

Multilevel Modelling of Income and Market Participation

Termpaper submitted to the Dina Research School for the PhD course on Likelihood-based inference for hierarchical/mixed statistical models, Greve, Denmark

by

Bharat P Bhatta and Soleiman Mohammadi Limaie

August 2005

1-Introduction

Isolation contributes to rural poverty. Rural life stagnates and local development prospects remain limited without a minimum of reliable and efficient access to locations of basic social and economic activities (Schelling and Lebo, 2001). Rural infrastructure is therefore a major priority for reducing rural poverty (World Bank, 1994).

Various studies provide evidence that socioeconomic development is closely associated with road transportation. Poverty is more pervasive in areas with no or unreliable road access. For example, in Nepal, where the percentage of people below poverty line is as high as 42 percent, the incidence of poverty in areas not connected by roads is 70 percent (Schelling and Lebo, 2001). In Bangladesh, the enrollment of girls in primary schools is three times as high in connected villages compared to unconnected areas (ibid). In Andhra Pradesh, India, the female literacy rate is 13 percent higher in villages with all weather road access compared to those with unreliable road access (Liu, 2000). Plenty of further evidence of the economic impact of rural roads exists (Pankaj, 1999). We therefore attempted to examine the impact of some variables like road accessibility , number of oxen, amount of own land and etc. on household income and market participation in Tigray, the northern highlands of Ethiopia. The aim of the project was to develop a hierarichal model and to analyze the fixed and random effects about the detrimiation of income and market participation.

2- Data and Estimation Methods

Data and variables

The data used in this project were collected from household survey undertakend in Tigray in 2003. 372 households were selected adopting a stratified random sampling approach based on market access, population density, rainfall, and irrigation projects from 16 communities located in four different zones of Tigray. The data set that we used in the project contains different socioeconomic varialbes such as access to information and public goods, household income and asset status, household characteristics, and market participation. Specifically, the varialbes used in the project and thier definition are as follows:

income	=	Household income
road	=	Road accessibility measured by the walking time to get to the nearest all weather road from the household
hsex	=	Sex of household head
hedu	=	Education of household head
Age	=	Age of household head
Oxen	=	Number of oxen in a household
Ownland	=	Land owned by each household
Cmp	=	Credit market participation by the hh
Cwr	=	Consumer worker ratio defined as the household size divided by the number of workers in a hh
Mlab	=	Male labor endowments in hh
Flab	=	Female labor endowments in hh
Tlab	=	Total labor endowments in hh
Tlu	=	Tropical livestock units in a hh

Here household income, denoted by income, and credit market participation by a household, denoted by cmp, are our dependent variables used in the project. We have identified different factors which determine or are correlated with income and cmp based on theory and empirical studies.

Hierarchical structure, and fixed and random effects

Principally there are five hierarchies in our data, viz from bottom to top household, village, district, zone, and state. But we do not have such clear and distinct hierarchies in our data so we developed our model in two hierarchies, viz household level and village level. There are 372 household level observations and 4 village level observations.

Data analysis and software

The MLwiN software was used to analyse the data. The model constructed on based of hierarchical data structure like: village level 16 and household level 372.

Model formulation

The mixed model can be written as

$$y = X\beta + Z\gamma + \varepsilon$$

Where y denotes the vector of observed y 's, X is the known matrix of x_{ij} 's, β is the unknown fixed-

effects parameter vector, Z is the known design matrix and γ is the vector of unknown random-effects parameter. We have formulated two models for our project. They are as follows.

Model I: Income

The factors determining the rural farm households' income may be broadly divided into four constructs: market and access to information, household resource endowments, household characteristics, and village level variables and public goods.

Model II: Market Participation

The factors influencing or correlated with the rural farm households' market participation may be broadly divided into four constructs: market and access to information, household income and resource endowments, household characteristics, and village level variables and public goods. Here market participation by households is a binary variable taking only two values: either participate or not i.e. 1 for market participation and 0 for not participation. So the distribution is binomial and the regression is logistic. We used the logit model to examine the market participation by the households. Based on theory and empirical evidence market participation may depend on road accessibility, income, amount of own land, number of oxen, age, sex and education of the household head.

4- Results and Discussion

Model I: Income

Model interpretation :

We have two levels of hierarchy in the model: household at the first level and village at the second level. The model we have now set up is a member of the variance component family. Because there is only one set up random effects (intercepts) for each level in the model.

