Biodiesel and social inclusion: a fuzzy set and regression analysis in the north of Brazil

Marcus Vinicius Alves Finco[1]

Werner Doppler[2]

Abstract: *After the launch of Brazilian Biodiesel use and production Program (PNPB) in 2004, the oil seed production became an economic alternative for family farmers since they might have the chance to improve their income and supply biodiesel enterprises with raw material. Theoretically speaking this picture is easy to draw, but few recent little scientific research was carried out aiming at identifying the socio-economic profile of farmers that are being included in the Brazilian biodiesel chain. Thus, this study aims to assess the links between the family farmers' living standards and the adoption of oil seed activity. Thus, an analysis of Jatropha curcas and Ricinus communis small-scale seed production was done in the Tocantins state, located in the north of Brazil. A fuzzy set theory was applied to construct deprivation indicators based on living standard criteria and a nonlinear probit model was formulated to analyze the inclusion of poor families in the biodiesel chain. The preliminary results point towards a positive relationship between degree of deprivation and adoption of oil seed activity in the case of Ricinus communis system production and a negative relationship in the case of Jatropha curcas system.*

Keywords: *Biodiesel; social inclusion; fuzzy set; regression analysis; Brazil.*

1   PhD in Agricultural Economics (University of Hohenheim, Germany). Lecturer at Department of Business Administration (Federal University of Tocantins, Brazil). Email:marcus.finco@gmail.com
2   Professor at Department of Agricultural Economics (University of Hohenheim, Germany).

Biodiesel e inclusão social: uma análise baseada em conjuntos fuzzy e regressão no norte do Brasil

Resumo: Após o lançamento do Programa Nacional de uso e produção de biodiesel (PNPB) em 2004, a produção de oleaginosas se tornou uma alternativa econômica para os agricultores familiares uma vez que eles tem a chance de aumentar a renda bem como fornecer matéria prima para as indústrias produtoras de biodiesel. Teoreticamente esse quadro é simples de visualizar, mas até recentemente poucos trabalhos científicos foram realizados a fim de identificar o perfil socio-econômico dos agricultores que efetivamente estavam sendo incluídos na cadeia brasileira de biodiesel. O presente estudo, portanto, busca entender as relações entre a condição de vida dos agricultores e a adoção do cultivo da Mamona e do Pinhão Manso como alternativa econômica no meio rural. A teoria dos conjuntos fuzzy foi aplicada para construir indicadores de privação e um modelo não linear probit foi formulado para avaliar a inclusão de famílias pobres na cadeia de biodiesel. Os resultados preliminares apontam na direção de uma relação positiva entre grau de privação e adoção da Mamona e uma relação negativa no caso do Pinhão Manso.

Palavras-chave: Biodiesel, inclusão social, conjuntos fuzzy, análise de regressão, Brasil

Introduction

Global concern for the depredation and exhaustion of natural resources has led governments and scientists around the world to identify alternatives and solutions to the problem. Since the 21st century beginning, an international debate has taken shape, which is currently discussed at 10 out of 10 meetings on sustainable development around the world: pros and cons of biofuels (Dubois 2008; FAO 2008a; FAO 2008c). From a socio-economic point of view, biofuels[3] can positively impact rural development and diversify the utilization of the local environment. For example, enhancing rural space multi-functionality, where farms are used not only for crop production, but also for eco and rural tourism and, of course, generating strategies that reduce and alleviate poverty, so-called 'pro-poor' strategies (UN Energy 2007; FAO 2008b; FAO 2008d).

---

3   Biofuel essentially refers to liquid fuels derived from agricultural crops used for transport and energy generation (FAO, 2008b).

Based on this, in 2004 the Brazilian government launched a biodiesel program, National biodiesel use and production program (PNPB), which was designed based on a scenario of high oil prices, a growing demand for fuels from renewable sources, and the country's comparative advantage in natural resources (Nass *et al.* 2007). The PNPB is a Brazilian government interdepartmental program with goals, such as, implementing production and use of biodiesel while focusing on social inclusion and regional development[4]. Moreover, the PNPB has several specific objectives: (i) implement a sustainable program, promoting social inclusion; (ii) guarantee competitive prices, quality and supply; (iii) produce biodiesel from different vegetable oil species[5].

After 5 years of PNPB implementation, data released by the oil, Natural gas and Biofuels National Agency (ANP) demonstrated that biodiesel in Brazil, at the macro level, is being produced mainly by soybean and animal fat (roughly 80% and 15%, respectively) (ANP, 2010) which are, in most cases, produced by large-scale farmers. The impacts of PNPB at a micro level, however, are uncertain. Especially regarding the effectiveness of implementing social inclusion, i.e. there is uncertainty about whether poor rural families are being included in the national biodiesel chain. Social inclusion is a process that aims to offer opportunities to access goods and services to **the most excluded people within a system** and therefore that benefit all and not only those in better conditions (Silver & Miller 2002; Pierson 2002) [Emphasis added].

In this context, the present study aims to analyze the relationship and linkages between poor rural families and the adoption of small-scale oil seed production from two different farming systems in the north of Brazil. Therefore, assessing whether PNPB's 'pro-poor' strategy is effective or not. The research will support the regional and national governments in improving the PNPB and social inclusion targets. Moreover, the results presented can play an important role in integrating efforts to consolidate a new (and 'clean') economic alternative for the small-scale farmers.

---

4   According to Garcez and Vianna (2009), the biodiesel policy has the clear objective of promoting the social inclusion of family farming. This is a very commendable and much needed proposition, specifically in Brazil, a country that has one of the worst distributions of land in the American continent and great disparities between the wealth of large, industrial-scale farming versus family-based counterparts.

