Bayesian Inference for a Stratified Categorical Variable Allowing All Possible Category Choices

Balgobin Nandram[1], Maria Criselda S. Toto[1], Myron Katzoff[2]
Department of Mathematical Sciences, Worcester Polytechnic Institute[1]
Office of Research and Methodology, CDC National Center for Health Statistics[2] *

Abstract

We consider problems of inference from survey data for proportions when sample individuals have been asked to mark all responses that apply to them, the table with mutually exclusive categories is sparse, the number of individuals to whom none of the listed categories apply is missing and the category proportions are to be compared across population strata. We consider an example from the Kansas Farm Survey which is described in Loughin and Scherer (1998), where the reader may find the data presented in tabular form. We use a Bayesian product multinomial-Dirichlet model to fit the count data both within and across farmer's education levels. We estimate the proportions of individuals with each choice; show how to estimate the most frequently indicated choice; and show, using the Bayes factor, how to test that these proportions are the same over different levels of farmers' education. Our Bayesian procedure uses a sampling based method with independent samples.

KEY WORDS: multiple category choices; surveys; Bayes factor; Monte Carlo integration; sparse table

1. Introduction

A common type of multiple choice survey question provides respondents with a list of categories from which they are asked to mark all that apply to them. Analyses of survey responses to this type of question should give careful consideration to several concerns of which we mention only a few. First, in general, the list may be structured so that some combinations of categories and responses are intentionally excluded. For example, the response "none of the categories listed applies to me" might be excluded by questionnaire design. Second, for even a modest number of categories, say c, the investigator may not have a large enough budget for obtaining "adequate" numbers of responses for each of the possible 2^c combinations or sequences of responses, especially rare ones. Thus, he might have reason to expect that the contingency table in which the totals for each response sequence are tabulated would be sparse. Third, as with any survey data, nonresponse is likely to be a problem. Finally, the analysis should account for important design features such as stratification and clustering.

In this paper, we initiate the study of Bayesian analyses of the data resulting from questions of the above types. We consider the statistical problems of estimating the proportions of individuals who select each category under a general hypothesis and determining the most popular category. We also consider the problem of testing the hypothesis of the equality of the category proportions across strata which is solved using the Bayes factor. For stratified random sampling, we examine the situation where:

(1) the categorical variable has c levels;

(2) the population has been stratified into r subgroups;

(3) the response sequence corresponding to the case that none of the choices offered applies is excluded;

(4) there are no nonrespondents in the usual sense — that is, the only reason for no response is that none of the choices applies; and

(5) the contingency table for response sequences is sparse so that the use of the chi-squared test for homogeneity is questionable.

Key features of the data and the format in which they are often reported (which we subsequently refer to as the m-table) are illustrated by a situation where $r = 3$ and $c = 2$, for levels A and B of a categorical variable. [See Table 1.] For each stratum, the n-table contains the counts for the collection of mutually exclusive and exhaustive response sequences. However, the m-table entries contain the counts for each instance in which the response level was A or B or both. That is to say, the m-table counts do not include anything for the $(0, 0)$ sequence but the counts for A and B are both incremented when the response sequence is $(1, 1)$.

2. Bayesian Methodology

For r strata and c categories, $\{\pi_{j\ell}|j = 1, \ldots, r \text{ and } \ell = 1, \ldots, L\}$ will denote the n-table cell probabilities and $\{p_{jk}|j = 1, \ldots, r \text{ and } k = 1, \ldots, c\}$, the m-table cell probabilities. If E_{jk} denotes the event that category k is chosen in stratum j, we assume $\{E_{jk}|k = 1, \ldots, c\}$ is a set of independent events. We also assume that the response sequences for each stratum are the same, so that

$$\pi_{j\ell} = \prod_{k=1}^{c} p_{jk}^{I_{k\ell}} \, (1 - p_{jk})^{(1-I_{k\ell})}$$

*The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of the National Center for Health Statistics, Centers for Disease Control and Prevention.

Table 1: Illustration of the n-table and the m-table for two categories and three strata.