As we can see from the results of the regression output (Fig 4.1) road accessibility, oxen and land holding significantly contributed more income of the households in the study area. But there is no significant relationship between household income and hsex, and hedu. The results also show that there is high variation of income between households than between villages.

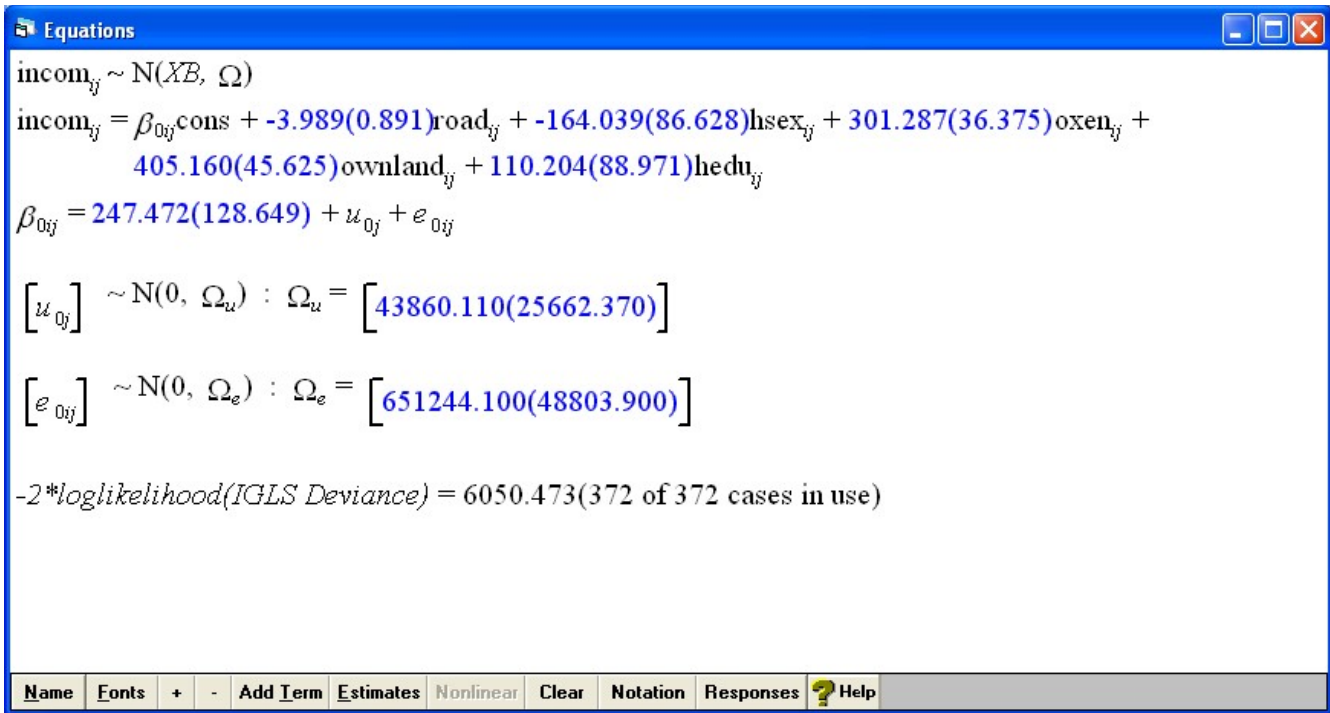


Fig.4.1. Model of household income

Residuals of the data also were calculated in household and village levels (Fig.4.2).

