

Hindawi Publishing Corporation
Journal of Probability and Statistics
Volume 2012, Article ID 617678, 26 pages
doi:10.1155/2012/617678

Research Article

Bayesian Approach to Zero-Inflated Bivariate Ordered Probit Regression Model, with an Application to Tobacco Use

Shiferaw Gurmu¹ and Getachew A. Dagne²

¹ Department of Economics, Andrew Young School of Policy Studies, Georgia State University, P.O. Box 3992, Atlanta, GA 30302, USA

² Department of Epidemiology and Biostatistics, College of Public Health, University of South Florida, Tampa, FL 33612, USA

Correspondence should be addressed to Shiferaw Gurmu, sgurmu@gsu.edu

Received 13 July 2011; Revised 18 September 2011; Accepted 2 October 2011

Academic Editor: Wenbin Lu

Copyright © 2012 S. Gurmu and G. A. Dagne. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a Bayesian analysis of bivariate ordered probit regression model with excess of zeros. Specifically, in the context of joint modeling of two ordered outcomes, we develop zero-inflated bivariate ordered probit model and carry out estimation using Markov Chain Monte Carlo techniques. Using household tobacco survey data with substantial proportion of zeros, we analyze the socioeconomic determinants of individual problem of smoking and chewing tobacco. In our illustration, we find strong evidence that accounting for excess zeros provides good fit to the data. The example shows that the use of a model that ignores zero-inflation masks differential effects of covariates on nonusers and users.

1. Introduction

This paper is concerned with joint modeling of two ordered data outcomes allowing for excess zeros. Economic, biological, and social science studies often yield data on two ordered categorical variables that are jointly dependent. Examples include the relationship between desired and excess fertility [1, 2], helmet use and motorcycle injuries [3], ownership of dogs and televisions [4], severity of diabetic retinopathy of the left and right eyes [5], and self-assessed health status and wealth [6]. The underlying response variables could be measured on an ordinal scale. It is also common in the literature to generate a categorical or grouped variable from an underlying quantitative variable and then use ordinal response regression model (e.g., [4, 5, 7]). The ensuing model is usually analyzed using the bivariate ordered probit model.

Many ordered discrete data sets are characterized by excess of zeros, both in terms of the proportion of nonusers and relative to the basic ordered probit or logit model. The zeros may be attributed to either corner solution to consumer optimization problem or errors in recording. In the case of individual smoking behavior, for example, the zeros may be recorded for individuals who never smoke cigarettes or for those who either used tobacco in the past or are potential smokers. In the context of individual patents applied for by scientists during a period of five years, zero patents may be recorded for scientists who either never made patent applications or for those who do but not during the reporting period [8]. Ignoring the two types of zeros for nonusers or nonparticipants leads to model misspecification.

The univariate as well as bivariate zero-inflated count data models are well established in the literature for example, Lambert [9], Gurmu and Trivedi [10], Mullahy [11], and Gurmu and Elder [12]. The recent literature presents a Bayesian treatment of zero-inflated Poisson models in both cross-sectional and panel data settings (see [13, 14], and references there in). By contrast, little attention has been given to the problem of excess zeros in the ordered discrete choice models. Recently, an important paper by Harris and Zhao [15] developed a zero-inflated univariate ordered probit model. However, the problem of excess zeros in ordered probit models has not been analyzed in the Bayesian framework. Despite recent applications and advances in estimation of bivariate ordered probit models [1–6], we know of no studies that model excess zeros in bivariate ordered probit models.

This paper presents a Bayesian analysis of bivariate ordered probit model with excess of zeros. Specifically, we develop a zero-inflated ordered probit model and carry out the analysis using the Bayesian approach. The Bayesian analysis is carried out using Markov Chain Monte Carlo (MCMC) techniques to approximate the posterior distribution of the parameters. Bayesian analysis of the univariate zero-inflated ordered probit will be treated as a special case of the zero-inflated bivariate order probit model. The proposed models are illustrated by analyzing the socioeconomic determinants of individual choice problem of bivariate ordered outcomes on smoking and chewing tobacco. We use household tobacco prevalence survey data from Bangladesh. The observed proportion of zeros (those identifying themselves as nonusers of tobacco) is about 76% for smoking and 87% for chewing tobacco.

The proposed approach is useful for the analysis of ordinal data with natural zeros. The empirical analysis clearly shows the importance of accounting for excess zeros in ordinal qualitative response models. Accounting for excess zeros provides good fit to the data. In terms of both the signs and magnitudes of marginal effects, various covariates have differential impacts on the probabilities associated with the two types of zeros, nonparticipants and zero-consumption. The usual analysis that ignores excess of zeros masks these differential effects, by just focusing on observed zeros. The empirical results also show the importance of taking into account the uncertainty in the parameter estimates. Another advantage of the Bayesian approach to modeling excess zeros is the flexibility, particularly computational, of generalizing to multivariate ordered response models.

The rest of the paper is organized as follows. Section 2 describes the proposed zero-inflated bivariate probit model. Section 3 presents the MCMC algorithm and model selection procedure for the model. An illustrative application using household tobacco consumption data is given in Section 4. Section 5 concludes the paper.

