UNIVERSITA' DEGLI STUDI DI TRENTO - DIPARTIMENTO DI ECONOMIA

MODELS FOR NON-EXCLUSIVE MULTINOMIAL CHOICE, WITH APPLICATION TO INDONESIAN RURAL HOUSEHOLDS

Christopher L.. Gilbert
Francesca Modena

Discussion Paper No. 24, 2007

Dott. Stefano Comino
Dipartimento di Economia
Università degli Studi
Via Inama 5
38100 TRENTO ITALIA

Università degli Studi di Trento

Models for Non-Exclusive Multinomial Choice, with Application to Indonesian Rural Households

Christopher L. Gilbert and Francesca Modena
GRADE, University of Trento

20 September 2007

Abstract

Textbook discussions of discrete choice modelling focus on binomial and multinomial choice models in which agents select a single response. We consider the situation of non-exclusive multinomial choice. The widely used Marginal Logit Model imposes independence and has other disadvantages. We propose two models which account for non-exclusive and dependent multiple responses and require at least one response. In the first and simpler specification, the Poisson-multinomial, households first choose the number of responses to a specific shock, and then the specific choices are identified to maximize household utility conditional on the former choice. The second specification, the threshold-multinomial, generalizes the standard multinomial logit model by supposing that agents will choose more than one response if the utility they derive from other choices is "close" to that of the utility-maximizing choice. We apply these two approaches to reported responses of rural Indonesian rural households to demographic and economic shocks.

Keywords: discrete choice models, marginal logit, shocks, risk coping strategies

JEL Classification: C25, C51, O12

1. Introduction

Textbook discussions of discrete choice modelling focus on binomial and multinomial choice models in which agents select a single response. We consider the situation of non-exclusive multinomial choice. One possibility is to adopt the so-called Marginal Logit Model (MLM) which posits an independent binomial model for each choice (Agresti and Liu, 1999). The MLM has two disadvantages: it allows the possibility of no response which may not be realistic, and it supposes that response decisions are independent. We propose two alternative models which allow for interdependence and require at least one response.

We apply these models to the responses of Indonesian rural households to demographic and economic shocks. The structure of the interviews from which we take our data requires a shock to have a response. While the majority of shocks elicit only a single response, some shock instances elicit multiple responses. It appears that multiple responses are to a large extent associated with particular interviewers employed in the survey.

In a multinomial context, the standard random utility model supposes that the agent will pick the choice which maximizes his utility. Depending on the stochastic specification, one obtains either the multinomial logit or the multinomial probit model. When the number of potential responses exceeds three, multinomial probit becomes computationally infeasible and, absent any natural ordering of choices, one is obliged to use multinomial logit despite the well-known irrelevant alternatives problem. This is the situation from which we start.

We develop two models. In the first and simpler specification, choices are modelled as sequential: a household first chooses the number of responses to a specific shock, and then the specific choices are identified to maximize household utility conditional on the former choice. In this case, we may think of the interviewer as selecting the number of responses and the interviewee identifying the particular responses. In the second specification, we generalize the standard multinomial logit model by supposing that agents will choose more than one response if the utility they derive from other choices is “close” to that of the utility-maximizing choice. In effect, this supposes selection of a utility maximizing band, which will contain at least one choice but may contain more than one. This specification makes choice of the number of responses joint with choice of the particular responses.

2. Non-exclusive multinomial choice

The standard multiple choice model is posed in terms of maximization of a random utility function. Our application is to adjustment to shocks. Any such adjustment imposes costs. We adapt the random utility framework by modelling choice as resulting from minimization of a random cost function. We suppose that households may experience one of a number $S \geq 1$ of shocks and respond to each shock experienced from a choice set comprising $M \geq 2$ of adjustment modes. (In our data, all responses are available for all shock types). This gives a total of MS shock-response pairs.

Write c_{hms} for the cost of adjustment mode m to shock s for household h. These costs will depend on a vector x_h of household characteristics. There are H households in the sample. Following the random utility approach, we assume that adjustment costs have a deterministic and a stochastic component and write

$$c_{hms} = f_{hms} + \varepsilon_{hms} \tag{1}$$

The household chooses its adjustment mode(s) to minimize adjustment costs. Satisfaction of the budget constraint forces at least one response. Standard microeconomic theory suggests that it will be optimal for the household to make multiple responses, such that marginal adjustment cost is equalized across modes. Either discreteness (for example, in taking an extra job) or fixed costs may result in zero adjustment in one or more modes.

The Marginal Logit Model

We start with the case of a single response and then move on to multiple responses. Write $r_{hms} = 1$ if household h chooses response $m \in \{1, 2, ..., M\}$ in response to a shock of type s. Define $p_{hms} = \Pr\left(r_{hms} = 1 \mid x_h\right)$ as the probability that response m is the cost-minimizing response to shock s. For simplicity, focus on the first response p_{h1s}. Henceforth, we omit the shock subscript s where this does not result in ambiguity. Ignoring the possibility of ties

$$p_{h1} = \Pr\left(r_{h1} = 1\right) = \Pr\left(c_{h1} < c_{hm}, m = 2,.., M \mid x_h\right) \tag{2}$$

and similarly for the remaining M-1 choices. Following Domenich and McFadden (1975) assume that the stochastic cost components ε_{hm} follows an extreme value (Gnedenko) distribution. Then $c_{hm} = f_{hm} + \varepsilon_{hm}$ also has the same distribution as does the cost of the minimizing choice $c_h^* = \min_{m \in \{1,..,M\}} c_{hm}$. Hence the probabilities p_{hm} are logistic:

$$p_{hm} = \frac{e^{-f_{hm}}}{\sum_{j=1}^{M} e^{-f_{hj}}} = \frac{a_{hm}}{\sum_{j=1}^{M} a_{hj}} \quad (m = 1, 2, .., M) \tag{3}$$

where $a_{hm} = e^{-f_{hm}} \quad (m = 1, 2, .., M)$. (The minus signs reflect the fact that we are minimizing costs rather than maximizing utilities as in the standard random utility model).

In many circumstances it will be possible for agents to have multiple responses. The most simple way in which to model responses as independent.

$$p_{hm} = \frac{e^{-f_{hm}}}{1 + e^{-f_{hm}}} = \frac{a_{hm}}{1 + a_{hm}} \tag{4}$$

This is the Marginal Logit Model of Agresti and Liu (1999) – see also Loughin and Scherer (1998) and Agresti (2003). Treating choices as independent implies that the probability of choosing any one alternative does not affect the probability of choosing the others.

In principle, this independence across alternatives allows us to estimate each probability relationship separately. However, this approach has two disadvantages: it allows the possibility of a null response (not possible in our dataset) and it does not take into account possible interdependence across choices.[1] In contexts, such as that in which we find ourselves, in which most agents choose a single approach, the independence assumption appears implausible. In other contexts in which multiple choice is the norm and null responses are possible, the MLM may be completely satisfactory.

In contexts in which choices appear dependent, it may be more appropriate to develop a version of the multinomial model modified to allow multiple response. We develop two models that account for non-exclusive and dependent multiple responses: a Poisson model and a threshold model that generalizes the random utility approach.

The Poisson model supposes that the household makes a sequential decision, first choosing the number $m_h = \#(\Omega_h)$ of responses to a shock, and then identifying the best (i.e. cost-minimizing) m_h response, i.e. the set Ω_h conditional on this choice m_h. By contrast, the threshold model supposes that the number m_h of responses is an outcome of the response identification decision.

[1] Agresti and Liu (1999, p.943) note the absence of null responses in the Kansas livestock farmer dataset analyzed by Loughin and Scherer (1998).

Dependent Multiple Responses: A Poisson-Multinomial Model[2]

The survey design obliges households to identify at least one response to any shock. Hence the number m_h of responses is a non-zero integer: $m_h \in \{1, 2, .., M\}$. If $m_h - 1$ follows a Poisson process with mean $\mu_h = \mu(x_h)$, where the unit displacement reflects the impossibility of a null response, we may write

$$\Pr(m_h \mid x_h) = \frac{e^{-\mu(x_h)}\mu(x_h)^{m_h-1}}{(m_h-1)!} \tag{5}$$

Consider first the case in which $m_h = 1$ and response j is selected. In the logit framework

$$\Pr(\Omega_h = \{j\} \mid m_h = 1, x_h) = p_{hj}^1 = \frac{e^{-f_{hj}}}{\sum_{m=1}^{M} e^{-f_{hm}}} = \frac{a_{hj}}{\sum_{m=1}^{M} a_{hm}} \tag{6}$$

Combining equations (5) and (6)

$$\Pr(\Omega_h = \{j\} \mid x_h) = p_{hj} = \frac{\theta_h a_{hj}}{\sum_{m=1}^{M} a_{hm}} \tag{7}$$

where $\theta_h = e^{-\mu_h}$.