Stratum	n-table				m-table	
	Sequence				Level	
	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$	A	B
1	n_{11}	n_{12}	n_{13}	n_{14}	m_{11}	m_{12}
2	n_{21}	n_{22}	n_{23}	n_{24}	m_{21}	m_{22}
3	n_{31}	n_{32}	n_{33}	n_{34}	m_{31}	m_{32}

For $j = 1, 2, 3$, note that $m_{j1} = n_{j2} + n_{j4}$ and $m_{j2} = n_{j3} + n_{j4}$.

where $p_{jk} = \Pr\{E_{jk}\}$ and

$$I_{k\ell} = \begin{cases} 1, & \text{if category } k \text{ is selected} \\ & \text{in response sequence } \ell \\ 0, & \text{otherwise} \end{cases}$$

Note here that the indicator variables are notational devices and not random variables. In addition, we have that

$$p_{jk} = \sum_{\ell=1}^{L} \pi_{j\ell} I_{k\ell}.$$

Thus, one might expect to be able to use either the n-table or the m-table to make inferences for the p's or the π's.

2.1 The Likelihood Function

Under the null hypothesis that for each stratum j the $\{p_{jk} \text{ for } k = 1, \ldots, c\}$ are not the same, let $\underset{\sim}{n}'_j = (n_{j1}, n_{j2}, \ldots, n_{jL}) \stackrel{\text{def}}{=} (n_{j1}, \underset{\sim}{n}'_{j(1)})$, $\underset{\sim}{\pi}'_j = (\pi_{j1}, \pi_{j2}, \ldots, \pi_{jL})$ and $N_j = \underset{\sim}{n}'_j \underset{\sim}{\mathbf{1}}$ where $j = 1, \ldots, r$ and $\underset{\sim}{\mathbf{1}}$ is an L-vector of ones. n_{j1} will denote the n-table count for stratum j for the sequence $(0, 0, \ldots, 0)^{1\times c}$ and $I_{k1} = 0$ for $k = 1, 2, \ldots, c$. We assume independence over strata and that

$$\underset{\sim}{n}_j | \underset{\sim}{\pi}_j, N_j \sim \text{Multinomial}(N_j, \underset{\sim}{\pi}_j) \text{ for } j = 1, \ldots, r.$$

For $\underset{\sim}{p}'_j = (p_{j1}, \ldots, p_{jc})$, using the formulas for expressing the π's as functions of the p's, we then have that factor j of the likelihood is

$$p(\underset{\sim}{n}_j | \underset{\sim}{p}_j, N_j) = p(n_{j1}, \underset{\sim}{n}_{j(1)} | \underset{\sim}{p}_j, N_j)$$
$$= \frac{(n_{j1} + \sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}! \prod_{\ell=2}^{L} n_{j\ell}!} \times$$
$$\prod_{k=1}^{c} \left\{ p_{jk}^{\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}} (1 - p_{jk})^{n_{j1} + \sum_{\ell=2}^{L} n_{j\ell}(1 - I_{k\ell})} \right\}$$

Defining $\underset{\sim}{\vec{n}}'_1 = (n_{11}, n_{21}, \ldots, n_{r1})$, $\underset{\sim}{n}'_{(1)} = (\underset{\sim}{n}'_{1(1)}, \ldots, \underset{\sim}{n}'_{r(1)})$, $\underset{\sim}{p}' = (\underset{\sim}{p}'_1, \underset{\sim}{p}'_2, \ldots, \underset{\sim}{p}'_r)$ and $\underset{\sim}{N} = (N_1, N_2, \ldots, N_r)$, we have the combined data distribution

$$p(\underset{\sim}{\vec{n}}_1, \underset{\sim}{n}_{(1)} | \underset{\sim}{p}, \underset{\sim}{N}) = \prod_{j=1}^{r} p(\underset{\sim}{n}_j | \underset{\sim}{p}_j, N_j)$$
$$= \prod_{j=1}^{r} \left[\frac{(n_{j1} + \sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}! \prod_{\ell=2}^{L} n_{j\ell}!} \right.$$
$$\left. \times \prod_{k=1}^{c} \left\{ p_{jk}^{\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}} (1 - p_{jk})^{n_{j1} + \sum_{\ell=2}^{L} n_{j\ell}(1 - I_{k\ell})} \right\} \right]$$

Note that both $\underset{\sim}{p}$ and $\underset{\sim}{\vec{n}}_1$ (and, hence, $\underset{\sim}{N}$) are unknown.