2. Zero-Inflated Bivariate Ordered Probit Model

2.1. The Basic Model

We consider the basic Bayesian approach to a bivariate latent variable regression model with excess of zeros. To develop notation, let \tilde{y}_{1i}^* and \tilde{y}_{2i}^* denote the bivariate latent variables. We

consider two observed ordered response variables \tilde{y}_{1i} and \tilde{y}_{2i} taking on values $0, 1, \dots, J_r$, for $r = 1, 2$. Define two sets of cut-off parameters $\alpha^r = (\alpha_{r2}, \alpha_{r3}, \dots, \alpha_{rJ_r})$, $r = 1, 2$, where the restrictions $\alpha_{r0} = -\infty$, $\alpha_{rJ_r+1} = \infty$, and $\alpha_{r1} = 0$ have been imposed. We assume that $(\tilde{y}_{1i}^*, \tilde{y}_{2i}^*)' \equiv \tilde{\mathbf{y}}_i^*$ follows a bivariate regression model

$$\tilde{\mathbf{y}}_{ri}^* = \mathbf{x}'_{ri} \boldsymbol{\beta}_r + \varepsilon_{ri}, \quad r = 1, 2, \quad (2.1)$$

where \mathbf{x}_{ri} is a K_r -variate of regressors for the i th individual ($i = 1, \dots, N$) and ε_{ri} are the error terms. For subsequent analysis, let $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2)'$, $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i}, \varepsilon_{2i})'$, and

$$\mathbf{X}_i = \begin{pmatrix} \mathbf{x}'_{1i} & 0' \\ 0' & \mathbf{x}'_{2i} \end{pmatrix}. \quad (2.2)$$

Analogous to the univariate case, the observed bivariate-dependent variables are defined as

$$\tilde{\mathbf{y}}_{ri} = \begin{cases} 0 & \text{if } \tilde{\mathbf{y}}_{ri}^* \leq 0, \\ 1 & \text{if } 0 < \tilde{\mathbf{y}}_{ri}^* \leq \alpha_{r2}, \\ j & \text{if } \alpha_{rj} < \tilde{\mathbf{y}}_{ri}^* \leq \alpha_{rj+1}, \quad j = 2, 3, \dots, J_r - 1, \\ J_r & \text{if } \tilde{\mathbf{y}}_{ri}^* \leq \alpha_{rJ_r}, \end{cases} \quad (2.3)$$

where $r = 1, 2$. Let $\tilde{\mathbf{y}}_i = (\tilde{y}_{1i}, \tilde{y}_{2i})'$.

We introduce inflation at the point $(\tilde{y}_{1i} = 0, \tilde{y}_{2i} = 0)$, called the zero-zero state. As in the univariate case, define the participation model:

$$\begin{aligned} s_i^* &= \mathbf{z}'_i \boldsymbol{\gamma} + \mu_i, \\ s_i &= I(s_i^* > 0). \end{aligned} \quad (2.4)$$

In the context of the zero-inflation model, the observed response random vector $\mathbf{y}_i = (y_{1i}, y_{2i})'$ takes the form

$$\mathbf{y}_i = s_i \tilde{\mathbf{y}}_i. \quad (2.5)$$

We observe $\mathbf{y}_i = \mathbf{0}$ when either the individual is a non-participant ($s_i = 0$) or the individual is a zero-consumption participant ($s_i = 1$ and $\tilde{\mathbf{y}}_i = \mathbf{0}$). Likewise, we observe positive outcome (consumption) when the individual is a positive consumption participant for at least one good ($s_i = 1$ and $\tilde{\mathbf{y}}_i \neq \mathbf{0}$).

Let $\Phi(a)$ and $\phi(a)$ denote the respective cumulative distribution and probability density functions of standardized normal evaluated at a . Assuming normality and that μ_i is

uncorrelated with $(\varepsilon_{1i}, \varepsilon_{2i})$, but $\text{corr}(\varepsilon_{1i}, \varepsilon_{2i}) = \rho_{12} \neq 0$, and each component with unit variance, the zero-inflated bivariate ordered probit (ZIBOP) distribution is

$$f_b(\mathbf{y}_i^*, \mathbf{y}_i, s_i^*, s_i \mid \mathbf{X}_i, \mathbf{z}_i, \Psi) = \begin{cases} \Pr(s_i = 0) + (1 - \Pr(s_i = 0))\Pr(\tilde{y}_{1i} = 0, \tilde{y}_{2i} = 0), & \text{for } (\tilde{y}_{1i}, \tilde{y}_{2i}) = (0, 0) \\ (1 - \Pr(s_i = 0))\Pr(\tilde{y}_{1i} = j, \tilde{y}_{2i} = l), & \text{for } (\tilde{y}_{1i}, \tilde{y}_{2i}) \neq (0, 0), \end{cases} \quad (2.6)$$

where $j = 0, 1, \dots, J_1$, $l = 0, 1, \dots, J_2$, $\Pr(s_i = 0) = \Phi(-\mathbf{z}'_i\boldsymbol{\gamma})$, $\Pr(s_i = 1) = \Phi(-\mathbf{z}'_i\boldsymbol{\gamma})$. Further, for $(\tilde{y}_{1i}, \tilde{y}_{2i}) = (0, 0)$ in (2.6), we have $\alpha_{r0} = -\infty$, $\alpha_{r1} = 0$ for $r = 1, 2$ so that