Turning to the case in which $m_h = 2$ with responses i and j selected, we need to consider the probability that j is the overall cost minimizing choice and that i is the next best, and the converse situation in which i is the overall cost minimizing choice and that j is the next best. Following the derivation shown in Appendix B, the probability of choosing responses i and j given that $m_h = 2$ is given by:

$$\Pr(\Omega_h = \{i, j\} \mid x_h) = p_{hij} = \frac{a_{hi}a_{hj}\theta_h\mu_h}{\sum_{m=1}^{M} a_{hm}}.\left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)}\right) \tag{8}$$

The argument is similar in the case that three responses are selected - see Appendix B.

[2] There exists a literature on so-called multinomial-Poisson models in which individuals make multiple responses across a range of response modes. An example is transport mode frequencies for different transport modes in which households may use different modes on different occasions – see Terza and Wilson (1990). These models replace the multinomial response probabilities with Poisson frequencies. Our models differ with the polar opposite case in which responses remain categorical but the number of responses is variable and is modeled as Poisson.

Dependent Multiple Responses: A Threshold Model

As previously, let Ω_h be the set of responses made by household *h* to a particular shock. We generalize the random utility framework by introducing a household-specific threshold $t_h = t\left(x_h\right) \geq 0$. Within this framework, the household may choose the first response to shock *s* either because this is the cost-minimizing choice or because one of the other choices is cost-minimizing but the cost of the first choice is sufficiently close.

First consider households which make a single response. Define

$$p_{h1} = \Pr\left(\Omega_h = \{1\}\right)$$

Using the notation already established

$$\begin{aligned} p_{h1} &= \Pr\left(c_{h1} \leq c_{hm} - t_h, m = 2,...,M\right) \\ &= \Pr\left(f_{h1} + \varepsilon_{h1} \leq f_{hm} + \varepsilon_{hm} - t_h, m = 2,...,M\right) \\ &= \Pr\left(\varepsilon_{h1} - \varepsilon_{hm} \leq f_{hm} - f_{h1} - t_h, m = 2,...,M\right) \end{aligned} \tag{9}$$

If the errors have the extreme value distribution, and generalizing to the case in which response *j* is chosen:

$$p_{hj} = \frac{e^{-f_{hj}}}{e^{-f_{hj}} + \sum_{m=1, m \neq j}^{M} e^{-\left(f_{hm} - t_h\right)}} = \frac{a_{hj}}{a_{hj} + \lambda_h \sum_{m=1, m \neq j}^{M} a_{hm}} \tag{10}$$

where $\lambda_h = e^{t_h} \geq 1$.

Now consider a household which responds using two modes, say 1 and 2:

$$p_{h12} = \Pr\left(\Omega_h = \{1, 2\}\right)$$

We need to consider two cases, that in which choice mode 1 is cost minimizing while mode 2 is sufficiently close to be also chosen, and the converse case in which 2 is cost minimizing and 1 is also chosen. Using the same notation

$$\begin{aligned} p_{h12} &= \Pr\left(c_{h1} < c_{h2} \leq c_{h1} + t_h \;\&\; c_{h1} \leq c_{hm} - t_h, m = 3,...,M\right) \\ &+ \Pr\left(c_{h2} \leq c_{h1} \leq c_{h2} + t_h \;\&\; c_{h2} \leq c_{hm} - t_h, m = 3,...,M\right) \end{aligned} \tag{11}$$

In the general case for two choices in which $\Omega_h = \{j, k\}$, this probability is given by

$$p_{hjk} = \frac{\left(\lambda_h - 1\right) a_{hj} a_{hk}}{\left(a_{hj} + a_{hk} + \lambda_h \sum_{m \notin \Omega_h}^{M} a_{hm}\right)} \cdot \left[\frac{1}{\left(a_{hj} + \lambda_h \sum_{m=1, m \neq j}^{M} a_{hm}\right)} + \frac{1}{\left(a_{hk} + \lambda_h \sum_{m=1, m \neq k}^{M} a_{hm}\right)}\right] \tag{12}$$

See Appendix B for derivation of this result and also for the case in which three responses are selected.

3. Shocks, household responses and their consequences

The analysis of the uncertainty affecting households, and their responses to this uncertainty, is a key issue in developing countries, where poor people are exposed to risks that affect household living conditions (Morduch, 1994; Dercon, 2005). Shocks, defined as large adverse movements in their incomes or consumption requirements, can have a major impact on the possibility of the household escaping poverty or may induce a non-poor household to enter poverty. Uncertainty is therefore central to our understanding of vulnerability. World Bank (2000) notes the importance of policies that help poor people to manage the risks they face. A growing theoretical and empirical literature focuses on the analysis of income variability and on the ability of households to overcome income risks. A related literature looks at vulnerability.

Poor people have developed mechanisms to deal with hardships. Often these involve informal insurance arrangements between individuals and entire communities. Although these strategies offer some cushion against shocks , they are not always sufficient with the consequence that shocks may push households into poverty or exacerbate their existing poverty status. Even when households are able to deal with risk, the risk-management strategies they adopt may have negative consequences. They may, for example, destroy or reduce the physical, financial, human or social capital of the household (Dercon, 2005), in this way increasing the risk of entering poverty when faced with future . Transient shocks can give rise to permanent effects when children are required to drop out of school or, are required to work while remaining in school, (de Janvry et al., 2006). In this sense, short run income maintenance may be at the expense of longer-term well-being. Furthermore, fear of risk can force poor households to choose safe but less profitable choices (Morduch, 1990; Alderman and Paxson, 1992; Rosenzweig and Wolpin (1993); with the result that risk-coping strategies may divert resources away from directly productive activities and may prevent households from exploiting comparative advantage. Hence, if we are to design an appropriate income protection framework, it is important to understand how households cope with actual shocks and the possibility of future shocks, and to evaluate which responses are costlier for households.

4. Data

The data used for this study are from the 1993 Indonesia Family Life Survey data (IFLS1). 7224 households were interviewed over a wide range of issues. Our focus is on the section of the survey relating to demographic and economic shocks. Respondents were asked whether their household had experienced an economic shock in the past five years, the type of the shock, when it happened (year and month), what measures were taken and the costs of overcoming the shocks.

Six types of shock are analyzed in the IFLS dataset:

i) death of a household member,
ii) sickness of a household member
iii) crop loss,
iv) household or business loss due to a disaster,
v) unemployment of a household member,
vi) fall in the price of a crop.

We distinguish between demographic and economic shocks – demographic shocks are death and sickness, while economic shocks are the remaining four categories. The nature of the shock is important because it has implications for the ability to cope with its consequences (see Dercon, 2002), and influences the response adopted. A related distinction is between idiosyncratic and common shocks which is correlated with but not implied by the demographic-economic distinction.

Turning to the measures adopted to cope with the shocks, the survey allowed us to distinguish six possible responses:

i) extra job,
ii) loan, (including a loan from families or friends),
iii) asset sale (sale of next harvest, food, cattle or poultry, jewellery or other assets),
iv) family assistance,
v) use of savings,
vi) cut expenditures.

The survey questionnaire was not explicit as to whether single or multiple responses to a shock were sought.[3]

[3] The survey also includes an explicit question on the costs associated with each shock. We do not use the answers to this question in this paper for two reasons. First, response is partial. Second, it is unclear whether this variable measures the pre- or post-response (i.e. gross or net) costs associated with shocks. (The variable we have defined in equation (1) is on a gross basis).

The responses identified in the survey are all ex-post risk-coping strategies.[4] They can be divided into two categories: risk-sharing strategies that smooth consumption across households, and intertemporally smoothing strategies, that smooth consumption over time. Risk sharing responses involve either formal institutions, such as formal credit transactions, or informal mechanisms (e.g. transfers between families or friends). Instead, households smooth consumption intertemporally by saving and borrowing, or by accumulating and selling non-financial assets (Alderman and Paxson, 1992; Bardhan and Udry, 1999; Dercon, 2002).

We include only those households in our dataset that supplied a complete set of income and demographic data. After dropping income outliers (1.3% of the total sample), and considering only rural households, the sample reduces to 3246 households. 1116 households (34.4% of the total sample) experienced at least one shock in the five years reporting period, 697 of them (21,5% of the total sample) experienced at least one shock in 1992-93. Table 1 reports the number of households that experienced each type of shock over the five years 1989-93 and the two years 1992-93.

Table 1
Reported Household Shock Experience

Type of shock	1989-93		1992-93	
	# rural households	per cent	# rural households	per cent
Death	254	7.8%	111	3.5%
Sickness	325	10.0%	169	5.2%
Crop loss	544	16.8%	340	10.5%
Business loss	60	1.8%	35	1.0%
Unemployment	54	1.7%	28	0.9%
Price falls	231	7.0%	147	4.5%

The table reports the number of households, and the percentage of all households sampled, reporting shocks of each type over the five year period 1989-93 and the two year sub-period 1992-93 used in the subsequent analysis.

The most frequent shocks are sickness and crop loss. Business loss and unemployment affect only a few households. In view of the low incidence of these shocks in our data, we aggregate these into a single category reducing the number M of shock types to five for the purposes of econometric analysis.