5   The PNPB has also a differential to other biodiesel programs created during the 1980's in Brazil, such as, the National program of vegetable oils to energy production (Pr**ó-Ó**leo) and the Program of vegetable oils (OVEG). Apart from the production and use of biodiesel in the country, the PNPB focuses mainly on social inclusion and also takes into account the environmental aspects of biodiesel production. Due to importance of the social inclusion target, the national government established a range of different instruments aiming at fulfilling this target, such as, fiscal exemptions, financial subsidies, and specific credit lines to the biodiesel companies.

1. Research area and methodological aspects

The research was carried out in Tocantins State located in northern Brazil and known as the Brazilian Legal Amazon. The State is situated in a transition area, containing climate and vegetation patterns from both the Amazon rain forest (15% of the territory) and the Cerrado (85% of the territory) (IBAMA 2004). This transition area, so-called Ecotone zone, is the home to traditional communities (family agriculture, indigenous, as well as, quilombolas[6]) and comprises rich biodiversity, which is responsible for numerous environmental services. For this reason, scientific studies and research in the area are extremely important. They are often focused on understanding the different farming systems and their connections to the local economy and the very diverse environment.

The foundation for the database was formed through a comprehensive survey, which was carried out between April and September 2008 in two sub-study regions within Tocantins State. In one sub-study region, *Ricinus communis* (castor bean and, also well known as, mamona in Brazil) oil seed is cultivated and in the other sub study-region *Jatropha curcas* (well known as pinhão manso in Brazil) is cultivated. Specific questionnaires were applied to smallholders, who were randomly selected: 27 in the case of *Jatropha curcas* farming system and 25 in the case of *Ricinus communis* farming system. It is important to highlight that the selection of smallholders followed statistical procedures and that the sample can be considered representative since it comprises more than 90% of small-scale oil seed producers in the region in question at the time the research was carried out.

Aiming at better understanding the links between the farmers' living standards and the adoption of oil seed activity and, therefore, checking the effectiveness of PNPB's 'pro-poor' strategy, the *fuzzy* set theory was chosen, and non-parametric tests, as well as, a nonlinear *probit* model were used to demonstrate this relationship with more accuracy. For this purpose, the software XL Stat was used to support the *fuzzy* analysis, and the software STATA was used to support the statistical and the econometric analysis.

---

6   Descendants of African slaves.

2. The *fuzzy* set theory

According to Doppler (1993, 2004), the farming system is as a complex system, which includes several features and is characterized by interdependence and interactions between different agents. These interactions usually result in nonlinear effects, which are difficult to assess due to the behaviors that appear when small pieces are reassembled into the larger systems they comprise. Thus, farming systems should be studied in a holistic and systemic way and the outputs of any treatment must be palatable to the final user.

Therefore, one of the most promising avenues on farming systems research seems to rest in the use of various measures of uncertainty and information that emerge from the broad framework of *fuzzy* set theory (Reys 2003). Thus, when assessing *fuzzy* membership, the researcher's goal is to assert each case's degree of membership in a set and not only to determine its position on a continuum relative to other cases, as in variable-oriented social research. Regarding farming system studies, for example, instead of measuring the household income (quantitative variable), a researcher might want to emphasize a degree of membership in the set of households that are financially secure (RAGIN, 2000)[7]. In this context, the concept of *fuzzy* set is related to a class with a continuum of grades of membership. According to Zadeh (1965), the framework generated by the *fuzzy* set provides a natural way of dealing with problems. Imprecision arises from the absence of a sharply defined criterion of class membership rather than the presence of random variables[8], which can be the case in farming systems research (Reys 2003; Blair 2007).

Based on this, an expression in which linguistic variables are related to linguistic terms represents a linguistic statement in *fuzzy* logic, and the determination of matching input variables with the linguistic terms is referred to as *fuzzification* (Zadeh 1972). To this end, the actual degree of membership of input variables is determined for each linguistic term of the corresponding linguistic variable.

Data from the interviewers, for instance, are tabled and the process of inference starts, which consists of three sub-functions: aggregation, activation and accumulation. Finally, after the inference process, which provides a membership function as a result, one can convert this result in a crisp number through the *defuzzification* (Zadeh 1965; Zadeh 1972),

7   It is important to say that, so far, it is not surprisingly to recognize that there are few scientific works dealing with *fuzzy* set theory and farming systems approach. Some of them were carried out by Reys (2003), and Blair (2007), whose work has great influence in the present study.
8   *Fuzzy* set theory presented by Zadeh (1965) describes physical variables using linguistic values and expressions.

and the crisp number generated by the *defuzzification* process, in turn, should provide a good representation of the information contained in the *fuzzy* set (figure 1).
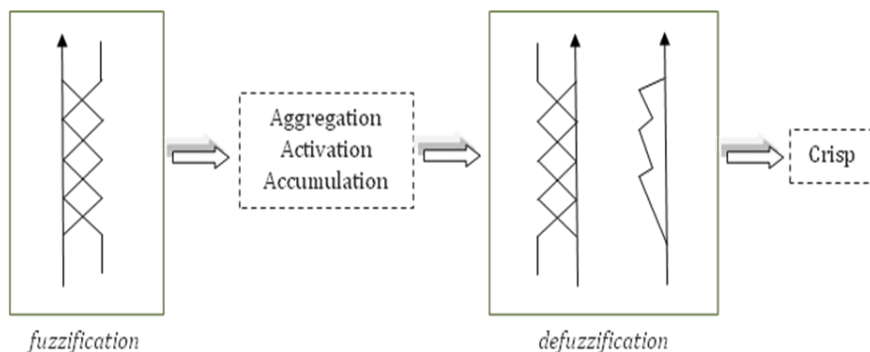
FIGURE 1: INFERENCE PROCESS.

SOURCE: ADAPTED FROM REYS (2003).