$$\Pr(\tilde{y}_{1i} = 0, \tilde{y}_{2i} = 0) = \Phi_2(-\mathbf{x}'_{1i}\boldsymbol{\beta}_1, -\mathbf{x}'_{2i}\boldsymbol{\beta}_2, \rho_{12}), \quad (2.7)$$

where $\Phi_2(\cdot)$ is the cdf for the standardized bivariate normal. Likewise, $\Pr(\tilde{y}_{1i} = j, \tilde{y}_{2i} = l)$ in (2.6) are given by

$$\begin{aligned} \Pr(\tilde{y}_{1i} = j, \tilde{y}_{2i} = l) &= \Phi_2(\alpha_{1j+1} - \mathbf{x}'_{1i}\boldsymbol{\beta}_1, \alpha_{2l+1} - \mathbf{x}'_{2i}\boldsymbol{\beta}_2; \rho_{12}) \\ &\quad - \Phi_2(\alpha_{1j} - \mathbf{x}'_{1i}\boldsymbol{\beta}_1, \alpha_{2l} - \mathbf{x}'_{2i}\boldsymbol{\beta}_2, \rho_{12}) \quad \text{for } j = 1, \dots, J_1 - 1; l = 1, \dots, J_2 - 1; \\ \Pr(\tilde{y}_{1i} = J_1, \tilde{y}_{2i} = J_2) &= 1 - \Phi_2(\alpha_{1J_1} - \mathbf{x}'_{1i}\boldsymbol{\beta}_1, \alpha_{2J_2} - \mathbf{x}'_{2i}\boldsymbol{\beta}_2, \rho_{12}). \end{aligned} \quad (2.8)$$

The ensuing likelihood contribution for N -independent observations is

$$\begin{aligned} \mathcal{L}_b(\mathbf{y}^*, \mathbf{y}, s^*, s \mid \mathbf{X}, \mathbf{z}, \Psi_b) &= \prod_{i=1}^N \prod_{(j,l)=(0,0)} [\Pr(s_i = 0) + (1 - \Pr(s_i = 0))\Pr(\tilde{y}_{1i} = 0, \tilde{y}_{2i} = 0)]^{d_{ijl}} \\ &\quad \times \prod_{i=1}^N \prod_{(j,l) \neq (0,0)} [(1 - \Pr(s_i = 0))\Pr(\tilde{y}_{1i} = j, \tilde{y}_{2i} = l)]^{d_{ijl}}, \end{aligned} \quad (2.9)$$

where $d_{ijl} = 1$ if $\tilde{y}_{1i} = j$ and $\tilde{y}_{2i} = l$, and $d_{ijl} = 0$ otherwise. Here, the vector Ψ_b consists of $\boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha^1, \alpha^2$, and the parameters associated with the trivariate distribution of $(\boldsymbol{\varepsilon}, \boldsymbol{\mu})$.

Regarding identification of the parameters in the model defined by (2.1) through (2.5) with normality assumption, we note that the mean parameter (joint choice probability associate with the observed response vector \mathbf{y}_i) depends nonlinearly on the probability of zero inflation ($\Phi(-\mathbf{z}'_i\boldsymbol{\gamma})$) and choice probability ($\Pr(\tilde{y}_{1i} = j, \tilde{y}_{2i} = l)$) coming from the BOP submodel. Since the likelihood function for ZIBOP depends separately on the two regression components, the parameters of ZIBOP model with covariates are identified as long as the model is estimated by full maximum likelihood method. The same or different sets of covariates can affect the two components via \mathbf{z}_i and \mathbf{x}_{ri} . When using quasi-likelihood estimation or generalized estimating equations methods rather than full ML, the class of identifiable zero-inflated count and ordered data models is generally more restricted; see, for example, Hall and Shen [16] and references there in. Although the parameters in the ZIBOP

model above are identified through a nonlinear functional form estimated by ML, for more robust identification we can use traditional exclusion restrictions by including instrumental variables in the inflation equation, but excluding them from the ordered choice submodel. We follow this strategy in the empirical section.

About 2/3 of the observations in our tobacco application below have a double-zero-state, ($y_1 = 0, y_2 = 0$). Consequently, we focused on a mixture constructed from a point mass at $(0,0)$ and a bivariate ordered probit. In addition to allowing for inflation in the double-zero-state, our approach can be extended to allow for zero-inflation in each component.

2.2. Marginal Effects

It is common to use marginal or partial effects to interpret covariate effects in nonlinear models; see, for example, Liu et al. [17]. Due to the nonlinearity in zero-inflated ordered response models and in addition to estimation of regression parameters, it is essential to obtain the marginal effects of changes in covariates on various probabilities of interest. These include the effects of covariates on probability of nonparticipation (zero-inflation), probability of participation, and joint and/or marginal probabilities of choice associated with different levels of consumption.

From a practical point of view, we are less interested in the marginal effects of explanatory variables on the joint probabilities of choice from ZIBOP. Instead, we focus on the marginal effects associated with the marginal distributions of y_{ri} for $r = 1, 2$. Define a generic (scalar) covariate w_i that can be a binary or approximately continuous variable. We obtain the marginal effects of a generic covariate w_i on various probabilities assuming that the regression results are based on ZIBOP. If w_i is a binary regressor, then the marginal effect of w_i on probability, say P , is the difference in the probability evaluated at 1 and 0, conditional on observable values of covariates: $P(w_i = 1) - P(w_i = 0)$. For continuous explanatory variables, the marginal effect is given by the partial derivative of the probability of interest with respect to w_i , $\partial P(\cdot) / \partial w_i$.