2.2 Estimation of Cell Proportions and The Most Popular Choice

Here we assume that $\underset{\sim}{p}$ and $\underset{\sim}{N}$ are independent and that for $j = 1, \ldots, r$ and $k = 1, \ldots, c$

$$\begin{aligned} p_{jk} &\overset{\text{ind}}{\sim} U(0, 1) \\ Pr\{n_{j1}\} &= 1 \text{ for } n_{j1} \geq 0 \end{aligned}$$

Our use of uniform and improper priors does not cause a problem when estimating cell proportions and determining the most popular choice because the posteriors are proper. In what follows it will sometimes be more convenient to work with N_j rather than n_{j1} or *vice versa*.

2.2.1 Joint Posterior Density: Unrestricted Model

By Bayes Theorem, the joint posterior density

$$\pi(\underset{\sim}{p}, \underset{\sim}{\vec{n}}_1 | \underset{\sim}{n}_{(1)}) \propto$$
$$\prod_{j=1}^{r} \left[\frac{(n_{j1} + \sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}! \prod_{\ell=2}^{L} n_{j\ell}!} \prod_{k=1}^{c} \left\{ p_{jk}^{\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}} (1 - p_{jk})^{n_{j1} + \sum_{\ell=2}^{L} n_{j\ell}(1 - I_{k\ell})} \right\} \right]$$

When analyzing data from the Kansas Farm Survey, we used a sampling-based method to obtain samples from the joint posterior density. The posterior conditional density for $\underset{\sim}{p}$ is a product of the densities

$$p_{jk}|n_{j1}, \underset{\sim(1)}{n} \sim \text{Beta}\{\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}+1, n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell})+1\}$$

This is easily sampled once samples of $\vec{n}_1$ are obtained from $\pi(\vec{n}_1|\underset{\sim(1)}{n}) = \prod_{j=1}^{r} \pi(n_{j1}|\underset{\sim j(1)}{n})$, by construction, and

$$\pi(n_{j1}|\underset{\sim j(1)}{n}) \propto \frac{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}!} \prod_{k=1}^{c} \left\{ \frac{\{n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell})\}!}{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}+1)!} \right\}.$$

Since the terms on the right-hand side of this relationship are those of a convergent series, $\lim_{n_{j1}\to\infty} \pi(n_{j1}|\underset{\sim j(1)}{n}) = 0$; so, there must exist a value of n_{j1} beyond which the distribution has negligible probability to the right. We use this idea in approximating the distributions of the n_{j1}.

2.2.2 *Joint Posterior Density: Restricted Model*

When $p_{jk} = p_k$, we respecify $\underset{\sim}{p}' = (p_1, p_2, \ldots, p_c)$ so that

$$\pi(\underset{\sim}{p}, \vec{n}_1|\underset{\sim(1)}{n}) \propto \prod_{j=1}^{r} \left[\frac{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}! \prod_{\ell=2}^{L} n_{j\ell}!} \right] \times \prod_{k=1}^{c} \left\{ p_k^{\sum_{j=1}^{r}\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}} (1-p_k)^{\sum_{j=1}^{r}[n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell})]} \right\}$$

and

(i) $p_k|\vec{n}_1, \underset{\sim(1)}{n} \sim \text{Beta}(\nu_k+1, \gamma_k+1)$
where $\nu_k = \sum_{j=1}^{r}\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}$
and $\gamma_k = \sum_{j=1}^{r} \left\{ n_{j1} + \sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell}) \right\}$.

but

(ii) $\pi(\vec{n}_1|\underset{\sim(1)}{n}) \propto \left[\prod_{j=1}^{r} \frac{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}!}\right] \times \prod_{k=1}^{c} \left\{ \frac{[\sum_{j=1}^{r}\{n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell})\}]!}{(\sum_{j=1}^{r}\{n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}\}+1)!} \right\}$

From (ii) it is clear that, for the restricted model, the n_{j1} are correlated. When analyzing data from the Kansas Farm Survey, we used the Gibbs sampler to draw values for the n_{j1} from the distribution determined from (ii) and, again, used the composition method to draw values for the p_k.