Table 2 shows the percentage of multiple responses for each shock. The majority of households report a single response. This is consistent with the view either that responses are interdependent - the fact of having chosen (or reported) one response mode reduces the probability of choosing (or

[4] Ex ante risk-management strategies include diversification across crops, the use of a variety of production techniques, etc.

reporting) others – or that many interviewers interpreted the survey question as requiring a single response.

Table 2 **Percentage of Multiple Responses Reported**		
	1989-93	1992-93
Death	19.7%	22.5%
Sickness	19.7%	16.6%
Crop loss	18.6%	17.4%
Business loss due to a disaster	15.0%	8.6%
Unemployment	24.0%	21.5%
Price falls	15.0%	17.7%

The table reports the percentage of those households which experienced each type of shock who reported multiple responses over the five year period 1989-93 and the two year sub-period 1992-93 used in the subsequent analysis.

Table 3 shows the percentage of households that responded in each manner for each shock. These statistics suggest that household responses differ between demographic shocks (death, illness) and economic shocks (crop loss, business loss or unemployment, price falls). The data suggest an important role for family and community assistance in the case of demographic shocks, while this measure appears relatively less important as a response to crop loss and price falls.

.

Table 3 **Shock Responses by Household (1992-93)**					
	Death	Sickness	Crop loss	Business loss or unemployment	Price falls
Extra job	13.5%	7.7%	47.9%	34.4%	42.2%
Take loan	28.8%	36.0%	18.8%	29.5%	17.0%
Sell assets	27.0%	27.2%	15.6%	23.0%	17.7%
Family assistance	36.0%	21.9%	7.0%	13.0%	4.0%
Use savings	15.3%	17.7%	5.0%	3.3%	4.0%
Cut down on household expenses	4.5%	4.7%	22.6%	14.7%	32.6%

The table summarizes the percentages of those households which experienced each type of shock who identify each response mode the two year sub-period 1992-93 used in the subsequent analysis. Because of multiple responses, percentages sum to more than 100%.

Cameron and Worswick (2003) have argued that labour supply responses help Indonesian households to smooth consumption in the face of a crop loss. This response appears particularly important for economic shocks. It is also apparent that economic shocks are more likely to lead to a decline in consumption than are demographic shocks.

We turn now to the explanatory variables we will use in the econometric model. Microeconomic theory indicates that two variables are potentially important in explaining shock responses. The first is the extent to which shocks are common across households. Standard discussions indicate that informal insurance mechanisms are better able to cope with idiosyncratic shocks than with common shocks. The dataset allows us to identify the village in which a household is resident. We define a commonality variable as the weighted percentage of households (other than the household in question) that experienced the same shock in the same village. Let Z_{hs} be the percentage of families that experience shock *s* in the village in which household *h* is resident. The commonality variable, modifies to exclude the household in question, is defined as $z_{hs} = Z_{hs}\frac{n_h}{n_h - 1} - \frac{1}{n_h - 1}\delta_{hs}$, where n_h is the number of households surveyed in the village and δ_{hs} is a dummy variable equals to one if the household *h* has experienced shock. The modification is important because the unmodified variable Z_{hs} will not be independent of δ_{hs} in villages in which there is a small number of reporting households.

The second potentially important variable is household permanent income. Certain shock responses are more easily available to rich households than to poor households. We should therefore expect that the probability of choosing a specific mode will be affected by the household's wealth. For example, poor people are less able to save and accumulate assets, and they have restricted access to credit because of lack of collaterals. As suggested in World Bank (2000), poor health and bad nutrition limit the possibility to work more or send more household members to work. The poor households are thus more vulnerable and have limited means to deal with a crisis. We measure household wealth through estimated permanent income. Construction of the permanent income variable is discussed in Appendix A.

In addition to these economic variables, the survey design (in particular, incomplete instructions) may result that the identity of the interviewer plays a role in determining the number of responses chosen by the household. Since we are able to identify the interviewer for each respondent, we relate the number of responses chosen by each household to the average number of responses elicited by the same interviewer, excluding responses given by the household in question.

5. The empirical model

We define the cost associated with a certain measure and a certain shock as a random cost with a stochastic and a deterministic component – see section 2. We assume that households choose the response mode that minimizes this adjustment cost or, in the case of multiple responses, that has a cost sufficiently close to the best choice. To implement the models we need to specify the deterministic cost component (f_{hms} in equation (1)), the household specific threshold (t_{hs}), and the Poisson parameter μ_{hs}. Hence the empirical strategy has two components: the first is the definition and estimation of the deterministic cost, and the second involves the adjustment of the multinomial model to account for non-exclusivity of responses.

The three variables we use vary across households and shocks but not across response modes. This implies that the response probabilities to a particular shock all depend on the same three variables. Alternative models imply different nonlinear mappings of these variables into the unit interval. In principles, each such mapping might depend on a large number of parameters, but with a limited number of observations it is difficult to identify all parameters. (In particular, a number of shock-response pairs are poorly represented in the data). We impose structure on these mappings by jointly estimating the response probabilities across shocks and by imposing a degree of response homogeneity.

We adopt a linear specification for the deterministic cost component:

$$f_{hms} = \kappa_{ms} + \gamma_m z_{hs} + \alpha_{ms} y_h^P \tag{13}$$

z_{hs} is the variable defined in section 4 that captures the commonality of the shock in each village, and y_h^P is the estimated household permanent income. We impose homogeneity on the intercepts for the demographic and economic shocks respectively

$$\begin{aligned} \kappa_{ms} &= \kappa_m^d \quad (s = 1, 2) \\ \kappa_{ms} &= \kappa_m^e \quad (s = 3, 4, 5) \end{aligned} \tag{14}$$

In the Poisson model we posit

$$\mu_{hs} = \exp(\mu_s + \beta_s v_{hs}) \tag{15}$$

where v_{hs} is the average number of responses elicited by household *h*'s interviewer to shock type *s*. Similarly, for the threshold model we suppose that the threshold, beneath which response costs are regarded as indistinguishable, is influenced by the identity of the interviewer, motivating the specification

$$t_{hs} = \exp(\tau_s + \varphi_s v_{hs}), \tag{16}$$

where $s = \{d, e\}$ (demographic and economic)

The estimated cost functions are latent and therefore have an arbitrary zero. This implies that we need to normalize the parameters. We do this by setting the adjustment costs c_{h1s} to zero for each shock type *s*. The implied parameter restrictions are

$$\kappa_{1s} = \gamma_1 = \alpha_{1s} = 0 \quad \left(s = 1,..,S\right) \tag{17}$$

To evaluate the importance of modelling multiple responses as interdependent, we compare the results obtained from these models with those from the MLM which treats each response as an independent decision. Both the threshold and the Poisson-multinomial models use the identical adjustment costs expression (13) but substitutes the MLM probability (4) for the threshold and the Poisson probabilities. Note that parameter normalization is not required in the MLM case – the alternative to responding in a particular manner is not making that response. The implication is that it is only the intra-response mode differences that are comparable across the two models. To obtain comparability between the MLM, the threshold and Poisson probabilities, we re-normalize equation (13) for the MLM model, as

$$f_{hms} = (k_{1s} + \kappa_{ms}) + (g_1 + \gamma_m) z_{hs} + (a_{1s} + \alpha_{ms}) y_h^P + \beta_s v_{hs} \tag{18}$$

in conjunction with the restrictions given by equations (17).

6. Results

We have estimated the parameters of the deterministic component of the cost function (13) using the MLM, the Poisson-multinomial models and the threshold multinomial. Results are reported in Table 4 (see end of paper).

- Looking first at the MLM specification, the greater the shock commonality, the lower is the cost of the labour supply response implying a higher probability of adopting this response. The effect of commonality is positive for the remaining responses, implying that they are less likely to be adopted, but the coefficient is only significant for asset sales. The signs of commonality coefficients are the same in the Poisson and threshold models as in the MLM case, but statistical significance is greater.
- The estimated models provide very clear evidence that the probability of responding to shocks through use of savings increases with permanent income (i.e. the cost of responding through the use of savings is significantly negatively related to permanent income)[5].

[5] We ran the Poisson and threshold models with the full set of α coefficients. All apart from α_5 ("use savings") were close to zero. Even in the MLM model we cannot reject the hypothesis that all the coefficients on permanent income except α_5 are equal to zero ($\chi^2(5) = 3.81$, tail probability 0.58).

- The average number of responses for the interviewer is highly significant for economic shocks, and fairly significant for demographic shocks irrespective of model specification confirming that interviewer identity plays an important role in determining the number of responses.

These results are all in line with the qualitative conclusions drawn from Tables 1-3 in section 4.

7. Testing the model specification

We have considered three models – the first (MLM) supposes independence of alternatives, the second (Poisson-multinomial) and the third (threshold multinomial) treat choices as dependent but not mutually exclusive. These models have allowed us to estimate the probabilities of choosing different alternative using both the MLM and the multinomial specification. The MLM and the multinomial specifications answer two different questions and maximize different likelihood functions.