Regressor w_i can be a common covariate in vectors of regressors \mathbf{x}_{ri} and \mathbf{z}_i or appears in either \mathbf{x}_{ri} or \mathbf{z}_i . Focusing on the continuous regressor case, the marginal effects of w_i in each of the three cases are presented below. First, consider the case of common covariate in participation and main parts of the model, that is, w_i in both \mathbf{x}_{ri} and \mathbf{z}_i . The marginal effect on the probability of participation is given by

$$M_i(s_i = 1) = \frac{\partial \Pr(s_i = 1)}{\partial w_i} = \phi(\mathbf{z}_i' \boldsymbol{\gamma}) \gamma_{w_i, r} \quad (2.10)$$

where again $\phi(\cdot)$ is the probability density function (pdf) of the standard normal distribution and γ_{w_i} is the coefficient in the inflation part associated with variable w_i . In terms of the zeros category, the effect on the probability of nonparticipation (zero inflation) is

$$M_i(s_i = 0) = \frac{\partial \Pr(s_i = 0)}{\partial w_i} = -\phi(-\mathbf{z}_i' \boldsymbol{\gamma}) \gamma_{w_i, r} \quad (2.11)$$

while

$$\begin{aligned} M_i(s = 1, \tilde{y}_{ri} = 0) &= \frac{\partial \Pr(s_i = 1) \Pr(\tilde{y}_{ri} = 0)}{\partial w_i} \\ &= \Phi(-\mathbf{x}'_{ri} \boldsymbol{\beta}_r) \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i} - \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \phi(-\mathbf{x}'_{ri} \boldsymbol{\beta}_r) \beta_{rw_i}, \quad r = 1, 2, \end{aligned} \quad (2.12)$$

represents the marginal effect on the probability of zero-consumption. Here the scalar β_{rw_i} is the coefficient in the main part of the model associated with w_i .

Continuing with the case of common covariate, the marginal effects of w_i on the probabilities of choice are given as follows. First, the total marginal effect on the probability of observing zero-consumption is obtained as a sum of the marginal effects in (2.11) and (2.12); that is,

$$M_i(y_{ri} = 0) = [\Phi(-\mathbf{x}'_{ri} \boldsymbol{\beta}_r) - 1] \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i} - \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \phi(-\mathbf{x}'_{ri} \boldsymbol{\beta}_r) \beta_{rw_i}. \quad (2.13)$$

The effects for the remaining choices for outcomes $r = 1, 2$ are as follows:

$$\begin{aligned} M_i(y_{ri} = 1) &= [\Phi(\alpha_{r2} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \Phi(-\mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i} \\ &\quad - \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) [\phi(\alpha_{r2} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \phi(-\mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \beta_{rw_i}; \\ M_i(y_{ri} = j) &= [\Phi(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \Phi(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i} \\ &\quad - \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) [\phi(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \phi(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \beta_{rw_i}, \quad \text{for } j = 2, \dots, J_r - 1; \\ M_i(y_{ri} = J_r) &= [1 - \Phi(\alpha_{r,J_r} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i} + \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \phi(\alpha_{r,J_r} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) \beta_{rw_i}. \end{aligned} \quad (2.14)$$

Now consider case 2, where a generic independent variable w_i is included only in \mathbf{x}_{ri} , the main part of the model. In this case, covariate w_i has obviously no direct effect on the inflation part. The marginal effects of w_i on various choice probabilities can be presented as follows:

$$\begin{aligned} M_i(y_{ri} = j) &= \frac{\partial \Pr(y_{ri} = j)}{\partial w_i} \\ &= -\Phi(\mathbf{z}'_i \boldsymbol{\gamma}) [\phi(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \phi(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \beta_{rw_i}, \quad \text{for } j = 0, 1, \dots, J_r, \end{aligned} \quad (2.15)$$

with $\alpha_{r0} = -\infty$, $\alpha_{r1} = 0$, and $\alpha_{r,J_r+1} = \infty$. The marginal effects in (2.15) can be obtained by simply setting $\gamma_{w_i} = 0$ in (2.13) and (2.14).

For case 3, where w_i appears only in \mathbf{z}_i , its marginal effects on participation components given in (2.10) and (2.11) will not change. Since $\beta_{rw_i} = 0$ in case 3, the partial effects of w_i on various choice probabilities take the form:

$$M_i(y_{ri} = j) = [\Phi(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \Phi(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i} \quad \text{for } j = 0, 1, \dots, J_r. \quad (2.16)$$

Again, we impose the restrictions $\alpha_{r0} = -\infty$, $\alpha_{r1} = 0$ and $\alpha_{r,J_r+1} = \infty$.