2.3 Testing The Equality of Strata Proportions

Let $p_{M_1}(\underset{\sim(1)}{n})$ be the marginal likelihood of $\underset{\sim(1)}{n}$ under the unrestricted model, M_1. Let $p_{M_0}(\underset{\sim(1)}{n})$ be the marginal likelihood of $\underset{\sim(1)}{n}$ under the restricted model $M_0 : p_{jk} = p_k$ for each k and j. The Bayes factor is

$$BF = p_{M_0}(\underset{\sim(1)}{n}) / p_{M_1}(\underset{\sim(1)}{n})$$

For use of the Bayes factor in hypothesis testing, since we must work with marginal densities for $\underset{\sim(1)}{n}$, we need a proper prior on $\vec{n}_1$. Therefore, we take the n_{j1} to be independent with probability mass function

$$p(n_{j1}) = \frac{1}{a_j^{(0)}+1}$$

for $n_{j1} = 0, \ldots, a_j^{(0)}$ for some integer $a_j^{(0)}$ determined from the posterior distribution for n_{j1} such that

$$a_j^{(0)} = \max\{n_{j1}|\pi(n_{j1}|\underset{\sim j(1)}{n}) \geq .001\}$$

We also use $\text{Beta}(\alpha_k, \beta_k)$ as a more flexible choice than uniform priors for the p_{jk}, in the unrestricted model, and p_k, in the restricted model.

Upon specifying the joint densities under each model and integrating to obtain the marginals for $\underset{\sim(1)}{n}$ we have that, under the unrestricted model, M_1,

$$p_{M_1}(\underset{\sim(1)}{n}) = \sum_{\vec{n}_1 \in \mathcal{N}} \left[\prod_{j=1}^{r} \left\{ \frac{1}{a_j^{(0)}+1} \frac{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}! \prod_{\ell=2}^{L} n_{j\ell}!} \right. \right. \times \left. \left. \prod_{k=1}^{c} \left\{ \frac{(\sum_{\ell=2}^{L} n_{j\ell} I_{k\ell}+\alpha_k-1)!\{n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell})+\beta_k-1\}!}{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}+\alpha_k+\beta_k-1)! B(\alpha_k,\beta_k)} \right\} \right\} \right],$$

where $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$, $\mathcal{N} = \{\vec{n}_1 | 0 \leq n_{j1} \leq a_j^{(0)}$ for $j = 1, \ldots, r\}$; and, under the restricted model, M_0,

$$p_{M_0}(\underset{\sim(1)}{n}) = \sum_{\vec{n}_1 \in \mathcal{N}} \left[\prod_{j=1}^{r} \left\{ \frac{1}{a_j^{(0)}+1} \frac{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell})!}{n_{j1}! \prod_{\ell=2}^{L} n_{j\ell}!} \right\} \times \prod_{k=1}^{c} \left\{ \frac{\{\nu_k+\alpha_k-1\}![\gamma_k+\beta_k-1]!}{[\nu_k+\gamma_k+\alpha_k+\beta_k-1]! B(\alpha_k,\beta_k)} \right\} \right],$$

where ν_k and γ_k are as previously defined in (i) of section 2.2.2.