3. The *fuzzy* set theory and its applications on poverty issues

The *fuzzy* set theory was selected because it enables one to construct a multi-dimensional poverty index based on living standard criteria approach that is used to assess farming systems (Doppler 1993; Doppler 2004). A major advantage of using a multi-dimensional measure of poverty over a traditional one (poverty line) is that it not only takes account of the material situation of individuals, but it also captures their general living conditions (Miceli 1998). Often in the traditional approach, poverty is measured by a poverty line, i.e. all individuals whose income does not exceed a certain level of income are stated to be poor (Costa 2002; Deutsch & Silber 2005; Betti & Verma 1998). According to Cerioli and Zani (1995) this approach has clear merits, taking not just the relative number of poor, but both their absolute and relative de-privation. Nevertheless, it fails to capture a number of features, which may be relevant in thoroughly understanding poverty.

In this context, the *fuzzy* set theory can be considered a suitable approach for

one strong reason: it allows the construction of a multi-dimensional index without a sharp division between poor and non poor populations, i.e. the degree of person's membership can vary between 0 (non membership or fully out) and 1 (full membership or fully in). In addition, *fuzzy* is a broad analysis including a range of indicators for living standards, while also adopting mathematical tools that can represent the complex character and nuances of the multi-dimensional phenomenon (Cheli & Lemmi 1995).

For example, let X be a set of elements $x \in X$. Then, a *fuzzy* subset A of X is a set of ordered pairs:

$$A = \{x, \mu_A(x)\} \tag{1}$$

Where μ_A is a mapping from X to the closed interval [0,1], so that each value $\mu_A(x)$ represents the grade of membership of x in A. In other words, the *fuzzy* subset A of X is characterized by a membership function $\mu_A(x)$ associating a real number in the interval [0,1] to each point of X. The value $\mu_A(x)$ represents the degree of belonging of x to A.

If A is an ordinary set, its membership function can then take only the values 0 and 1, and for this reason is also called a *crisp* set. In that case:

$$\mu_A(x) = \begin{cases} 1 \ if \ x \in A \\ 0 \ if \ x \notin A \end{cases} \tag{2}$$

Similarly, when A is a *fuzzy* set, if $\mu_A(x)$ = 0 then x does not belong to A, whilst $\mu_A(x)$ =1 if x fully belongs to A. Elements x such that 0 < $\mu_A(x)$ < 1 belongs to A only to a certain degree, and their membership in the *fuzzy* subset increases the closer the values $\mu_A(x)$ are to 1.

Therefore, because the concept of poverty is not as sharp, as it is sometimes assumed to be, the same procedure can be used to define the *fuzzy* set of the poor. Let N = {1,...,n} be the set of individuals in a population and P be the *fuzzy* subset of the poor, defined as follows:

$$P = \{i, \mu_p(i)\} \tag{3}$$

With $i = 1,...,n$ and where $\mu_p(i)$ represents the degree of membership of each individual, i, to the *fuzzy* subset of the poor and denote for each individual a grade of membership in the poor set. Then:

$$\begin{cases} \mu_p(i) = 0 \\ \mu_p(i) = 1 \\ 0 < \mu_p(i) < 1 \end{cases}$$

(4)

The basic problem of this approach is providing an appropriate definition of the membership function. Therefore, it is unclear whether an individual is 100% poor, only poor to a certain degree, or not poor at all. (Lelli 2001; Maertinetti 2000).

In this context, linear membership functions are often used in common empirical analyses as they are easy to specify, interpret and visualize. These representations presuppose the variables' modalities to be equidistant from one another and assume a direct proportionality between the elements of the domain and the membership grade and are usually used to represent polytomous (qualitative in nature) variables.

If the variable X_j is measured on an ordinal scale, for example, the different amount of deprivation inherent in its increasing levels must be taken into account. According to Cerioli and Zani (1995), this can be readily done by assigning scores $\psi_j^{(1)}$ $(r = 1,2, ... s_j)$ to the s_j ordered categories of X_j, so that:

$$\psi_j^{(1)} < \cdots < \psi_j^{(r)} < \cdots < \psi_j^{(s_j)}$$

(5)

A simple and common specification for such scores consists of the first s_j integers:

$$\psi_j^{(r)} = r \quad r = 1, ... , s_j$$

(6)

Definition (6) is appropriate whenever the categories of X_j may be supposed

to be equally spaced.

Based on this, Ξ can be considered the subset of households undergoing some privation according to a set of indicators ξ_j, with j = 1,...,*k*. So, regarding ξ_j it is possible to state a modality associated with such bad living conditions that poverty couldn't be denied. When ξ_j takes the value one, for instance, an individual is found to be fully deprived regarding the service or good *j*. However, when it has the value zero, an individual is considered non poor regarding the same service or good *j* in question.

If ψ_j^{min} and ψ_j^{min} are the scores corresponding to the boundaries for an individual to be considered non poor and poor, respectively, i.e.when ψ_j^{min} corresponds to better living conditions and ψ_j^{max} corresponds to worse living conditions, the memberships can be expressed as:

$$u_1(\Xi_1(i))\ (i) - \{\blacksquare(0 \qquad\qquad if\ \Psi_{1ij} - \Psi_{ij}^{min} @ (\Psi_{1ij} - \Psi_{ij}^{min})/$$

$$(7)$$

Where Ψ_{ij} is the score for an individual *i* ensuing from indicator ξ_{ij}. With this specification, the membership function increases linearly as the risk of poverty rises.