As noted by a referee, it is important to understand the sources of covariate effects and the relationship between the marginal effects and the coefficient estimates. Since

$$\Pr(y_{ri} = j) = [\Pr(s_i = 1)\Pr(\tilde{y}_{ri} = j)] \quad (2.17)$$

for $j = 0, 1, \dots, J_r$, the total effect of a generic covariate w_i on probability of consumption at level j comes from two (weighted) sources: the participation part ($\Pr(s_i = 1)$) and the main ordered probit part ($\Pr(\tilde{y}_{ri} = j)$) such that

$$\frac{\partial \Pr(s_i = 1)}{\partial w_i} = \phi(\mathbf{z}'_i \boldsymbol{\gamma}) \gamma_{w_i}; \quad (2.18)$$

$$\frac{\partial \Pr(\tilde{y}_{ri} = j)}{\partial w_i} = -[\phi(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \phi(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)] \beta_{rw_i} \quad (2.19)$$

with $\alpha_{r0} = -\infty$, $\alpha_{r1} = 0$, s and $\alpha_{r,J_r+1} = \infty$. This shows that $\text{sign}(\gamma_{w_i})$ is the same as $\text{sign}(\partial \Pr(s_i = 1) / \partial w_i)$ —the participation effect in (2.18)—but $\text{sign}(\beta_{rw_i})$ is not necessarily the same as the sign of $(\partial \Pr(\tilde{y}_{ri} = j) / \partial w_i)$. The latter is particularly true in the left tail of the distribution, where the coefficient (β_{rw_i}) and the main (unweighted) effect in (2.19) have opposite signs because

$$\{-[\phi(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r) - \phi(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)]\} \equiv \varpi \quad (2.20)$$

is negative. In this case, a positive effect coming from the main part requires β_{rw_i} to be negative. By contrast, ϖ is positive in the right tail, but can be positive or negative when the terms $(\alpha_{rj} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)$ and $(\alpha_{r,j+1} - \mathbf{x}'_{ri} \boldsymbol{\beta}_r)$ are on the opposite sides of the mode of the distribution. This shows that a given covariate can have opposite effects in the participation and main models. Since the total effect of an explanatory variable on probability of choice is a weighted average of (2.18) and (2.19), interpretation of results should focus on marginal effects of covariates rather than the signs of estimated coefficients. This is the strategy adopted in the empirical analysis below.

2.3. A Special Case

Since the zero-inflated univariate ordered probit (ZIUOP) model has not been analyzed previously in the Bayesian framework, we provide a brief sketch of the basic framework for ZIUOP. The univariate ordered probit model with excess of zeros can be obtained as a special case of the ZIBOP model presented previously. To achieve this, let $\rho_{12} = 0$ in the ZIBOP model and focus on the first ordered outcome with $r = 1$. In the standard ordered response approach, the model for the latent variable \tilde{y}_{1i}^* is given by (2.1) with $r = 1$. The observed ordered variable \tilde{y}_{1i} can be presented compactly as

$$\tilde{y}_{1i} = \sum_{j=0}^J j I(\alpha_{1j} < \tilde{y}_{1i}^* \leq \alpha_{1j+1}), \quad (2.21)$$

where $I(w \in A)$ is the indicator function equal to 1 or 0 according to whether $w \in A$ or not. Again $\alpha_{10}, \alpha_{11}, \dots, \alpha_{1J_1}$ are unknown threshold parameters, where we set $\alpha_{10} = -\infty$, $\alpha_{11} = 0$, and $\alpha_{1J_1+1} = \infty$.

Zero-inflation is now introduced at point $\tilde{y}_{1i} = 0$. Using the latent variable model (2.4) for the zero inflation, the observed binary variable is given by $s_i = I(s_i^* > 0)$, where $I(s_i^* > 0) = 1$ if $s_i^* > 0$, and 0 otherwise. In regime 1, $s_i = 1$ or $s_i^* > 0$ for participants (e.g., smokers), while, in regime 0, $s_i = 0$ or $s_i^* \leq 0$ for nonparticipants. In the context of the zero-inflation model, the observed response variable takes the form $y_{1i} = s_i \tilde{y}_{1i}$. We observe $y_{1i} = 0$ when either the individual is a non-participant ($s_i = 0$) or the individual is a zero-consumption participant ($s_i = 1$ and $\tilde{y}_{1i} = 0$). Likewise, we observe positive outcome (consumption) when the individual is a positive consumption participant ($s_i = 1$ and $\tilde{y}_{1i}^* > 0$).

Assume that ϵ_1 and μ are independently distributed. Harris and Zhao [15] also consider the case where ϵ_1 and μ are correlated. In the context of our application, the correlated model did not provide improvements over the uncorrelated ZIOP in terms of deviance information criterion. The zero-inflated ordered multinomial distribution, say $\Pr(y_{1i})$, arises as a mixture of a degenerate distribution at zero and the assumed distribution of the response variable \tilde{y}_{1i} as follows:

$$f_1(y_{1i}^*, y_{1i}, s_i^*, s_i | \mathbf{x}_{1i}, \mathbf{z}_i, \Psi_1) = \begin{cases} \Pr(s_i = 0) + \Pr(s_i = 1)\Pr(\tilde{y}_{1i} = 0), & \text{for } j = 0 \\ \Pr(s_i = 1)\Pr(\tilde{y}_{1i} = j), & \text{for } j = 1, 2, \dots, J_1, \end{cases} \quad (2.22)$$

where, for any parameter vector Ω_{10} associated with the distribution of (ϵ_1, μ) , $\Psi_1 = (\boldsymbol{\beta}_1, \boldsymbol{\gamma}, \boldsymbol{\alpha}^1, \Omega_{10})$ with $\boldsymbol{\alpha}^1 = (\alpha_{12}, \dots, \alpha_{1J_1})$. For simplicity, dependence on latent variables, covariates, and parameters has been suppressed on the right-hand side of (2.22). The likelihood based on N -independent observations takes the form