The MLM model treats choices over different responses as independent. The likelihood is defined in terms of the probability of each response being selected. If household *h* selects response *m* to shock *s*, $r_{\text{hms}} = 1$. This outcome occurs with probability p_{hms}. Similarly, the outcome and $r_{\text{hms.}}= 0$ occurs with probability ($1-p_{\text{hms}}$). The overall probability can be written in the binomial form $p_{hms}^{r_{hms}}(1-p_{hms})^{1-r_{hms}}$ and the log-likelihood is $r_{hms}\ln p_{hms} + (1-r_{hms})\ln(1-p_{hms})$. The overall log-likelihood function is

$$LB = \sum_{h=1}^{H}\sum_{s=1}^{S}\sum_{m=1}^{M}\left[r_{hms}\ln p_{hms} + (1-r_{hms})\ln(1-p_{hms})\right] \tag{19}$$

In the multinomial specification the likelihood function is maximized over the entire set of all the possible combinations of responses. The entire set possible combinations of up to three choices is given by Q = 41 possibilities and we index these by q such that $\Omega_{hs}^{1} = \{1\}$, …, $\Omega_{hs}^{6} = \{6\}$, $\Omega_{hs}^{7} = \{1,2\}$ …, $\Omega_{hs}^{11} = \{1,6\}$ etc. Define $\tilde{r}_{hqs} = 1$ if the combination Ω^{q} is chosen, 0 otherwise, with q = 1,…,Q and let $\tilde{p}_{hqs}$ and $1-\tilde{p}_{hqs}$ be the associated probabilities. The log-likelihood for the Poisson and the threshold models is defined as:

$$LJ = \sum_{h=1}^{H}\sum_{s=1}^{S}\sum_{q=1}^{Q}\left[\tilde{r}_{hqs}\ln\tilde{p}_{hqs} + (1-\tilde{r}_{hqs})\ln(1-\tilde{p}_{hqs})\right] \tag{20}$$

The MLM and multinomial likelihood functions are not directly comparable although either can be transformed into the other. Given the independent choice probabilities p_{hms} estimated from the MLM model, the corresponding multinomial probabilities $\tilde{p}_{hs}^{q}$ may be computed as

$$\tilde{p}_{hqs} = \prod_{m \in \Omega^q} p_{hms} \prod_{m \notin \Omega^q} \left(1 - p_{hms}\right) \tag{21}$$

Equivalently, given the multinomial probabilities $\tilde{p}_{hs}^{q}$ we may compute the associated probabilities p_{hms} associated with each choice as

$$p_{hms} = \sum_{q=1}^{Q} 1\left(m \in \Omega^q\right) \tilde{p}_{hqs} \tag{22}$$

where the function $1(v)$ returns the value unity if v is true and zero if false.

We use expressions (21) and (22) to calculate the multinomial (joint choice-based) likelihood based on the estimated MLM probabilities and the MLM (independent choice-based) probabilities for the two multinomial models. Table 5 lists the maximized log-likelihoods on both bases for all three specifications. The two multinomial models have higher log-likelihoods irrespective of the choice basis. On the independent choice basis, the threshold-multinomial model slightly out-performs the Poisson-multinomial model, but the ranking is reversed on the joint choice basis.

Table 5
Log-likelihoods

	Independent choice basis *LB*	**Joint choice basis *LJ***
Marginal logit model	-2125.75	-2518.34
Poisson-multinomial	-2124.66	-2454.87
Threshold-multinomial	-2123.93	-2456.67

The table records the results maximized log-likelihoods from equations (19) for the three models considered .

The three models we have considered are not nested and comparison of likelihoods is therefore at best a criterion for good fit and not a test. In what follows, we first use a version of the paired *J* test introduced by Davidson and Mackinnon (1981).[6]

Index the three specifications as (*b*,*p*,*t*) for the MLM, Poisson-multinomial and threshold-multinomial models respectively. Write the estimated individual choice-based probabilities as

[6] A "paired" non-nested test is a test between a pair of two hypotheses from a larger set of hypotheses (McAleer, 1995). Our test should be thought of as a *J*-type test rather than a pure *J* test since our models do not fall within the linear regression class.

p^{j}_{hms} $\left(j = b, p, t \right)$ and the estimated joint choice-based probabilities as $\tilde{p}^{j}_{hqs}$ $\left(j = b, p, t \right)$. Construct the two sets of differences

$$d^{j}_{hms} = p^{j}_{hms} - p^{b}_{hms} \quad \left(j = p, t \right) \tag{23}$$

To perform the *J*-type test we include these differences d^{j}_{hms} additively in the augmented MLM model. The *J* test statistic for the MLM null against alternative *j* is the one-sided *t* statistic on the coefficient λ_j of the variable d^{j}_{hms}.

The procedure for testing the two multinomial models is identical. We construct the four set of differences[7]

$$\tilde{d}^{jk}_{hqs} = \tilde{p}^{j}_{hqs} - \tilde{p}^{k}_{hqs} \quad \left((j,k) = (p,b), (t,b), (p,t), (t,p) \right) \tag{24}$$

Regarding model *j* as the null, we re-estimate the model including the difference $\tilde{d}^{jk}_{hqs}$ as an additive regressor. The *J* test statistic for null *j* against alternative *k* is the *t* statistic on the coefficient λ_{jk}. Monte Carlo evidence has established that *J* tests have a pronounced tendency to over-reject in finite samples – see McAleer and Pesaran (1986) and McAleer (1987).

Table 6
***J* Test Results**

		Alternative hypothesis		
		Marginal logit model	Poisson -multinomial	Threshold -multinomial
Null hypothesis	Marginal logit model	-	1.80 [3.60%]	1.41 [7.93%]
	Poisson-multinomial	1.20 [11.5%]	-	2.94 [0.16%]
	Threshold-multinomial	2.23 [1.29%]	2.58 [0.50%]	-

The table records the results of the *J* tests for each pair of models. The test statistics are distributed as Student *t*. Tail probabilities are given in parentheses. The hypotheses tests are all one-sided so rejection of the null at the conventional 95% level is appropriate if the tail probability is inferior to 5%. The tests are calculated using equations (23) and (24).

Test outcomes are listed in Table 6. At the 5% level, neither of the multinomial models rejects the MLM model whereas the two multinomial models reject each other with the threshold model also being rejected by the MLM model. These outcomes are not easy to reconcile with the likelihood values reported in Table 5. We note that the properties of the *J*-style test have not been established

[7] In effect three since $\tilde{d}^{tp}_{hqs} = -\tilde{d}^{pt}_{hqs}$.

for nonlinear environments and it also seems possible that our sample, although large, is insufficient to give reliable results.

We obtain clearer results from an alternative approach using a linear probability (LPM) framework. To test the MLM null, consider the six regressions

$$r_{hms} = \delta_j p_{hms}^j + \delta_k p_{hms}^k + u_{hms}^{jk} \quad \left(j,k = b,p,t; k \neq j \right) \tag{25}$$

If the MLM model (H_j say) is valid, we should find $\delta_j = 1$ and $\delta_k = 0$. Conversely, if hypothesis k (H_k) is valid we should find $\delta_j = 0$ and $\delta_k = 1$. Similarly, in the multinomial framework, we consider the six regressions

$$\tilde{r}_{hqs}^j = \delta_j \tilde{p}_{hqs}^j + \delta_k \tilde{p}_{hqs}^k + \tilde{u}_{hqs}^{jk} \quad \left(j,k = b,p,t; k \neq j \right) \tag{26}$$

The tests have the same form. However, as is well-known, the LPM suffers from heteroscedasticity and so in all cases we use a heteroscedasticity-robust estimate of the variance-covariance matrix.

Table 7
Linear Probability Model Test Results

Individual choice basis		δ_b	δ_p	δ_t	H_k versus H_j	H_j versus H_k
H_j	Marginal logit model	0.398	0.603	-	1.21	0.53
H_k	Poisson-multinomial	(1.03)	(1.55)		[29.9%]	[59.0%]
H_j	Marginal logit model	0.315	-	0.692	1.77	0.42
H_k	Threshold-multinomial	(0.86)		(1.88)	[17.0%]	[65.9%]
H_j	Poisson-multinomial	-	0.334	0.672	0.83	0.24
H_k	Threshold-multinomial		(0.64)	(1.29)	[43.7%]	[78.3%]
Joint choice basis						
H_j	Poisson-multinomial	0.011	1.005	-	32.7	0.10
H_k	Marginal logit model	(0.09)	(8.09)		[0.00%]	[90.8%]
H_j	Threshold-multinomial	0.049	-	0.975	36.8	0.26
H_k	Marginal logit model	(0.42)		(8.58)	[0.00%]	[77.0%]
H_j	Poisson-multinomial	-	0.287	0.733	2.18	0.43
H_k	Threshold-multinomial		(0.78)	(2.01)	[11.3%]	[65.0%]

The table reports the results of the test based using the linear probability (LPM) framework, described in equations (25) and (26). Hypothesis test statistics (column 4 and 5) are heteroscedasticity-corrected F tests. Heteroscedasticity-robust t-statistics are in "(.)" parentheses and tail probabilities in "[.]" parentheses.