Direct computation of the Bayes factor is certainly possible but it was expected to be very time consuming since the cardinality of $\mathcal{N}$ is very large. For this reason, we opted for using Monte Carlo integration upon observing that the Bayes factor can be expressed as a conditional expectation; that is,

$$BF = E\left\{ \frac{W_0(\vec{n}_1)}{W_1(\vec{n}_1)} \,\middle|\, \underset{\sim(1)}{n} \right\},$$

where the expectation is taken with respect to the conditional distribution $\vec{n}_1|\underset{\sim(1)}{n}$ under model M_1. Here

$$W_0(\vec{n}_1) = \prod_{k=1}^{c} \left\{ \frac{\{\nu_k+\alpha_k-1\}![\gamma_k+\beta_k-1]!}{[\nu_k+\gamma_k+\alpha_k+\beta_k-1]!} \right\}$$

and

$W_1(\vec{n}_1) =$

$$\prod_{j=1}^{r}\prod_{k=1}^{c}\left\{\frac{(\sum_{\ell=2}^{L} n_{j\ell}I_{k\ell}+\alpha_k-1)!\{n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}(1-I_{k\ell})+\beta_k-1\}!}{(n_{j1}+\sum_{\ell=2}^{L} n_{j\ell}+\alpha_k+\beta_k-1)!}\right\}$$

Since we already have available samples from the posterior densities of the components of $\vec{n}_1$, for a large sample $\vec{n}_1^{(h)}, h = 1, \ldots, M$, we can obtain a simulation consistent estimator of BF as

$$\widehat{BF} = \frac{1}{M}\sum_{h=1}^{M} e^{W(\vec{n}_1^{(h)})},$$

where $W(\vec{n}_1^{(h)}) = \ln\{W_1(\vec{n}_1^{(h)})\} - \ln\{W_0(\vec{n}_1^{(h)})\}$.

When analyzing the Farm Survey Data, we worked with the logarithms of the Bayes factor for which we obtained approximate numerical standard errors (NSE's) using a first-order Taylor series expansion which yielded $NSE[\ln(\widehat{BF})] \doteq S_W/\sqrt{M}$, where $S_W^2 = \sum_{h=1}^{M}\{W(\vec{n}_1^{(h)}) - \bar{W}\}^2/(M-1)$ and $\bar{W} = \frac{1}{M}\sum_{h=1}^{M} W(\vec{n}_1^{(h)})$.

3. Analysis of The Kansas Farm Data

Our data were the replies to the question: What are your primary sources of veterinary information? Farmers were allowed to pick all the sources (professional consultant [A], veterinarian [B], state or local extension service [C], magazines [D], feed companies and representatives [E]) that applied to them. The data did not include the number of farmers to whom none of the choices applied and farmers were classified by education level: high school or less, vocational school, 2-year college, 4-year college, other. The inference problems of interest were: the number of farmers who said "no" to all choices, the probabilities of each choice, the most frequent choice and determining if the distribution of information sources differs by education level.

We now outline our approach for finding the most popular choice. For the unrestricted model, given $j = 1, \ldots, r$, let ρ_{jk} denote the ascending order rank of p_{jk} for $k = 1, \ldots, c$. (Then, $\rho_{jk'} = c$ if $\rho_{jk'} = \max\{p_{jk} | k = \ldots, c\}$.) If $\underset{\sim}{p}^{(h)}$ for $h = 1, 2, \ldots, M$ are M samples from the posterior distribution of $\underset{\sim}{p}$, let $r_{jk}^{(h)}$ denote the ascending order rank of $p_{jk}^{(h)}$ among $p_{j1}^{(h)}, \ldots, p_{jc}^{(h)}$ for each h and j. A 95% credible interval (CI) for ρ_{jk} is determined from $\{r_{jk}^{(h)} | h = 1, 2, \ldots, M\}$. The category that is the most popular is the one corresponding to the CI which lies to the right of all of the other intervals.

Tables 2, 3 and 4, respectively, provide for each education level estimates of the number of farmers who said "no" to all choices, the probabilities of each choice, and the most frequent choice. In the order of the farmer's education levels listed in the first paragraph of this section, the n_{j1} for $j = 1, \ldots, 5$ denote the estimated numbers of farmers for whom none of the veterinary information sources offered as questionnaire choices was applicable. In Table 3, subtables (a) - (e) provide results under the unrestricted model. In general, the posterior standard deviations are much smaller under the restricted model. Under the unrestricted model, note that the estimated proportions can be very different by education level; look, for example, at the estimates for source C in this table. In Table 4, note that, except for one case, all the CI's for each education level overlap under the unrestricted model indicating that there is no most popular choice. However, it is clear that for "2 year college" source A is least popular. Under the restricted model, it appears that choice D is the most popular one.