$$\begin{aligned} \mathcal{L}_1(\mathbf{y}_1^*, \mathbf{y}_1, \mathbf{s}^*, \mathbf{s} | \mathbf{x}_1, \mathbf{z}, \Psi_1) &= \prod_{i=1}^N \prod_{j=0}^{J_1} [\Pr(y_{1i} = j | \mathbf{x}_{1i}, \mathbf{z}_i, \Psi_1)]^{d_{ij}} \\ &= \prod_{i=1}^N \prod_{j=0}^{J_1} [\Pr(s_i = 0) + \Pr(s_i = 1)\Pr(\tilde{y}_{1i} = j)]^{d_{ij}} \\ &\quad \times \prod_{i=1}^N \prod_{j>0} [\Pr(s_i = 1)\Pr(\tilde{y}_{1i} = j)]^{d_{ij}}, \end{aligned} \quad (2.23)$$

where, for example, $\mathbf{y}_1^* = (y_1^*, \dots, y_N^*)'$, and $d_{ij} = 1$ if individual i chooses outcome j , or $d_{ij} = 0$ otherwise.

Different choices of the specification of the joint distribution of (ϵ_{1i}, μ_i) give rise to various zero-inflated ordered response models. For example, if the disturbance terms in the latent variable equations are normally distributed, we get the zero-inflated ordered probit model of Harris and Zhao [15]. The zero-inflated ordered logit model can be obtained by assuming that ϵ_{1i} and μ_i are independent, each of the random variables following the logistic distribution with cumulative distribution function defined as $\Lambda(a) = e^a / (1 + e^a)$. Unlike the ordered probit framework, the ordered logit cannot lend itself easily to allow for correlation

between bivariate discrete response outcomes. Henceforth, we focus on the ordered probit paradigm in both univariate and bivariate settings.

Assuming that ϵ_{1i} and μ_i are independently normally distributed, each with mean 0 and variance 1, the required components in (2.22) and consequently (2.23) are given by:

$$\begin{aligned}\Pr(s_i = 0) &= \Phi(-\mathbf{z}'_i \boldsymbol{\gamma}), \\ \Pr(\tilde{y}_{1i} = 0) &= \Phi(-\mathbf{x}'_{1i} \boldsymbol{\beta}_1), \\ \Pr(\tilde{y}_{1i} = j) &= \Phi(\alpha_{1j+1} - \mathbf{x}'_{1i} \boldsymbol{\beta}_1) - \Phi(\alpha_{1j} - \mathbf{x}'_{1i} \boldsymbol{\beta}_1), \quad \text{for } j = 1, \dots, J_1 - 1 \text{ with } \alpha_{10} = 0, \\ \Pr(\tilde{y}_{1i} = J_1) &= 1 - \Phi(\alpha_{1J_1} - \mathbf{x}'_{1i} \boldsymbol{\beta}_1).\end{aligned}\tag{2.24}$$

The marginal effects for the univariate ZIOP are given by Harris and Zhao [15]. Bayesian analysis of the univariate ZIOP will be treated as a special case of the zero-inflated bivariate order probit model in the next section.

3. Bayesian Analysis

3.1. Prior Distributions

The Bayesian hierarchical model requires prior distributions for each parameter in the model. For this purpose, we can use noninformative conjugate priors. There are two reasons for adopting noninformative conjugate priors. First, we prefer to let the data dictate the inference about the parameters with little or no influence from prior distributions. Secondly, the noninformative priors facilitate resampling using Markov Chain Monte Carlo algorithm (MCMC) and have nice convergence properties. We assume noninformative (vague or diffuse) normal priors for regression coefficients $\boldsymbol{\beta}$, with mean $\boldsymbol{\beta}^*$ and variance $\boldsymbol{\Omega}_\beta$ which are chosen to make the distribution proper but diffuse with large variances. Similarly, $\boldsymbol{\gamma} \sim N(\boldsymbol{\gamma}^*, \boldsymbol{\Omega}_\gamma)$.

In choosing prior distributions for the threshold parameters, α 's, caution is needed because of the order restriction on them. One way to avoid the order restriction is to reparameterize them. Following Chib and Hamilton [18] treatment in the univariate ordered probit case, we reparameterize the ordered threshold parameters

$$\tau_{r2} = \log(\alpha_{r2}); \quad \tau_{rj} = \log(\alpha_{rj} - \alpha_{rj-1}), \quad j = 3, \dots, J_r; \quad r = 1, 2 \tag{3.1}$$

with the inverse map

$$\alpha_{rj} = \sum_{m=2}^j \exp(\tau_{rm}), \quad j = 2, \dots, J_r; \quad r = 1, 2. \tag{3.2}$$

For $r = 1, 2$, let $\boldsymbol{\tau}^r = (\tau_{r2}, \tau_{r3}, \dots, \tau_{rJ_r})'$ so that $\boldsymbol{\tau} = (\boldsymbol{\tau}^1, \boldsymbol{\tau}^2)$. We choose normal prior $\boldsymbol{\tau} \sim N(\boldsymbol{\tau}^*, \boldsymbol{\Omega}_\tau)$ without order restrictions for τ_r 's.