Alternatively, the trapezoidal specification is usually applied to continuous variables (quantitative in nature) and postulates that the membership function to the poor set can be defined by fixing a value *y'* up to which an individual is definitely poor, and a value *y"* above which an individual is definitely not poor, for instance:

$$u_1(\Xi_1(i))\ (i) - \{\blacksquare(1 \qquad\qquad if\ 0 \leq y \leq y' @ (y^{1''} - y)/(y^{1''} - y')$$

$$(8)$$

In this case, below *y'* the individual is considered fully in (poor) and above *y"* the individual is considered fully out (non poor). For values between *y'* and

y'' the membership function takes on values in [0, 1], and it seems natural to make it a decreasing function of the deprivation indicator. Therefore:

$$f(y'') = 0 \ and \ \lim_{y \downarrow y'} f(y) = 1 \qquad (9)$$

After the establishment of the memberships of each indicator $\xi_j \xi_j$, the degree of belonging for each household to the *fuzzy* subset of the poor is taken by the weighted arithmetic mean of their degrees of memberships to the set of deprived people, according to k indicators. Poverty is then regarded as an accumulation of deprivation situations where k *fuzzy* sets Ξ_1, Ξ_2,..., Ξ_k are defined over the set of individuals.

$$u_p(i) = \sum_{j=1}^{k} \omega_j u_{\Xi_j}(i) \qquad (10)$$

An appropriate system of weighting could be found by making each weight an inverse function of the number of individuals in the reference population which shows the corresponding poverty symptom, i.e. an inverse function of the average deprivation level:

$$\omega_j = \frac{ln\left(\frac{1}{\overline{u}_{\Xi_j}}\right)}{\Sigma_{j=1}^{k} ln\left(\frac{1}{\overline{u}_{\Xi_j}}\right)} \qquad (11)$$

The system of weights presented in equation (11) gives more importance to those indicators of deprivation associated with less frequent symptoms of

poverty, where the choice of the logarithm is justified by not giving importance to the modalities with a very low frequency (Martinetti 2000). This can be also justified by the fact that, in a relative definition of poverty, people have a stronger feeling of deprivation when they do not own a widespread good (Miceli 1998).

The equation (12) represents the fuzzy proportion of deprived individual according to indicator ξ_j.

$$\bar{u}_{\xi_j} = \frac{1}{n} \sum_{i=1}^{n} \bar{u}_{\xi_j}(i) \qquad (12)$$

Based on this, a general index of poverty is constructed by averaging the individual membership functions with *fuzzy* index of poverty (FIP) $\in$ [0, 1], which represents the proportion of individuals "belonging", in a fuzzy sense, to the poor subset.

$$FIP = \frac{1}{n} \sum_{i=1}^{n} u_p(i) \qquad (13)$$

According to Cerioli and Zani (1995), the *fuzzy* approach mentioned hitherto has the merit of providing a more realistic portrayal of poverty than the traditional income-based dichotomy of poor and non poor. Moreover, the *fuzzy* set theory enables one to capture many facets of an individual standard of living and emphasizes the different degree to which each subject is considered to be poor (Bantilan *et al.* 1992).

4. The *probit* model[9]

After the establishment of the *fuzzy* set theory and the factor analysis, the next step is to select an econometric model to analyze the PNPB's social inclusion target and, therefore, its effectiveness as a 'pro-poor' strategy. In this context, the analysis was done by the estimation of a nonlinear regression model because the dependent variable is a binary variable, which expresses the adoption or the non-adoption of the oil seed activity by the smallholders.

In the presence of binary dependent variables, the simplest model is the linear probability model, estimated by the ordinary least square method. In this case:

$$\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \qquad (14)$$

p being the probability of y, assuming a value equal to one:

$$E(y) = p = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \qquad (15)$$
$$E(y) = p = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

In the linear probability model, the estimated coefficients from the regression, which express the effect of unit variations on the independent variables on the probability of the dependent variable, have a value equal to one. The problem is that these effects are constant and as x_i increases p also increases (when β_i is positive, otherwise it continues to decrease) under a constant ratio. However, since $0 \leq p \leq 1$, it is impossible to have a constant increase rate. In addition, the linear probability model also presents heteroskedastic errors and, therefore, the estimated coefficients are not sufficiently robust, and therefore the hypothesis tests and confidence intervals may not be valid.

9   This section is based on Greene (2008) and Hill *et al.* (2008).

Aiming to overcome those problems, one can choose between the nonlinear *probit* and *logit* models. In these models the slope is not constant and the probabilities are restricted to the [0, 1] interval based on probability density functions. The *probit* function is related to the standardized normal probability distribution, whereas, the *logit* function is related to the logistic distribution. As these models are nonlinear models, the coefficients estimation should be done by the maximum likelihood method. In general, the estimated coefficients in *probit* and *logit* models are slightly different and the choice between them might be done based on the researcher convenience. On the present study the *probit* model was chosen considering that the errors have a normal distribution.

In this context:

$$p = F(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) \qquad (16)$$

Where p is the probability of y take value equal to one and F is the *probit* function, nonlinear in the β_i. The effect on a unit variation in x_i about the probability p of y to be equal to one is given by:

$$\frac{\partial p}{\partial x_i} = \frac{\partial F_{(t)}}{\partial_{(t)}} \cdot \frac{\partial t}{\partial x_i} = f(.)\beta_i \qquad (17)$$

And

as $f(\cdot)$ is a probability density function, and its value is always positive. Moreover, the sign of $\partial p/\partial x_i$ is determined by the sign of β_i;

as x_i varies, the value of $f(\cdot)$ also does. Therefore, the effect of a unit variation in the independent variables on the probability of the dependent variable relies in the independent variables' levels. Hence, aiming at estimating the coefficients' effects through the *probit* model it is necessary to choose some level to the independent variables as a reference.

5. Variables definition

Several approaches have been adopted in the field of poverty research in different regions and contexts, and comprise both quantitative and qualitative aspects. The poverty condition has been defined broadly as a phenomenon of multiple dimensions with economic, cultural and social aspects and it is characterized by: (a) insufficient income; (b) limited access to basic needs and (c) social exclusion and discrimination due to the ethnic or gender origin (ECHEVERRIA, 2000). The measurement of poverty through indicators that have only one dimension, such as, income can lead to discrepancies in the interpretation of a broader phenomenon[10].

To obtain more accurate analyses of the linkages between farmers' living standard and oil seed activity adoption, social, as well as, economic indicators were taken into account. Therefore, variables related to economic and social aspects of households were selected to better demonstrate the multi-dimensional poverty situation in both sub-study regions in question. The variables selected are described below.

5.1 Family income, per capita

Family income, *per capita*, was selected as standard of living indicator since it is the sum of on-farm and off-farm income divided by the members of the family and, thus, gives a picture of the economic situation of the family.