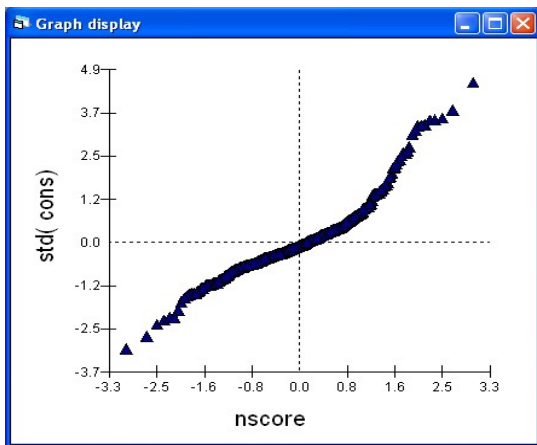


Fig.4.2. Residuals of the data on household level

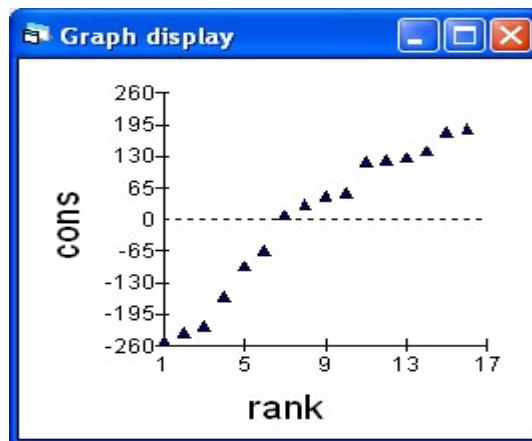


Fig.4.2. Residuals of the data on village level

Setting up a random intercept model in the equation:

After adding random effects on the independent variables we found that there are significant variation

for oxen and owned land in the households levels but there are not significant variation between the villages levels. There were not also any significant variations on hedu, hsex and road.

Predictions in the model:

First we added a quadratic own land in the model, this variable also is strongly significant with coefficient 96.39 and standard error 23.216.

The predicted equation is like:

$$\text{Income} = \beta_0 \text{cons} + \beta_6 \text{ownland} + \beta_8 (\text{ownland})^2.$$

The predicted equation and confidence interval were plotted to see how much uncertainty we had in our model. A linear prediction was also added to the model to compare the linear and quadratic own land equations (Fig.4.3).

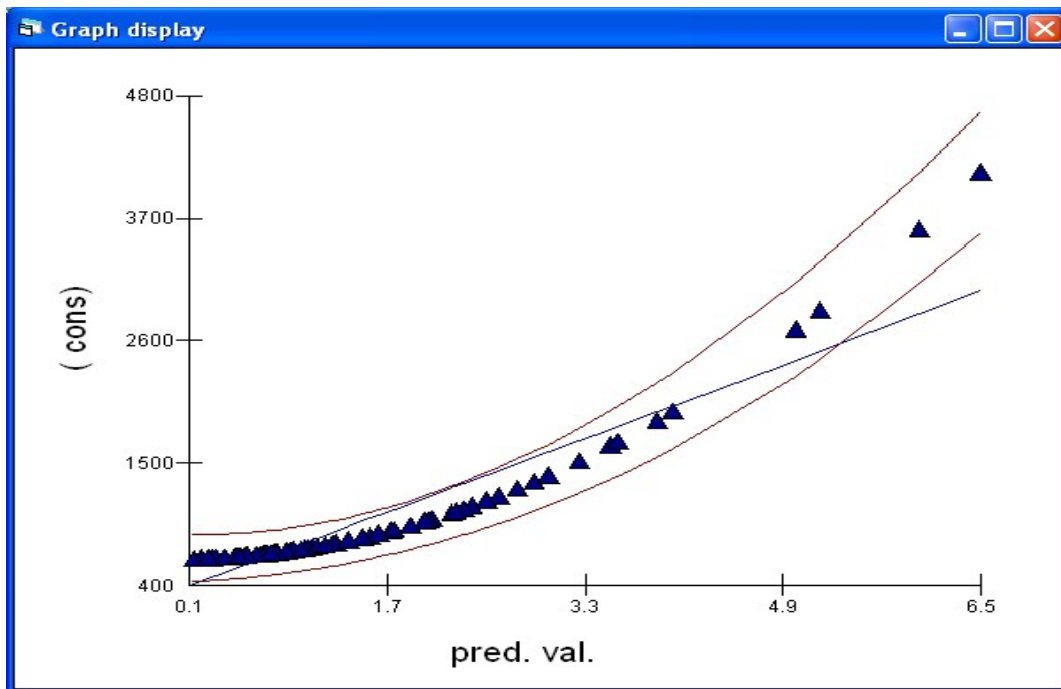


Fig.4.3: Quadratic, linear and confidence interval for ownland variable.

Model II: Market Participation

First we formulated the market participation model taking variables such as sex of household head, education of household head, number of oxen, amount of own land, road accessibility as explanatory

variables but we found that age and sex of the household head were only significant but other variables were not significant (Fig. 4.4).