The only unknown parameter associate with the distribution of $(\boldsymbol{\epsilon}, \boldsymbol{\mu})$ in (2.1) and (2.4) is ρ_{12} , the correlation between ϵ_1 and ϵ_2 . The values of ρ_{12} by definition are restricted to be in

the -1 to 1 interval. Therefore, the choice for prior distribution for ρ_{12} can be *uniform* $(-1, 1)$ or a proper distribution based on reparameterization. Let ν denote the hyperbolic arc-tangent transformation of ρ_{12} , that is,

$$\nu = a \tanh(\rho_{12}), \quad (3.3)$$

and taking hyperbolic tangent transformation of ν gives us back $\rho_{12} = \tanh(\nu)$. Then parameter ν is asymptotically normal distributed with stabilized variance, $1/(N-3)$, where N is the sample size. We may also assume that $\nu \sim N(\nu^*, \sigma_\nu^2)$.

3.2. Bayesian Analysis via MCMC

For carrying out a Bayesian inference, the joint posterior distribution of the parameters of the ZIBOP model in (2.6) conditional on the data is obtained by combining the likelihood function given in (2.9) and the above-specified prior distributions via Bayes' theorem, as:

$$\begin{aligned} f(\Psi_b | \mathbf{x}, \mathbf{z}) &\propto \prod_{i=1}^N \prod_{(j,l)=(0,0)} [\Phi(-\mathbf{z}'_i \boldsymbol{\gamma}) + \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \Phi_2(-\mathbf{x}'_{1i} \beta_1, -\mathbf{x}'_{2i} \beta_2, \rho_{12})]^{d_{ijl}} \\ &\times \prod_{i=1}^N \prod_{(j,l) \neq (0,0)} \left[\begin{array}{l} \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) [\Phi_2(\alpha_{1j+1} - \mathbf{x}'_{1i} \beta_1, \alpha_{2l+1} - \mathbf{x}'_{2i} \beta_2; \rho_{12})] \\ -\Phi_2(\alpha_{1j} - \mathbf{x}'_{1i} \beta_1, \alpha_{2l} - \mathbf{x}'_{2i} \beta_2, \rho_{12}) \end{array} \right]^{d_{ijl}} \\ &\times f(\Psi_b), \end{aligned} \quad (3.4)$$

where $f(\Psi_b) \propto f(\boldsymbol{\beta})f(\boldsymbol{\gamma})f(\boldsymbol{\tau})f(\nu)$ and the parameter vector Ψ_b now consists of $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2)'$, $\boldsymbol{\gamma}$, $\boldsymbol{\tau} = (\boldsymbol{\tau}^1, \boldsymbol{\tau}^2)$, s and $\nu = a \tanh(\rho_{12})$. Here $f(\boldsymbol{\beta}) \propto |\Omega_\beta|^{-1/2} \exp\{-1/2(\boldsymbol{\beta} - \boldsymbol{\beta}^*)' \Omega_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}^*)\}$; $f(\boldsymbol{\gamma}) \propto |\Omega_\gamma|^{-1/2} \exp\{-1/2(\boldsymbol{\gamma} - \boldsymbol{\gamma}^*)' \Omega_\gamma^{-1} (\boldsymbol{\gamma} - \boldsymbol{\gamma}^*)\}$; $f(\boldsymbol{\tau}) \propto |\Omega_\tau|^{-1/2} \exp\{-1/2(\boldsymbol{\tau} - \boldsymbol{\tau}^*)' \Omega_\tau^{-1} (\boldsymbol{\tau} - \boldsymbol{\tau}^*)\}$; τ_{rj} are defined in (3.1), and α_{rj} are given via the inverse map (3.2).

Full conditional posterior distributions are required to implement the MCMC algorithm [19–22], and they are given as follows:

(1) fixed effects:

(a) zero state:

$$f(\boldsymbol{\gamma} | \mathbf{x}, \mathbf{z}, \Psi_{-\gamma}) \propto |\Omega_\gamma|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{\gamma} - \boldsymbol{\gamma}^*)' \Omega_\gamma^{-1} (\boldsymbol{\gamma} - \boldsymbol{\gamma}^*)\right\} \times f(\Psi_b | \mathbf{x}, \mathbf{z}); \quad (3.5)$$

(b) nonzero state:

$$f(\boldsymbol{\beta} | \mathbf{x}, \mathbf{z}, \Psi_{-\beta}) \propto |\Omega_\beta|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{\beta}^*)' \Omega_\beta^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}^*)\right\} \times f(\Psi_b | \mathbf{x}, \mathbf{z}). \quad (3.6)$$

(2) thresholds:

$$f(\boldsymbol{\tau} \mid \mathbf{x}, \mathbf{z}, \Psi_{-\tau}) \propto |\Omega_{\tau}|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{\tau} - \boldsymbol{\tau}^*)' \Omega_{\tau}^{-1} (\boldsymbol{\tau} - \boldsymbol{\tau}^*)\right\} \\ \times \prod_{i=1}^N \prod_{(j,l) \neq (0,0)} \left[\Phi(\mathbf{z}'_i \boldsymbol{\gamma}) [\Phi_2(\alpha_{1j+1} - \mathbf{x}'_{1i} \boldsymbol{\beta}_1, \alpha_{2l+1} - \mathbf{x}'_{2i} \boldsymbol{\beta}_2; \rho_{12}) - \Phi_2(\alpha_{1j} - \mathbf{x}'_{1i} \boldsymbol{\beta}_1, \alpha_{2l} - \mathbf{x}'_{2i} \boldsymbol{\beta}_2, \rho_{12})] \right]^{d_{ijl}}. \quad (3.7)$$