Results are reported in Table 7. The upper block of tests relates to the individual choice basis. The tests fail to discriminate between the alternative models even though the estimated coefficients give a greater weight to the Poisson and threshold probabilities than to those form the MLM model. By contrast, using the joint choice basis (Table 7, lower block) the two multinomial models decisively reject the MLM model. Although it remains true that neither multinomial model is able to reject the other, the estimated coefficients give a higher weight to the threshold model in line with the log-likelihoods reported in Table 5.

In summary, the test outcomes depend on the way the model is framed. If the question is, "Which response modes will be adopted?", this motivates an individual choice approach. In this case, the standard MLM model appears adequate. If, instead, the question is, "How will households respond?", a joint choice is required. In this second context, it is important to explicitly acknowledge the joint nature of multiple response choices and the MLM model is clearly inadequate. The evidence is less decisive in relation to the choice between alternative multinomial specifications although there is some suggestion that the threshold-multinomial model is slightly superior to the Poisson-multinomial model.

8. Conclusions

Multinomial choice models have traditionally focussed on exclusive choice. Survey design may however permit multiple responses. One possibility is to model such responses as MLM, but this ignores possible interdependence of responses and allows the possibility of null response. We have developed two models which generalize the McFadden's now standard random utility framework to allow for the possibility of multiple response. In the first of these models, the respondent first decides on the number of responses and then chooses the actual responses to maximize utility conditional on that prior choice. In the second, threshold, model, the two decisions are made jointly, with the agent choosing multiple responses if utility outcomes are sufficiently close. These models are both relatively straightforward from a computational standpoint provided the number of responses selected remains small.

We apply this framework to modelling the responses of households in rural Indonesia to demographic and economic shocks. The survey design obliges respondents to nominate at least one response to any such shock. A minority of households nominate multiple responses. The incidence of multiple responses appears to be primarily a function of the identity of the interviewer, and it appears that interviewers may have interpreted the survey instructions differently. Both the Poisson and threshold multinomial models outperform the MLM model. Choice between the two multinomial models is less clear but the data appear to marginally more favourable to the threshold model.

There are also substantive conclusions. Macroeconomic theory emphasizes the role of the individual household's savings as a device for smoothing consumption ion the face of income shocks. Our data for rural Indonesian households demonstrates the importance of this mechanism but only for the richest households. By contrast, the theoretical literature on shock response in development economics has emphasized the role of informal insurance arrangements at the family and village level but has noted that these arrangements only work well when shocks are idiosyncratic. We develop a measure of the commonality of shocks and show that response choice does indeed depend on commonality. This provides strong confirmation of the importance of these informal arrangements which appear to provide the dominant coping mechanism for poorer households.

References

Agresti, A. (2003), *Categorical Data Analysis* (2nd edition), New York, Wiley.

Agresti, A., and I-M. Liu (1999), "Modeling a categorical variable allowing arbitrarily many category choices", *Biometrics,* **55**, 936-43.

Alderman H, and C. Paxson (1992), "Do the Poor Insurance? A Synthesis of the Literature on Risk and Consumption in Developing Countries", *World Bank, Policy Research Working Paper Series* #1008

Bardhan, P.K. and C. Udry (1999), *Development Microeconomics*, Clarendon Press, Oxford.

Cameron L, and C. Worswick (2003), "The labour market as a smoothing device: labour supply responses to crop loss", *Review of Development Economics*, **7,** 327-341.

Davidson, R., and J.B. MacKinnon (1981), "Several tests for model specification in the presence of alternative hypotheses", *Econometrica*, **49**, 551-65.

de Janvry A., F. Finan, and E. Sadoulet (2006), "Can conditional cash transfers serve as safety nets to keep children at school and out of the labor market?", *Journal of Development Economics*, **79**, 349-73

Dercon, S. (2002), "Income risk, coping strategies and safety nets", *WIDER Discussion Paper* #2002/22.

Dercon, S. (2005), "Risk, vulnerability and poverty in Africa", *Journal of African Economies*, **14,** 483-88.

Dercon, S. and J. Hoddinott (2003), "Health, shocks and poverty persistence", WIDER Discussion Paper # 2003/08.

Domencich, T. and D. McFadden (1975), *Urban Travel Demand: A Behavioral Analysis*, Amsterdam, North-Holland, Jacoby, H. and E. Skoufias (1997), “Risk, Financial Markets, and Human Capital in a Developing Country”, *Review of Economic Studies*, **64**, 311–36.

Liu, I., and A. Agresti (2005), “The analysis of ordered categorical data: An overview and survey of recent developments”, *Test*, **14**, 1-73.

Loughin, T.M., and P.N. Scherer (1998), “Testing for association in contingency tables with multiple column responses”, *Biometrics*, **54**, 630-37.

McAleer, M. (1987), “Specification tests for separate models: a survey”, in M.L. King and D.E.A. Giles eds., *Specification Analysis in the Linear Model*, London, Routledge.

McAleer, M. (1995), “The significance of testing empirical non-nested models”, *Journal of Econometrics*, **67**, 149-71.

McAleer, M., and M.H. Pesaran (1986), “Statistical inference in non-nested econometric models”, *Applied Mathematics and Computation*, **20**, 271-311.

Morduch, J (1990), “Risk, production and saving: theory and evidence from Indian households”, Harvard University. processed.

Morduch, J. (1994), “Poverty and vulnerability”, *American Economic Review,* **84**, 221-25.

Paxon, C. (1992), “Using weather variability to estimate the response of savings to transitory income in Thailand”, *American Economic Review*, **82**, 5-33.

Rosenzweig, M.R. and K.I. Wolpin (1993), “Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: investments in bullocks in India”, *Journal of Political Economy*, **101**, 223-24

Terza, J.V., and P.W. Wilson (1990), “Analyzing frequencies of several types of events: a mixed multinomial-Poisson approach”, *Review of Economics and Statistics*, **72**, 108-15.

World Bank (2000), *World Development Report 2000/01: Attacking Poverty*, Washington DC, World Bank.

Table 4 Estimated Models						
	Marginal logit model		**Poisson-multinomial model**		**Threshold-multinomial model**	
	Coeff.	*t*	*Coeff.*	*t*	*Coeff.*	*t*
κ *(demographic)*						
extra job	2.81	6.91				
loan	0.93	2.51	-1.54	-6.89	-1.49	-6.75
sell assets	1.13	2.99	-1.40	-6.05	-1.34	-5.89
family assistance	1.17	3.04	-1.36	-5.56	-1.33	-5.46
use savings	2.94	6.27	0.24	0.71	0.15	0.41
cut expenses	3.41	7.53	0.64	1.83	0.62	1.82
κ *(economic)*						
extra job	1.54	5.46				
loan	2.05	6.88	0.40	2.13	0.34	1.92
sell assets	2.05	6.73	0.38	1.93	0.34	1.83
family assistance	3.22	8.91	1.40	5.10	1.28	4.84
use savings	4.71	10.26	2.82	7.45	2.63	7.26
cut expenses	1.94	7.09	0.31	1.78	0.24	1.44
γ *(commonality)*						
extra job	-0.29	-4.18				
loan	0.08	1.17	0.27	3.32	0.26	3.40
sell assets	0.20	2.22	0.37	4.04	0.36	4.06
family assistance	0.12	0.97	0.31	2.53	0.31	2.63
use savings	0.11	0.72	0.31	2.00	0.30	2.01
cut expenses	0.01	0.15	0.20	2.72	0.19	2.79
α *(permanent income)*						
use savings	-0.55	-5.03	-0.52	-5.52	-0.47	-5.43
β *interviewer*						
demographic	-0.30	-1.05				
economic	-0.69	-3.46				
μ, τ *(intercept)*						
demographic			-2.84	-4.74	-3.00	-3.04
economic			-5.12	-10.05	-6.29	-9.27
μ,τ *(interviewer)*						
demographic			1.09	2.31	1.56	1.90
economic			2.73	7.29	4.01	7.84
log likelihood	-2125.7444		-2454.87		-2456.67	

The table reports the estimated parameters of the deterministic component of the cost function (13) using respectively the MLM, the Poisson-multinomial and the threshold multinomial models. In the Poisson and threshold models, parameters are normalized setting the adjustment cost c_{h1s} to zero for each shock type s ($\kappa_{1s} = \gamma_1 = \alpha_{1s} = 0 \quad (s = 1,..,S)$). The estimated threshold t_{hs} and Poisson parameters μ_{hs} are also reported.