So, based on the *fuzzy* set theory presented previously, one should choose the lower and upper anchors (scores) where, in this case, an individual can surely be considered poor and non poor, respectively. In this case, family income *per capita* is a quantitative variable by nature, and the lower anchor is the value of R$2400,00 (Brazilian reais)[11] per year, which is the half of minimum salary established by the national government. According to a recent research (PNUD, 2009), a person below this value can be considered poor.

10   Due to the complexity of the poverty concept and its different perceptions and definitions, there is a high variation in estimated poverty incidence even in studies that were carried out in the same region (BROAD, 1994).
11   At the time this study was conducted one Brazilian Real (R$) was equivalent to 0.5 American Dollars (US$).

On the other hand, the upper score is established by multiplying the lower score by two or R$4800,00 per year, i.e. the minimum salary established by law and which reflects, in the present study, that above this value a person cannot be considered poor.

5.2 Ownership of durable goods

The variable ownership of durable goods is an ordinal variable and ranges from 1 to 5, where 1 means that the family owns all the five goods in question - television, refrigerator, telephone, gas stove and a small wash machine - and 5 when the family owns just one of these items. In this case, the lower score is 1 and the upper score is 5.

5.3 Transport mean

The variable, transport mean, is an ordinal variable and ranges from 1 to 5. Where 1 reflects that the family has the car as their main mean of transportation; 2 when the family has the motorcycle as the main transport mean; 3 when has the bicycle is the main transport mean; 4 when the family does not possess any kind of transportation, but can afford a bus ticket and 5 when the family does not possess any means of transportation and cannot afford a bus ticket. Here, the lower and the upper scores are 1 and 5, respectively.

5.4 House infra-structure

The house infra-structure variable is an ordinal variable and ranges from 1 to 3, where 1 represents those families who have a toilet linked to sewer, as well as, systems of water piped into dwelling; 2 when families possess only one of them; and 3 when families possess none of them.

Herein the scores are established in the following way: families who have both facilities are considered non poor or fully out and thus receive value equal to zero. Families who do possess none of the facilities are considered poor or fully in and thus receive value equal to one. And therefore, families who possess just one of the facilities mentioned are considered half member (or half deprived) and receive value equal to 0.5.

5.5 Health condition

The variable, health condition, is ordinal and the values range from 1 to 3, where 1 reflects that the household is not sick (good health condition); 2 when the family is sick, but it is not considered severe (regular health condition); 3 when the household is severely sick (bad health condition).

Here the scores follow the same procedure of the variable social capital where the families who have good health condition are considered non poor or fully out and receive a value of 0. The families who have bad health condition are considered poor or fully and receive a value of 1. The families which have regular health conditions are found to be in the middle and receive a value of 0.5.

5.6 Resource dependence

The variable resource dependence is a dichotomous variable by nature and, thus, the families have been split into poor or full member (1) when they rely on the municipality and/or on the association's machinery to run the farm activities; and non poor families (0) when the families do not depend on the municipality and/or on the association's machinery to run the farm activities.

6. Results and discussion

In this section, the index of poverty estimated through the *fuzzy* set theory and based on standard of living indicators is presented. In addition, the results related to the linkages between the *fuzzy* index of poverty and the adoption of oil seeds in both sub-study regions, estimated by nonlinear *probit* models are also discussed. Nevertheless, it is important to indicate that the *fuzzy*

model was applied to each living standard indicator (herein called indicators of deprivation) with the aim of estimating the farmers' degree of membership. For this purpose, each indicator follows its own nature, which, in this case, is separated by dichotomous, polytomous and continuous. The trapezoidal model is related to equation (8), the linear model is related to equation (7) and the traditional model is related to equation (2) presented previously.

TABLE 1. NATURE OF INDICATORS OF DEPRIVATION (ξ_j)

Indicators of deprivation ξ_j	Definiton	
	Indicator's nature	***Fuzzy model***
1. Family income per capita	Continuous	Trapezoidal
2. Durable goods	Polytomous	Linear
3. Transport mean	Polytomous	Linear
4. House infra-structure	Dichotomous	Linear
5. Health situation	Polytomous	Linear
6. Resource dependence	Dichotomous	Traditional

Source: own elaboration.

It is worth to state, *ex ante*, that whenever one applies *fuzzy* set theory and chooses different types of models (trapezoidal, linear and traditional models in the present case) and different lower and upper scores, a sensitivity analysis would be advisable in order to check the robustness of the results (Miceli 1998; Lelli 2001). Based on this, two types of sensitive analysis were done in the present study. First, different *fuzzy* models, as well as, different upper and lower scores were applied to the indicators. The results point to similarities in the memberships' degree and, thus, the models and the scores presented herein could be considered robust.

Second, the *fuzzy* index of poverty was used as a dependent variable in two different regression models: (1) a linear model; and (2) a censored *tobit* model (corner solution model) since there is a grouping of observations near zero (limit point). The objective was to check whether the sign, as well as, the significance of the independent variables follow the economic theory. Again, the results from the regressions point out that the *fuzzy* index of poverty estimated in the present study can be considered robust (data not shown).

As the objective of the present study is to assess the effectiveness of the PNPB's social inclusion target and, therefore, its 'pro-poor' strategy, i.e. the inclusion of poor families in the biodiesel chain through the small-scale oil seed activity, the *fuzzy* index of poverty was estimated separately for the *Ricinus Communis* sub-study region and the *Jatropha Curcas* sub-study region. The analysis per sub-study region is necessary since the smallholders are located in different areas, produce different oil seeds and have contracts with different biodiesel companies[12]. Therefore, the weights for each indicator of deprivation were calculated per sub-study region, taking into account the peculiarities and nuances and therefore the relative importance of each indicator of deprivation in each region.

