparametric censored

regression models. Journal of Econometrics. 103. p. 73-110.

[8] Klein, R. and Spady,R. (1993). An efficient semiparametric estimator of the binary response model, Econometrica, 61, 2, p.387-423.

[8] L. Muhamad. Safiih (2007). Fuzzy Semi-parametric of a Sample Selection Models. Ph.D. dissertation. University Science of Malaysia. Penang. Malaysia.

[9] L. Muhamad Safiih, A.A. Basah Kamil and M.T. Abu Osman (2006). Fuzzy Approach to Sample Selection Model. WSEAS Transactions on Mathematics. Issue 6. Volume 5. p. 706-712.

[10] Martin, M.F.O. (2001). Parametric and semiparametric estimation of sample selection models: An empirical application to the female labor force in Portugal. Journal of Applied Econometrics, 16, p. 23-39.

[11] Powell, J.L. (1987). Semiparametric estimation of bivariate latent variable models. Social Systems Research Institute. University of Wisconsin-Madison, Working paper No.8704.

[12] Schafgans, M. (1996). Semiparametric estimation of a sample selection model: estimation of the intercept; theory and applications, Ph.D. dissertation, Yale University New Haven, USA.

[13] Stoker, T.M. (1986). Consistent estimation of scaled coefficients, Econometrica, 54. p. 1461-81.

[14] Tanaka, H., Hayashi, S. and Asai, K. (1982). Linear regression analysis with fuzzy model. IEEE Transactions on Systems, Man and Cybernetics, 12. p. 903-907.

[15] Terano, T. Asai, K. Sugeno, M. (1994). Applied fuzzy systems, Cambridge. AP Professional.

[16] Yen, K.K., Ghoshray. S. and Roig. G. (1999). A linear regression model using triangular fuzzy number coefficients. Fuzzy sets and systems. 106. p. 167-177.

[17] Zadeh, L.A. (1965). Fuzzy Sets and systems. In: Fox, J., ed., System Theory. Brooklyn, New York. Polytechnic Press.