The Loughin and Scherer (1998) bootstrap procedure [denoted by LS] and the log-Bayes factor procedure [denoted by LBF] for testing the hypothesis of equality of proportions exhibit a certain consistency with respect to each other when applied correctly or incorrectly. We now indicate this in our discussion of the numerical results. The LS method gives a p-value of 0.047 with a numerical standard error of 0.003; thus, it is marginally significant. The sample-based value for the LBF was 12.21 with a numerical standard error of 0.0268 showing very strong evidence in favor of the restricted model. [See Kass and Raftery (1995) for a discussion of Bayes factors.] When examining a reduced table which involved only the counts for response sequences in which farmers replied that one, and only one, of the sources provided veterinary information, the chi-squared p-value of 0.479 showed no evidence against the restricted model and the LBF value of 4.79 indicated strong (but not very strong) evidence in favor of that model. There also seems to be a consistent pattern when the testing procedures are incorrectly applied to m-table cross-classifications. The chi-squared value reported by Loughin and Scherer was 0.602 (no evidence) and the corresponding LBF value was 12.44 (very strong evidence).

REFERENCES

Agresti, A. and Liu, I. (1999), "Modeling a Categorical Variable Allowing Arbitrarily Many Category Choices," *Biometrics*, **29**, 403-434.

Agresti, A. and Liu, I. (2001), "Strategies for Modeling a Categorical Variable Allowing Multiple Category Choices," *Sociological Methods and Research*, **29**, 403-434.

Berger J. O. and Pericchi, L. R. (1996), "The Intrinsic Bayes Factor for Model Selection and Prediction," *Journal of the American Statistical Association*, **91**, 109-122.

Bilder, C. R. and Loughin, T. M. (2004), "Testing for Marginal Independence between two Categorical Variables with Multiple Responses," *Biometrics*, **60**, 241-248.

Bilder, C. R. and Loughin, T. M. (2002), "Testing for Conditional Multiple Marginal Independence," *Biometrics*, **58**, 200-208.

Gelfand, A. E. and Smith, A. F. M.(1990), "Sampling-based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, **85**, 398-409.

Kass, R. and Raftery, A. (1995), "Bayes factors," *Journal of the American Statistical Association*, **90**, 773-795.

Loughin, T. M. and Scherer, P. N. (1998), "Testing for Association in Contingency Tables with Multiple Column Responses," *Biometrics*, **54**, 630-637.

Nandram, B. (2005), "A Bayesian Subset Analysis of Sensory Evaluation Data," *Journal of Modern Applied Statistical Methods*, **4**, 482-499.

Nandram, B. (1997), "Bayesian inference for the best ordinal multinomial population in a taste test," **in Case Studies in Bayesian Statistics**, eds.: C. Gatsonis, J. S. Hodges, R. E. Kass, R. McCulloch and N. D. Singpurwalla, Volume III, Springer-Verlag, 399-418.

Nandram, B. and Choi, J. W. (2006), "A Bayesian analysis of a two-way categorical table incorporating intra-class correlation," *Journal of Statistical Computation and Simulation*, 76, 233-249.

Table 2: Estimates of the number of farmers with 'no' on all choices by model

	Unrestricted Model				Restricted Model			
EDL	PM	PSD	NSE	INT	PM	PSD	NSE	INT
n_{11}	10.7	4.29	0.040	(4.0, 20.0)	20.1	5.66	0.064	(10.0, 32.0)
n_{21}	2.9	2.40	0.025	(0.0, 9.0)	3.8	2.24	0.023	(0.0, 9.0)
n_{31}	5.2	3.22	0.032	(1.0, 13.0)	7.2	3.13	0.035	(2.0, 14.0)
n_{41}	46.6	12.48	0.124	(26.0, 74.0)	25.7	6.55	0.070	(14.0, 40.0
n_{51}	1.7	1.75	0.017	(0.0, 6.0)	3.4	2.09	0.021	(0.0, 8.0)

NOTE: EDL - education level, PM - posterior mean, PSD - posterior standard deviation, NSE - numerical standard error, INT - 95% credible intervals. We do not round out to integers.