Thus, after the application of a *fuzzy* model to a specific indicator, a *fuzzy* index of poverty (FIP) was estimated, as well as, the farmers' degree of membership according to indicators of deprivation stated previously can be seen in Table 2.

TABLE 2. *FUZZY* POVERTY DECOMPOSED BY *RICINUS* AND *JATROPHA* REGIONS

Indicators of de-privation ξ_j	*Ricinus communis* region		*Jatropha curcas* region	
	Fuzzy proportion of poor $\overline{u}(\xi_j)$	Weights ω_j	*Fuzzy* proportion of poor $\overline{u}(\xi_j)$	Weights ω_j
1. Family income per capita	0.7802	0.0535	0.6423	0.0825
2. Durable goods	0.2750	0.2782	0.3431	0.1992
3. Transport mean	0.4200	0.1870	0.3922	0.1743
4. House infra-structure	0.5400	0.1328	0.4804	0.1365
5. Health situation	0.3103	0.2524	0.3824	0.1790
6. Resource dependence	0.6412[a]	0.0962	0.2941[b]	0.2279
FIP	**0.4083**		**0.3907**	

Note: FIP is the *fuzzy* index of poverty.

12   A one-year-contract in the case of *Ricinus communis* producers and a ten-year-contract in the case of *Jatropha curcas* producers.

Different letters show significant statistical differences according to the Mann-Whitney test (p < 0.05).

The results show, for instance, that, on average, 27% and 34% of the families can be considered deprived regarding the possession of durable goods in the RC and JC region, respectively. Moreover, around 31% and 38% of the households do not present a good health condition (RC and JC, respectively). A considerable (and statistical) difference between regions can be seen regarding the indicator "resource dependence". Here, 64% of farmers in RC region depend upon external resources to carry out farm and family activities compared to 29% of farmers in the JC region (*p-value* < 0.05).

It is important to notice that the lower the number of households deprived regarding a certain indicator, the higher is the weight of this indicator in the FIP composition. As mentioned before, this happens due to the fact that the system of weights used in the present study gives more importance to those indicators of deprivation associated with less frequent symptoms of poverty, since people have a stronger feeling of deprivation when they do not own a very widespread good.

The interpretation of the final index is difficult because it combines indicators of different natures, but it can still give an insight into the major areas of living standard conditions considering groups of indicators, such as, housing infra-structure conditions, the possession of durable goods and resource dependence, as well as the final degree of deprivation. Nevertheless, even though the interpretation of the *fuzzy* index of poverty is not straightforward, it gives a glimpse to the degree of deprivation (poverty) in both sub-study regions. When one considers the overall *fuzzy* index of poverty, for instance, the result of 0.4083 corresponds to a significant 40.83% average rate of deprived families in the *Ricinus communis* region and 0.3907 or 39.07% average rate of deprived families in *Jatropha curcas* region, which is a value that should not be neglected.

As the objective of the present study is to assess the differences among farmers within each region, i.e. estimate the linkages between deprivation and adoption of oil seed activity, after the estimation of the *fuzzy* index poverty for both sub-study regions, the farmers were split into "poor" and "non-poor", according to their own *fuzzy* index of poverty (FIP). Based on this, in the case of *Ricinus communis* region, for example, the families who have, on average, a FIP above 0.41 were considered poor and in the case of *Jatropha Curcas* region, those families who have a FIP above 0.40 were also considered poor.

In addition, aiming at understanding the linkages between the farmers' *fuzzy* index of poverty and the adoption of oil seed activity and, therefore, their social inclusion (or not) in the biodiesel chain, a nonlinear *probit* model was formulated to estimate the relationship between the rural poverty and adoption of oil seeds activity in both sub-study regions in question. In this context, two regressions were estimated: one for the *Ricinus Communis* region and one for the *Jatropha Curcas* region.

The *probit* model formulated was:

$$Y_j = \alpha + \beta X_j + \mu_j \qquad (j = 1, \dots, n) \qquad (18)$$

Where the dependent variable *Y* refers to the adoption of oil seeds (*Y*=1) or otherwise (*Y*=0), α and β are the parameters of the equation, μ is the stochastic term, and the independent variable *X* is a *dummy* variable where 1 refers to when a family is considered poor and 0 when a family is considered non poor (based on the *fuzzy* index of poverty previously mentioned).

Thus, in the case of *Ricinus communis* region:

$$X = \begin{cases} 1 \ if X_{ij} > 0.41 \\ 0 \ if \ otherwise \end{cases} \qquad (19)$$

And in the case of *Jatropha curcas* region:

$$X = \begin{cases} 1 \ if X_{ij} > 0.40 \\ 0 \ if \ otherwise \end{cases} \qquad (20)$$

The results of regression can be seen in Table 3.

TABLE 3. REGRESSION RESULTS FOR NON LINEAR *PROBIT* MODEL

	Ricinus Communis			Jatropha Curcas		
	region			region		
	Coef.	SE	z-sta-tistic	Coef.	SE	z-sta-tistic
Constant	-0.3308	0.2460	-1.64**	0.4948	0.3697	2.03**
FIP	0.7220	0.3643	1.98**	-0.9676	0.2433	-2.62*
Number obs	50			51		
LR χ^2	4.00**			7.08*		

SE = standard error

* Significance level of 0.01; ** Significance level of 0.05.

It can be observed that the coefficients, in both sub-study regions, are significant at 1% and 5%. The likelihood ratio Chi-Square is also significant at 1% and 5%. Regarding the signs of the coefficients in the RC region, for instance, one can notice that the sign of the coefficient FIP is positive; pointing out that a poor family has, on average, a higher probability to adopt the oil seed activity than a non poor family in the same region. This outcome suggests that in this region the social inclusion advocated by the PNPB is taking place, or in other words, the poor families are being included in the Brazilian biodiesel chain. In addition, the changes in the predicted probability for the *dummy* variables were also estimated and shown that being a poor family, in the *Ricinus communis* region, increases the probability of adopting oil seed activity (and therefore be included in the national biodiesel chain) by 28%, on average.