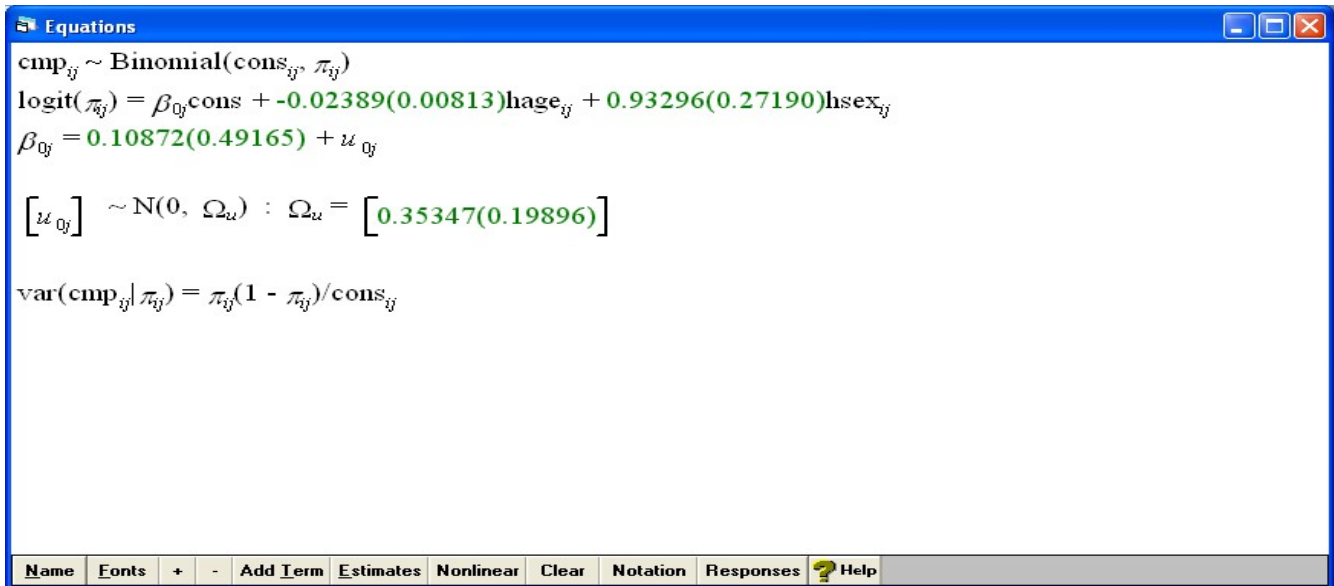


Fig. 4.4. Market participation model

From the model we see that the household level variability is more than that of the village level variability. The random effect for age of household head is significant but the random effect of sex of household head is not significant. We may conclude from the income and market participation models described above that the household variability is more than the village level variability.

Residuals of the data also were calculated in village levels (Fig.4.5)

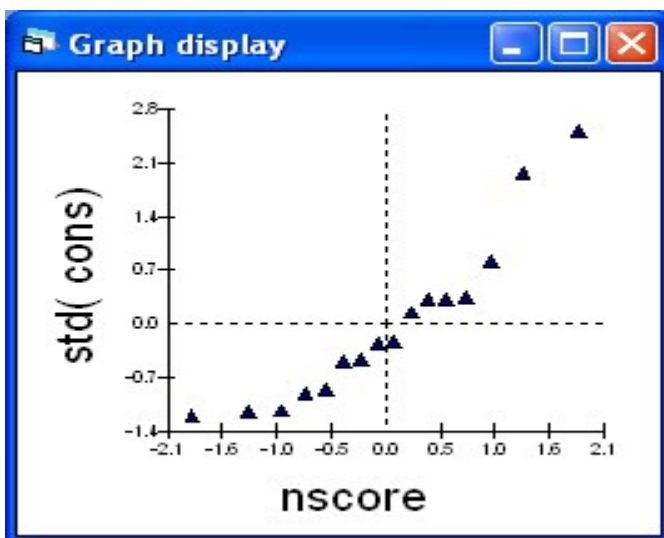


Fig.4.5. Residuals of the data on village level at market participation model.

5- Conclusion

In both income and market participation models, the household variability is more than the village level variability in determination of household income and market participation. As we hypothesized, road accessibility was positively significant to determine the household income. But contrary to our hypothesis Road accessibility was not significant factor in determining the market participation by the households in the study area. Age and sex of hh head were significant factors influencing hhs' market participation.