Appendix A – Income Equation Estimation

We adapt the methodology used by Paxson (1992) and Cameron and Worswick (2003) to decompose household income[8] into permanent and transitory components. We estimate the following equation:

$$Y_h = \alpha_0 + \alpha_1 X_h^P + \alpha_2 X_h^T + \nu_h \quad \text{(A1)}$$

where Y_h is the income of household h and X_h^P and X_h^T are variables viewed as determinants of permanent and the transitory income respectively. This allows us to decompose income as:[9]

$$\hat{Y}_h^P = \hat{\alpha}_0 + \hat{\alpha}_1 X_h^P + \hat{\alpha}_2 \overline{X}^T \quad \text{(A2)}$$

$$\hat{Y}_h^T = \hat{\alpha}_2 \left(X_h^T - \overline{X}^T \right) + \hat{\nu}_h \quad \text{(A3)}$$

We use the following variables X_h^P to identify the permanent component: the number of household members in each age categories, the number of adult members (age 18-64) in each education/gender category, dummies variables that indicate the occupation of the household head,[10] a dummy that identifies if there is a householder who has a non-farm business, the value of land and provincial dummies. X_h^2 includes dummy variables for the shocks experienced in the previous two years. There are two complications. First, not all shocks can be treated as transitory. For example, death of a household member may affect income in a permanent way. Hence, deaths occurred in the previous five years are included in the estimation of the permanent component. Second, households with a non-farm business are more likely to experience a household or business loss due to a disaster. This motivates the inclusion of an interaction term between the non-farm business dummy and the business loss variable in the vector.

[8] Household income (Y_h) is calculated as the sum of the following variables: wages earned by each household member, net profit generated by the farm, net profit generated by the household business, household income other than from business or employment (pension, scholarship loan, insurance claim, winnings, gift from family or friends, other), total income from household assets (other that farm and business assets).

[9] We subtract the sample mean of the X_h^T variables to ensure that transitory income has zero sample mean.

[10] Self employed workers, employees or family workers.

TableA1 Descriptive Statistics				
Variable	**Mean**	**s.d.**	**Min**	**Max**
Household income	1080	1598	- 104	20130
Death	0.078	0.27	0	1
Sickness	0.05	0.22	0	1
Crop loss	0.105	0.306	0	1
Business loss	0.011	0.103	0	1
Business loss x non farm business	0.006	0.078	0	1
Unemployment	0.009	0.092	0	1
Price falls	0.045	0.21	0	1
Non-farm business	0.32	0.466	0	1
Land value	3062	11077	0	200000
# household members aged 0 to 5	0.65	0.81	0	5
# household members aged 6 to 11	0.71	0.84	0	4
# household members aged 12 to 17	0.64	0.83	0	5
# household members aged 18 to 64	2.32	1.06	0	12
# household members aged over 64	0.2	0.5	0	8
# males 18-64 without education	0.17	0.4	0	3
# females 18-64 without education	0.37	0.54	0	3
# males 18-64 – primary school only	0.63	0.65	0	5
# females 18-64 – primary school only	0.64	0.62	0	5
# males 18-64 up to secondary school	0.29	0.55	0	4
# females 18-64 up to secondary school	0.19	0.44	0	4
# males 18-64 high school	0.025	0.17	0	3
# females 18-64 high school	0.012	0.11	0	2
Head employee	0.29	0.45	0	1
Head self-employed	0.69	0.46	0	1
Head family worker	0.016	0.12	0	1

Sample size: 3246

“Death” refers to a death in the previous five years. All other negative shock variables refer to the two years 1992-93. Household income and land value are in thousand rupiahs.

Table A2
Income decomposition equation estimates
Dependent variable: Household income

Permanent components	**Coefficient**	***t***
Intercept	242.1	1.21
Death	-267.32	-3.41
Non-farm business	587	10.11
Land value	0.021	4.17
# household members aged 6 to 11	83	2.31
# household members aged 12 to 17	97.73	2.95
# household members aged over 64	-37.40	-0.57
# males 18-64 up to secondary school	591.78	9
# females 18-64 up to secondary school	570.48	5.72
# males 18-64 high school	1737.66	6.23
# females 18-64 high school	1913.4	4.63
Head employee	736.84	4.5
Transitory components	**Coefficient**	***t***
Sickness	-52.54	-0.51
Crop loss	-88.2	-1.21
Business loss	340.17	0.98
Business loss x non farm business	-351.9	-0.76
Unemployment	-424.54	-2.32
Price falls	-57.82	-0.51
$F_{31,3320}$	25.61	
R^2	0.3172	

The table reports the OLS estimates from equation (A1). The household income is regressed on a set of variables that determine the permanent and the transitory income components. "Non-farm business" is a dummy that is equal to one if the household owns a non farm business. The dummy "Head employee" refers to the work status of the household head. "Death" refers to a death in the previous five years. All other negative shock variables refer to the two years 1992-93. Only the principal variables that enter in the permanent income estimation are reported in the table.
Sample: 3246 households. Robust *t* statistics in parentheses.

.

Appendix B – Calculation of multinomial probabilities

Poisson-multinomial model

In the case in which $m_h = 2$ with responses *i* and *j* selected, we need to consider the probability that *j* is the overall cost minimizing choice and that *i* is the next best, and the converse situation in which *i* is the overall cost minimizing choice and that *j* is the next best. Hence, using the notation already established:

$$\Pr\left(\Omega_h = \{i, j\} \mid m_h = 2, x_h\right) = p_{hij}^2 = \frac{a_{hj}}{\sum_{m=1}^{M} a_{hm}} \cdot \frac{a_{hi}}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{a_{hi}}{\sum_{m=1}^{M} a_{hm}} \cdot \frac{a_{hj}}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)}$$

$$= \frac{a_{hi} a_{hj}}{\sum_{m=1}^{M} a_{hm}} \cdot \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)} \right) \tag{B1}$$

Combining equations (5) and (B1),

$$\Pr\left(\Omega_h = \{i, j\} \mid x_h\right) = p_{hij} = \frac{a_{hi} a_{hj} \theta_h \mu_h}{\sum_{m=1}^{M} a_{hm}} \cdot \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)} \right) \tag{B2}$$

The argument is similar in the case that three responses are selected. We obtain

$$\Pr\left(\Omega_h = \{i, j, \ell\} \mid x_h\right) = p_{hij\ell} = \frac{a_{hi} a_{hj} a_{h\ell} \theta_h \mu_h^2}{2\sum_{m=1}^{M} a_{hm}} \cdot$$

$$\left[\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj}\right)} \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj} - a_{h\ell}\right)} \right) \right.$$

$$+ \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi}\right)} \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{hj}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{h\ell}\right)} \right)$$

$$\left. + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{h\ell}\right)} \left(\frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hi} - a_{h\ell}\right)} + \frac{1}{\left(\sum_{m=1}^{M} a_{hm} - a_{hj} - a_{h\ell}\right)} \right) \right] \tag{B3}$$

Threshold Multinomial Model

In the case of double responses, equation (11) gives the probability of choosing modes 1 and 2:

$$
\begin{aligned}
p_{h12} = &\Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \;\&\; c_{h1} \le c_{hm} - t_h, m = 3,\ldots,M\right) \\
&+ \Pr\left(c_{h2} \le c_{h1} \le c_{h2} + t_h \;\&\; c_{h2} \le c_{hm} - t_h, m = 3,\ldots,M\right)
\end{aligned}
$$

Consider the first term in this expression. We may split this into two further components as

$$
\begin{aligned}
&\Pr(c_{h1} < c_{h2} \le c_{h1} + t \;\&\; c_{h1} \le c_{hm} - t_h, m = 3,\ldots,M) \\
&= \Pr(c_{h1} < c_{hm} - (1-\delta_{m2})t_h, m = 2,\ldots,M) - \Pr(c_{h1} < c_{hm} - t_h, m = 2,\ldots,M)
\end{aligned}
$$

where δ_{ij} is the Kronecker delta, $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases}$

This is illustrated in terms of the distribution function of c_2 for the case of M =3 in Figure 1.