However, when one observes the *Jatropha curcas* region, one notices that the sign of the FIP coefficient is negative, i.e. the poor families have, on average, a lower probability to adopt the oil seed activity when compared to their counterparts who are considered non poor in this region. Here, it is clear that the social inclusion is not occurring since the more deprived families are not taking part of the biodiesel production process and, therefore, the effectiveness of the PNPB's 'pro-poor' strategy cannot be seen. The changes in the predicted probability for the *dummy* variables shows an opposite trend compared to RC region. In JC region, for instance, being a poor family decreases the probability of adopting oil seed activity (and therefore be a part of PNPB social inclusion target) by 37%, on average.

One possible explanation for the results from the *probit* model relies on the short horizon outlook and risk aversion embedded in the poverty condition. The bad living standards, e.g. the weak access to goods and services, and, therefore, the high vulnerability to disasters (natural and man-made) make the poor uncertain and insecure about the future. In addition, the risk aversion began from the limited assets available and the low educational level of family members. This put poor rural families into a precarious economic balance, where the families prefer short-term activities rather than medium to long-term projects (Reardon & Vosti 1995). This seems to be the case of *Ricinus communis* oil seed production. *Ricinus* is an annual crop and its production is based in one-year-contract with the biodiesel company, which allows the farmers to quit the activity in the following year in case of bad harvest and low productivity of the feedstock. However, this is not the case of *Jatropha curcas* oil seed. *Jatropha* is a perennial crop and its activity is based on a ten-year-contract with the biodiesel company. The oil seed becomes economically profitable after the fourth year and, thus, requires a diversified portfolio from the farmers' side. Therefore, the long-term horizon in the case of *Jatropha* production seems to be more attractive to the farmers in the region in question, which present, on average, a broad spectrum of capital assets.

In summary, apart the fact that the Brazilian biodiesel use and production program (PNPB) does possess the laudable objectives of promoting social inclusion and regional development, the results presented hitherto demonstrate and suggest that a lack in specific policy mechanisms has resulted in the nonfullfillment of one of its main targets: inclusion of marginalized people in the biodiesel chain. Based on this, it is imperative to say that the involvement of local agents seems to be a *sine qua non* condition to overcome the shortcomings in policy mechanisms by aiming to minimize the gaps in the national biodiesel policy and therefore reach the social inclusion and the sustainable development advocated by the PNPB.

Final remarks

The present study presents an example of the application of a multi-dimensional measurement of poverty using *fuzzy* sets. The empirical results obtained for Tocantins state, located in the north of Brazil, show that the use of several indicators not only helps in giving a more comprehensive and complete picture of living standards, but also gives an image of rural poverty that is closer to what is perceived by just observing reality. Nevertheless, for many

reasons, the results depend strongly on the choice of indicators for this kind of analysis, so it would be worth selecting and defining an appropriate set of indicators to include the data, covering all the relevant areas for a broader analysis of living conditions.

The outcomes from the present study contribute to society, in general, by enabling one to integrate efforts on the current debate about bioenergy and sustainable development in the country. The debate about bioenergy and social inclusion is currently considered a hot topic and the study reveals the dynamicity of the linkages between two multi-dimensional phenomenons. The ambiguity of this relationship, highlighted by the results of the *fuzzy* set theory and the *probit* model illustrates that the Brazilian program of biodiesel use and production (PNPB) target of social inclusion cannot yet be seen as successful. In one sub-study region, for instance, it is clear that the poor families are being included in the Brazilian biodiesel chain, but in the other sub study-region the poor families continue to be marginalized without access to this new economic alternative. Based on this, we strongly suggest that other studies be carried out, aiming at better understanding the reasons behind the inclusion and non-inclusion of farmers in the biodiesel chain.

This study is unprecedented in the northern region of Brazil and the results are extremely important in helping to obtain an appropriate method for regional and national governments to subsidize clean energy production with social inclusion. It is also important to say that the present study contributes to the dissemination of the application of *fuzzy* set theory to farming systems, especially in a region that faces a significant lack of scientific research. Although the present research focuses on Tocantins State, small-scale oil seed production can now be better gauged in other parts of the Brazilian Legal Amazon because our study highlights one of the most discussed topics in bioenergy debate: oil seed production and its linkages with farmers' living standard. As the present study focuses only at family level in a specific region, we also suggest that further studies should be carried out that are focused on biodiesel production, on both regional and national levels, in order to attain a broader idea of the biodiesel production in Brazil.

Acknowledgments

References

ANP (National Agency of oil, natural gas and biofuels). (2010). *Biodiesel statistics. URL:* www.anp.gov.br Acesso em: 10/03/2010.

BANTILAN, M. C. S.& BANTILAN Jr, F. T.& CASTRO, M. M. (1992). Fuzzy sub set theory in the Measurement of Poverty. *Journal of Philippine Development,* 19 (34).

BETTI, G. & VERMA, V. K. (1999)."Measuring the degree of poverty in a dynamic and comparative context: a multi-dimensional approach using fuzzy set theory**".** In Proceedings of the ICCS-VI, Lahore, Pakistan, August 27-31,Vol. 11, pp. 289-301.

BLAIR, R. A. (2007) "An Assessment of Agricultural Risk and Diversification on Farming Familie's Living Standard under Fuzzy Conditions. A case from Guyana." In DOPPLER, W. & BAUER, S. (eds.) *Farming and Rural Systems Economics.* Vol. 90. Weikersheim Margraf Verlag.

BROAD, R. (1994) "The poor and the environment: friends or foes?**" W***orld Development* 22(6): 811-22.

CERIOLI, A. & ZANI, S. (1995) "A Fuzzy Approach to the Measurement of Poverty." In DAGUM, C. & ZENGA, M. (eds.). *Income and Wealth Distribution, Inequality and Poverty.* Berlim/Springer: Verlag, p.272-284

CHELI, B. & LEMMI, A. (1995) "A "Totally" Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty." *Economic Notes* 24: 115-134.