Table 3: Posterior Distributions of Sources by Education Level & Model

Source	PM	PSD	NSE	INT	PM	PSD	NSE	INT
	(a) High School				(b) Vocational School			
A	.199	.040	.000	(.127, .284)	.147	.078	.001	(.032, .331)
B	.388	.051	.001	(.292, .489)	.339	.108	.001	(.154, .566)
C	.298	.047	.000	(.210, .396)	.436	.116	.001	(.221, .671)
D	.478	.054	.001	(.374, .584)	.437	.116	.001	(.222, .668)
E	.408	.052	.001	(.311, .511)	.243	.096	.001	(.085, .453)
	(c) Two-year College				(d) Four-year College			
A	.053	.036	.000	(.007, .144)	.125	.027	.000	(.077, .183)
B	.370	.083	.001	(.218, .543)	.187	.034	.000	(.126, .258)
C	.290	.077	.001	(.153, .453)	.255	.040	.000	(.182, .338)
D	.474	.088	.001	(.305, .646)	.337	.045	.000	(.253, .429)
E	.394	.085	.001	(.239, .565)	.186	.034	.000	(.125, .257)
	(e) Other				(f) Restricted Model			
A	.228	.099	.001	(.069, .445)	.139	.020	.000	(.103, .179)
B	.285	.108	.001	(.102, .521)	.281	.027	.000	(.230, .335)
C	.510	.125	.001	(.272, .748)	.296	.028	.000	(.244, .352)
D	.398	.119	.001	(.181, .643)	.408	.031	.000	(.349, .469)
E	.399	.119	.001	(.186, .650)	.290	.028	.000	(.238, .346)

Table 4: Posterior Distributions of Most Popular Source by EDL & Model

Source	PM	PSD	NSE	INT	PM	PSD	NSE	INT
		(a) High School				(b) Vocational School		
A	1.06	0.233	0.002	(1.00, 2.00)	1.30	0.612	0.006	(1.00, 3.00)
B	3.40	0.739	0.007	(2.00, 5.00)	3.20	1.003	0.010	(1.00, 5.00)
C	2.09	0.485	0.005	(1.00, 3.00)	4.14	0.912	0.010	(2.00, 5.00)
D	4.74	0.547	0.006	(3.00, 5.00)	4.15	0.890	0.009	(2.00, 5.00)
E	3.71	0.746	0.008	(2.00, 5.00)	2.20	0.899	0.009	(1.00, 4.00)
		(c) Two-year College				(d) Four-year College		
A	1.02	0.048	0.000	(1.00, 1.00)	1.12	0.370	0.004	(1.00, 2.00)
B	3.37	0.927	0.009	(2.00, 5.00)	2.51	0.690	0.007	(1.00, 4.00)
C	2.44	0.730	0.007	(2.00, 4.00)	3.91	0.473	0.005	(3.00, 5.00)
D	4.53	0.742	0.008	(3.00, 5.00)	4.95	0.231	0.002	(4.00, 5.00)
E	3.66	0.920	0.010	(2.00, 5.00)	2.50	0.696	0.007	(1.00, 4.00)
		(e) Other				(f) Restricted Model		
A	1.63	0.899	0.009	(1.00, 4.00)	1.00	0.000	0.000	(1.00, 1.00)
B	2.20	1.064	0.011	(1.00, 5.00)	2.72	0.783	0.008	(2.00, 4.00)
C	4.39	0.887	0.009	(2.00, 5.00)	3.24	0.784	0.008	(2.00, 4.00)
D	3.39	1.121	0.012	(1.00, 5.00)	5.00	0.054	0.001	(5.00, 5.00)
E	3.39	1.122	0.011	(1.00, 5.00)	3.04	0.804	0.008	(2.00, 4.00)