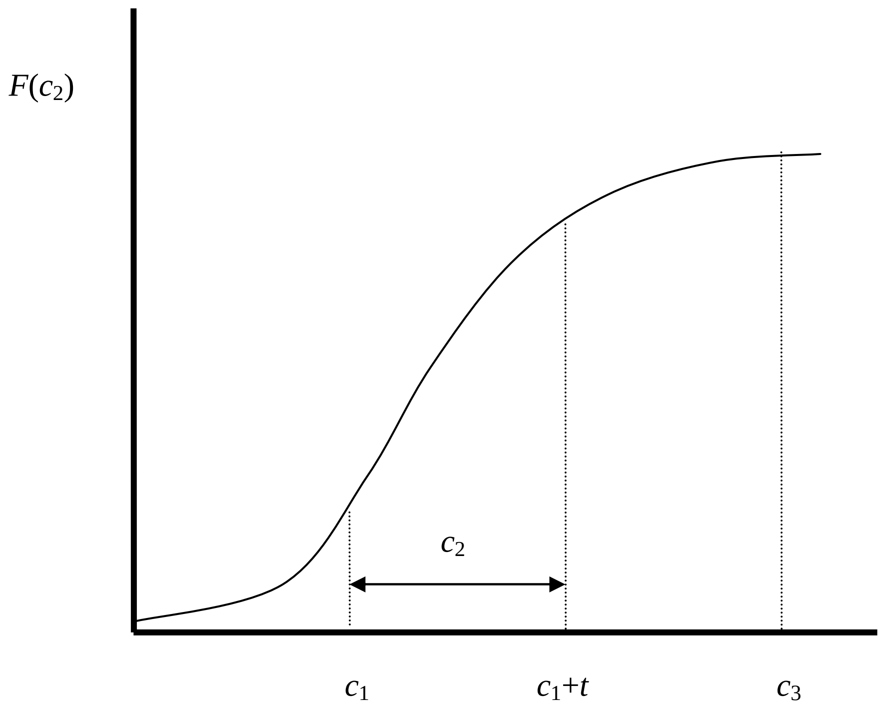

Figure 1: Distribution function of c2

Using equation (10), this probability becomes

$$
\Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \;\&\; c_{h1} \le c_{hm} - t_h, m = 3,\ldots,M\right) = \frac{a_{h1}}{a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}} - \frac{a_{h1}}{a_{h1} + \lambda_h \sum_{m=2}^{M} a_{hm}}
$$

$$
= \frac{(\lambda_h - 1)\, a_{h1} a_{h2}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h \sum_{m=2}^{M} a_{hm}\right)} \qquad \text{(B4)}
$$

The second component of equation (11) follows directly as

$$\Pr(c_{h2} < c_{h1} \le c_{h2} + t_h \,\&\, c_{h2} \le c_{hm} - t_h, m = 3,\ldots,M) = \frac{a_{h2}}{a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}} - \frac{a_{h2}}{a_{h2} + \lambda_h \sum_{m=1,m\neq 2}^{M} a_{hm}}$$

$$= \frac{(\lambda_h - 1) a_{h1} a_{h2}}{(a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm})(a_{h2} + \lambda_h \sum_{m=1,m\neq 2}^{M} a_{hm})} \quad \text{(B5)}$$

Combining equations (B4) and (B5), we obtain

$$p_{h12} = \frac{(\lambda_h - 1) a_{h1} a_{h2}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=3}^{M} a_{hm}\right)} \cdot \left[\frac{1}{\left(a_{h1} + \lambda_h \sum_{m=2}^{M} a_{hm}\right)} + \frac{1}{\left(a_{h2} + \lambda_h \sum_{m=1,m\neq 2}^{M} a_{hm}\right)} \right] \quad \text{(B6)}$$

Generalization to the case in which three modes are chosen is yet more complicated. For notational simplicity, let $\Omega_h = \{1,2,3\}$. Then

$$\begin{aligned} p_{h123} &= \Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 4,\ldots,M\right) \\ &+ \Pr\left(c_{h2} \le c_{h1}, c_{h3} \le c_{h2} + t_h \,\&\, c_{h2} \le c_{hm} - t_h, m = 4,\ldots,M\right) \\ &+ \Pr\left(c_{h3} \le c_{h1}, c_{h2} \le c_{h3} + t_h \,\&\, c_{h3} \le c_{hm} - t_h, m = 4,\ldots,M\right) \end{aligned} \quad \text{(B7)}$$

As previously, we analyze the three components separately. Consider the first component:

$$\Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 4,\ldots,M\right)$$

This probability depends on the values of both c_2 and c_3 which are taken as following independent and identical Gumbel distributions.

Hence

$$\begin{aligned} &\Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 4,\ldots,M\right) \\ &\quad = \Pr\left(c_{h1} < c_{h2} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 4,\ldots,M\right).\Pr\left(c_{h1} < c_{h3} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 4,\ldots,M\right) \end{aligned}$$

Using equation (6), we may write this joint probability as

$$\begin{aligned} &\Pr\left(c_{h1} < c_{h2}, c_{h3} \le c_{h1} + t_h \,\&\, c_{h1} \le c_{hm} - t_h, m = 4,\ldots,M\right) \\ &= \frac{(\lambda_h - 1)^2 a_{h1}^2 a_{h2} a_{h3}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h \sum_{m=2,m\neq 3}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h \sum_{m=3}^{M} a_{hm}\right)} \end{aligned} \quad \text{(B8)}$$

It follows that

$p_{h123} = \left(\lambda_h - 1\right)^2 a_{h1} a_{h2} a_{h3}.$

$$\left[\frac{a_{h1}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h1} + \lambda_h a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)}\right.$$

$$+\frac{a_{h2}}{\left(a_{h1} + a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + \lambda_h a_{h1} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + \lambda_h a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)}$$

$$\left.+\frac{a_{h3}}{\left(a_{h1} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h2} + a_{h3} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h3} + \lambda_h a_{h1} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)\left(a_{h3} + \lambda_h a_{h2} + \lambda_h \sum_{m=4}^{M} a_{hm}\right)}\right]$$

Elenco dei papers del Dipartimento di Economia

2000.1 *A two-sector model of the effects of wage compression on unemployment and industry distribution of employment*, by Luigi Bonatti

2000.2 *From Kuwait to Kosovo: What have we learned? Reflections on globalization and peace,* by Roberto Tamborini

2000.3 *Metodo e valutazione in economia. Dall'apriorismo a Friedman ,* by Matteo Motterlini

2000.4 *Under tertiarisation and unemployment.* by Maurizio Pugno

2001.1 *Growth and Monetary Rules in a Model with Competitive Labor Markets,* by Luigi Bonatti.

2001.2 *Profit Versus Non-Profit Firms in the Service Sector: an Analysis of the Employment and Welfare Implications,* by Luigi Bonatti, Carlo Borzaga and Luigi Mittone.

2001.3 *Statistical Economic Approach to Mixed Stock-Flows Dynamic Models in Macroeconomics,* by Bernardo Maggi and Giuseppe Espa.

2001.4 *The monetary transmission mechanism in Italy: The credit channel and a missing ring,* by Riccardo Fiorentini and Roberto Tamborini.

2001.5 *Vat evasion: an experimental approach,* by Luigi Mittone

2001.6 *Decomposability and Modularity of Economic Interactions,* by Luigi Marengo, Corrado Pasquali and Marco Valente.

2001.7 *Unbalanced Growth and Women's Homework,* by Maurizio Pugno

2002.1 *The Underground Economy and the Underdevelopment Trap,* by Maria Rosaria Carillo and Maurizio Pugno.

2002.2 *Interregional Income Redistribution and Convergence in a Model with Perfect Capital Mobility and Unionized Labor Markets,* by Luigi Bonatti.

2002.3 *Firms' bankruptcy and turnover in a macroeconomy,* by Marco Bee, Giuseppe Espa and Roberto Tamborini.

2002.4 *One "monetary giant" with many "fiscal dwarfs": the efficiency of macroeconomic stabilization policies in the European Monetary Union,* by Roberto Tamborini.

2002.5 *The Boom that never was? Latin American Loans in London 1822-1825,* by Giorgio Fodor.

2002.6 *L'economia senza banditore di Axel Leijonhufvud: le 'forze oscure del tempo e dell'ignoranza' e la complessità del coordinamento,* by Elisabetta De Antoni.

2002.7 *Why is Trade between the European Union and the Transition Economies Vertical?,* by Hubert Gabrisch and Maria Luigia Segnana.

2003.1 *The service paradox and endogenous economic gorwth,* by Maurizio Pugno.

2003.2 *Mappe di probabilità di sito archeologico: un passo avanti,* di Giuseppe Espa, Roberto Benedetti, Anna De Meo e Salvatore Espa.
(*Probability maps of archaeological site location: one step beyond,* by Giuseppe Espa, Roberto Benedetti, Anna De Meo and Salvatore Espa).

2003.3 *The Long Swings in Economic Understianding,* by Axel Leijonhufvud.

2003.4 *Dinamica strutturale e occupazione nei servizi,* di Giulia Felice.

2003.5 *The Desirable Organizational Structure for Evolutionary Firms in Static Landscapes,* by Nicolás Garrido.

2003.6 *The Financial Markets and Wealth Effects on Consumption* An Experimental Analysis, by Matteo Ploner.

2003.7 *Essays on Computable Economics, Methodology and the Philosophy of Science,* by Kumaraswamy Velupillai.

2003.8 *Economics and the Complexity Vision: Chimerical Partners or Elysian Adventurers?*, by Kumaraswamy Velupillai.

2003.9 *Contratto d'area cooperativo contro il rischio sistemico di produzione in agricoltura,* di Luciano Pilati e Vasco Boatto.

2003.10 *Il contratto della docenza universitaria. Un problema multi-tasking*, di Roberto Tamborini.

2004.1 *Razionalità e motivazioni affettive: nuove idee dalla neurobiologia e psichiatria per la teoria economica?* di Maurizio Pugno.
(*Rationality and affective motivations: new ideas from neurobiology and psychiatry for economic theory?* by Maurizio Pugno.