COSTA, M. (2002) "A Multidimensional Approach of the Measurement of Poverty." *IRISS Working Paper Series n. 2002-05.* URL: http://iriss.ceps.lu/documents/irisswps28.pdf. Acesso em: 26 October 2009.

DEUTSCH, J. & SILBER, J. (2005) "Measuring Multidimensional Poverty: An Empirical Comparison of Various Approaches." *Review of Income and Wealth*, 51 (51).

DOPPLER, W. (1993) "Contribution of the Farming Systems Approach to Regional Food Security and Rural Infrastructure." In THIMM, H. U. & HAHN, H. & REINHARD, I. & SCHULZ, D. (eds.) *Regional Food Security and Rural Infrastructure*, Vol. I: Münster-Hamburg: LIT Verlag.

DOPPLER, W. (2004) *Farming and Rural Systems Approaches.* Published Lecture Material. Hohenheim University, Stuttgart, Germany.

DUBOIS, O. (2008) *How Good Enough Biofuel Governance Can Help Rural Livelihoods: Making sure that Biofuel Development Works for Small Farmers and Communities.* Rome: FAO.

ECHEVERRIA, R. G.(2000) "Opciones para reducir la pobreza rural en América Latina y el Caribe." *Revista de la CEPAL* 70: 147-160.

FAO (Food and Agriculture Organization of the United Nations). (2008a) "Bioenergy, food security and Sustainability – Towards an International Framework." URL: http://www.fao.org/fileadmin/user_upload/foodclimate/HLCdocs/HLC08-inf-3-E.pdf. Acesso em 26/10/2009. Rome = FAO.

FAO (Food and Agriculture Organization of the United Nations). (2008b) "Climate Change, Bioenergy and Food Security: Civil Society and Private Sector Perspectives." URL: http://www.fao.org/fileadmin/user_upload/foodclimate/HLCdocs/HLC08-inf-6-E.pdf. Acesso em: 26/10/2009. Rome: FAO.

FAO (Food and Agriculture Organization of the United Nations). (2008c) "Climate Change, Bioenergy and Food Security: Options for Decision Makers identified by Expert Meetings." URL: http://www.fao.org/fileadmin/user_upload/food-climate/HLCdocs/HLC08-inf-5-E.pdf. Acesso em: 26 October 2009. Rome: FAO.

FAO (Food and Agriculture Organization of the United Nations). (2008d) "Bioenergy Policy, Markets and Trade and Food Security." URL: ftp://ftp.fao.org/docrep/fao/meeting/013/ai788e.pdf. Acesso em: 26 October 2009. Rome: FAO.

GARCEZ, C. A. G. & VIANNA, J. N. S. (2009) *Brazilian Biodiesel Policy: Social and environmental considerations*. Energy, doi: 10.1016/j.energy.2008.11.005.

GREENE, W. H. (2008) *Econometric analysis*. Sixth edition.

HILL, R. C. & GRIFFITHS, W. E. & LIM, G. C. (2008) P*rinciples of Econometrics*. Third edition.

IBAMA (Brazilian Institute of Environment and Ecossystems). URL: http://www.ibama.gov.br. Acesso em: 5 January 2008.

LELLI, S. (2001) "Factor Analysis vs Fuzzy Sets Theory: Assessing the Influence of Different Techniques on Sen's Functioning Approach." Discussion Paper. Faculty of Economics and Applied Economics, Katholieke Universiteit Leuven. URL: http://www.econ.kuleuven,ac.be/ew/academic/econover/Papers/DPS0121.pdf. Acesso em: 26 October 2009.

MARTINETTI, E. C. (2000) "A Multidimensional Assessment of Well-Being based on Sen's Functioning Approach." *Rivista Internazionale di Scienze Social*i 108: 207-239.

MICELI, D. (1998) "Measuring poverty using fuzzy sets." Discussion paper no. 38. National Centre for Social and Economic Modeling. University of Camberra.

NASS, L. & PEREIRA, P. & ELLIS, D. (2007) "Biofuels in Brazil: An Overview." *Crop Science* 47.

PIERSON, J. (2002) *Tackling Social Exclusion (Social Work Skills Series)*.

PNPB. (2005) "Programa Nacional de Produção e Uso de Biodiesel." URL: www.biodiesel.gov.br/programa.html. Acesso em: 12 November 2009.

PNUD (Programa das Nações Unidas para o Desenvolvimento). (2009) URL: http://www.pnud.org.br/gerapdf.php?id01=2802. Acesso em: 26 October 2009.

RAGIN, C. C. (2000) *Fuzzy-Set Social Science*. Chicago: The University of Chicago Press.

REARDON, T. & VOSTI, S. (1995) "Links between rural poverty and the environment in developing countries: asset categories and investment poverty". *World Development* 23(9): 1495-1506.

REYS, M. A. (2003) "Farming and Rural Systems Analyses in Forest Margins: An Application of Fuzzy Theory. The case of West Tocantins, Brazil." In DOPPLER, W. & BAUER, S. (eds.) *Farming and Rural Systems Economics*. Vol. 48. Wei-

kersheim: Margraf Verlag.

SILVER, H. & MILLER, M. (2002) "Social Exclusion: The European Approach to Social Disadvantage**."** Issue of Poverty and Race. URL: http://www.ag.ohio-state.edu/~hcrd/people/staff/Social%20Exclusion%20The%20European%20 approach.pdf . Acesso em: 15 march 2010.

UN-Energy. 2007) "Sustainable Bioenergy: A Framework for Decision Makers." URL: http://www.fao.org/docrep/010/a1094e/a1094e00.htm. Acesso em: 26 October 2009.

ZADEH, L. A. (1965) "Fuzzy sets." *Information and control*, 8: 338-353.

ZADEH, L. A. (1972) "A fuzzy set theoretic interpretation of linguistic hedges". *Journal of Cybernetics* 2:4-34.