2004.2 *The economic consequences of Mr. G. W. Bush's foreign policy. Can th US afford it?* by Roberto Tamborini

2004.3 *Fighting Poverty as a Worldwide Goal* by Rubens Ricupero

2004.4 *Commodity Prices and Debt Sustainability* by Christopher L. Gilbert and Alexandra Tabova

2004.5 *A Primer on the Tools and Concepts of Computable Economics* by K. Vela Velupillai

2004.6 *The Unreasonable Ineffectiveness of Mathematics in Economics* by Vela K. Velupillai

2004.7 *Hicksian Visions and Vignettes on (Non-Linear) Trade Cycle Theories* by Vela K. Velupillai

2004.8 *Trade, inequality and pro-poor growth: Two perspectives, one message?* By Gabriella Berloffa and Maria Luigia Segnana

2004.9 *Worker involvement in entrepreneurial nonprofit organizations. Toward a new assessment of workers? Perceived satisfaction and fairness* by Carlo Borzaga and Ermanno Tortia.

2004.10 *A Social Contract Account for CSR as Extended Model of Corporate Governance (Part I): Rational Bargaining and Justification* by Lorenzo Sacconi

2004.11 *A Social Contract Account for CSR as Extended Model of Corporate Governance (Part II): Compliance, Reputation and Reciprocity* by Lorenzo Sacconi

2004.12 *A Fuzzy Logic and Default Reasoning Model of Social Norm and Equilibrium Selection in Games under Unforeseen Contingencies* by Lorenzo Sacconi and Stefano Moretti

2004.13 *The Constitution of the Not-For-Profit Organisation: Reciprocal Conformity to Morality* by Gianluca Grimalda and Lorenzo Sacconi

2005.1 *The happiness paradox: a formal explanation from psycho-economics* by Maurizio Pugno

2005.2 *Euro Bonds: in Search of Financial Spillovers* by Stefano Schiavo

2005.3 *On Maximum Likelihood Estimation of Operational Loss Distributions* by Marco Bee

2005.4 *An enclave-led model growth: the structural problem of informality persistence in Latin America* by Mario Cimoli, Annalisa Primi and Maurizio Pugno

2005.5 *A tree-based approach to forming strata in multipurpose business surveys,* Roberto Benedetti, Giuseppe Espa and Giovanni Lafratta.

2005.6 *Price Discovery in the Aluminium Market* by Isabel Figuerola-Ferretti and Christopher L. Gilbert.

2005.7 *How is Futures Trading Affected by the Move to a Computerized Trading System? Lessons from the LIFFE FTSE 100 Contract* by Christopher L. Gilbert and Herbert A. Rijken.

2005.8 *Can We Link Concessional Debt Service to Commodity Prices?* By Christopher L. Gilbert and Alexandra Tabova

2005.9 *On the feasibility and desirability of GDP-indexed concessional lending* by Alexandra Tabova.

2005.10 *Un modello finanziario di breve periodo per il settore statale italiano: l'analisi relativa al contesto pre-unione monetaria* by Bernardo Maggi e Giuseppe Espa.

2005.11 *Why does money matter? A structural analysis of monetary policy, credit and aggregate supply effects in Italy*, Giuliana Passamani and Roberto Tamborini.

2005.12 *Conformity and Reciprocity in the "Exclusion Game": an Experimental Investigation* by Lorenzo Sacconi and Marco Faillo.

2005.13 *The Foundations of Computable General Equilibrium Theory*, by K. Vela Velupillai.

2005.14 *The Impossibility of an Effective Theory of Policy in a Complex Economy*, by K. Vela Velupillai.

2005.15 *Morishima's Nonlinear Model of the Cycle: Simplifications and Generalizations,* by K. Vela Velupillai.

2005.16 Using and Producing *Ideas* in Computable Endogenous Growth, by K. Vela Velupillai.

2005.17 *From Planning to Mature: on the Determinants of Open Source Take Off* by Stefano Comino, Fabio M. Manenti and Maria Laura Parisi.

2005.18 *Capabilities, the self, and well-being: a research in psycho-economics,* by Maurizio Pugno.

2005.19 Fiscal and monetary policy, unfortunate events, and the SGP arithmetics. Evidence from a *growth-gap* model, by Edoardo Gaffeo, Giuliana Passamani and Roberto Tamborini

2005.20 *Semiparametric Evidence on the Long-Run Effects of Inflation on Growth,* by Andrea Vaona and Stefano Schiavo.

2006.1 *On the role of public policies supporting Free/Open Source Software. An European perspective,* by Stefano Comino, Fabio M. Manenti and Alessandro Rossi.

2006.2 *Back to Wicksell? In search of the foundations of practical monetary policy,* by Roberto Tamborini

2006.3 *The uses of the past,* by Axel Leijonhufvud

2006.4 *Worker Satisfaction and Perceived Fairness: Result of a Survey in Public, and Non-profit Organizations,* by Ermanno Tortia

2006.5 *Value Chain Analysis and Market Power in Commodity Processing with Application to the Cocoa and Coffee Sectors,* by Christopher L. Gilbert

2006.6 *Macroeconomic Fluctuations and the Firms' Rate of Growth Distribution: Evidence from UK and US Quoted Companies,* by Emiliano Santoro

2006.7 *Heterogeneity and Learning in Inflation Expectation Formation: An Empirical Assessment,* by Damjan Pfajfar and Emiliano Santoro

2006.8 Good *Law & Economics* needs suitable microeconomic models: the case against the application of standard agency models: the case against the application of standard agency models to the professions, by Lorenzo Sacconi

2006.9 *Monetary policy through the "credit-cost channel". Italy and Germany,* by Giuliana Passamani and Roberto Tamborini

2007.1 *The Asymptotic Loss Distribution in a Fat-Tailed Factor Model of Portfolio Credit Risk,* by Marco Bee

2007.2 *Sraffa?s Mathematical Economics – A Constructive Interpretation,* by Kumaraswamy Velupillai

2007.3 *Variations on the Theme of Conning in Mathematical Economics,* by Kumaraswamy Velupillai

2007.4 *Norm Compliance: the Contribution of Behavioral Economics Models,* by Marco Faillo and Lorenzo Sacconi

2007.5 *A class of spatial econometric methods in the empirical analysis of clusters of firms in the space,* by Giuseppe Arbia, Giuseppe Espa e Danny Quah.

2007.6 *Rescuing the LM (and the money market) in a modern Macro course,* by Roberto Tamborini.

2007.7 *Family, Partnerships, and Network: Reflections on the Strategies of the Salvadori Firm of Trento,* by Cinzia Lorandini.

2007.8 *I Verleger serici trentino-tirolesi nei rapporti tra Nord e Sud: un approccio prosopografico,* by Cinzia Lorandini.

2007.9 *A Framework for Cut-off Sampling in Business Survey Design,* by Marco Bee, Roberto Benedetti e Giuseppe Espa

2007.10 *Spatial Models for Flood Risk Assessment,* by Marco Bee, Roberto Benedetti e Giuseppe Espa

2007.11 *Inequality across cohorts of households:evidence from Italy,* by Gabriella Berloffa and Paola Villa

2007.12 *Cultural Relativism and Ideological Policy Makers in a Dynamic Model with Endogenous Preferences,* by Luigi Bonatti

2007.13 *Optimal Public Policy and Endogenous Preferences: an Application to an Economy with For-Profit and Non-Profit,* by Luigi Bonatti

2007.14 *Breaking the Stability Pact: Was it Predictable?,* by Luigi Bonatti and Annalisa Cristini.

2007.15 *Home Production, Labor Taxation and Trade Account,* by Luigi Bonatti.

2007.16 *The Interaction Between the Central Bank and a Monopoly Union Revisited: Does Greater Uncertainty about Monetary Policy Reduce Average Inflation?,* by Luigi Bonatti.

2007.17 *Complementary Research Strategies, First-Mover Advantage and the Inefficiency of Patents,* by Luigi Bonatti.

2007.18 *DualLicensing in Open Source Markets,* by Stefano Comino and Fabio M. Manenti.

2007.19 *Evolution of Preferences and Cross-Country Differences in Time Devoted to Market Work,* by Luigi Bonatti.

2007.20 *Aggregation of Regional Economic Time Series with Different Spatial Correlation Structures,* by Giuseppe Arbia, Marco Bee and Giuseppe Espa.

2007.21 *The Sustainable Enterprise. The multi-fiduciary perspective to the EU Sustainability Strategy,* by Giuseppe Danese.

2007.22 *Taming the Incomputable, Reconstructing the Nonconstructive and Deciding the Undecidable in Mathematical Economics,* by K. Vela Velupillai.

2007.23 *A Computable Economist's Perspective on Computational Complexity,* by K. Vela Velupillai.

2007.24 *Models for Non-Exclusive Multinomial Choice, with Application to Indonesian Rural Households,* by Christopher L. Gilbert and Francesca Modena.

PUBBLICAZIONE REGISTRATA PRESSO IL TRIBUNALE DI TRENTO