(3) bivariate correlation:

$$f(\nu \mid \mathbf{x}, \mathbf{z}, \Psi_{-\nu}) \propto \sigma_{\nu}^{-1} \exp\left\{-\frac{(\nu - \nu^*)^2}{2\sigma_{\nu}^2}\right\} \times f(\Psi_b \mid \mathbf{x}, \mathbf{z}). \quad (3.8)$$

The MCMC algorithm simulates direct draws from the above full conditionals iteratively until convergence is achieved. A single long chain [23, 24] is used for the proposed model. Geyer [23] argues that using a single longer chain is better than using a number of smaller chains with different initial values. We follow this strategy in our empirical analysis.

The Bayesian analysis of the univariate ZIOP follows as a special case of that of the ZIBOP presented above. In particular, the joint posterior distribution of the parameters of the ZIOP model in (2.22) conditional on the data is obtained by combining the likelihood function given in (2.23) and the above-specified prior distributions (with modified notations) via Bayes' theorem, as follows:

$$f(\Psi \mid \mathbf{x}, \mathbf{z}, \nu) \propto \prod_{i=1}^N \prod_{j=0} \prod_{l=0} [\Phi(-\mathbf{z}'_i \boldsymbol{\gamma}) + \Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \Phi(-\mathbf{x}'_i \boldsymbol{\beta})]^{d_{ijl}} \\ \times \prod_{i=1}^N \prod_{j>0} \prod_{l=0} [\Phi(\mathbf{z}'_i \boldsymbol{\gamma}) \{\Phi(\alpha_{j+1} - \mathbf{x}'_i \boldsymbol{\beta}) - \Phi(\alpha_j - \mathbf{x}'_i \boldsymbol{\beta})\}]^{d_{ijl}} \\ \times f(\boldsymbol{\beta}) f(\boldsymbol{\gamma}) f(\boldsymbol{\tau}), \quad (3.9)$$

where, using notation of Section 2.3 for $\boldsymbol{\beta}$ and the other parameter vectors, $f(\boldsymbol{\beta}) \propto |\Omega_{\beta}|^{-1/2} \exp\{-1/2(\boldsymbol{\beta} - \boldsymbol{\beta}^*)' \Omega_{\beta}^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}^*)\}$; $f(\boldsymbol{\gamma}) \propto |\Omega_{\gamma}|^{-1/2} \exp\{-1/2(\boldsymbol{\gamma} - \boldsymbol{\gamma}^*)' \Omega_{\gamma}^{-1} (\boldsymbol{\gamma} - \boldsymbol{\gamma}^*)\}$; $f(\boldsymbol{\tau}) \propto |\Omega_{\tau}|^{-1/2} \exp\{-1/2(\boldsymbol{\tau} - \boldsymbol{\tau}^*)' \Omega_{\tau}^{-1} (\boldsymbol{\tau} - \boldsymbol{\tau}^*)\}$, $\tau_2 = \log(\alpha_2)$ and $\tau_j = \log(\alpha_j - \alpha_{j-1})$, $j = 3, \dots, J$. Apart from dropping the bivariate correlation, we basically replace the bivariate normal cumulative distribution $\Phi_2(\cdot, \cdot; \rho_{12})$ by the univariate counterpart $\Phi(\cdot)$. Details are available upon request from the authors.

Apart from Bayesian estimation of the regression parameters, the posterior distributions of other quantities of interest can be obtained. These include posteriors for marginal effects and probabilities for nonparticipation, zero-consumption, and joint outcomes of interest. These will be considered in the application section. Next, we summarize model selection procedure.

The commonly used criteria for model selection like BIC and AIC are not appropriate for the multilevel models (in the presence of random effects), which complicates the counting

of the true number of free parameters. To overcome such a hurdle, Spiegelhalter et al. [25] proposed a Bayesian model comparison criterion, called Deviance Information Criterion (DIC). It is given as

$$\text{DIC} = \text{goodness-of-fit} + \text{penalty for complexity}, \quad (3.10)$$

where the “goodness-of-fit” is measured by the deviance for $\theta = (\beta, \gamma, \alpha)$

$$D(\theta) = -2 \log \mathcal{L}(\text{data} \mid \theta) \quad (3.11)$$

and complexity is measured by the “effective number of parameters”:

$$\begin{aligned} pD &= E_{\theta|y}[D(\theta)] - D(E_{\theta|y}[\theta]) \\ &= \bar{D} - D(\bar{\theta}); \end{aligned} \quad (3.12)$$

that is, posterior mean deviance minus deviance evaluated at the posterior mean of the parameters. The DIC is then defined analogously to AIC as

$$\begin{aligned} \text{DIC} &= D(\bar{\theta}) + 2pD \\ &= \bar{D} + pD. \end{aligned} \quad (3.13)$$

The idea here is that models with smaller DIC should be preferred to models with larger DIC. Models are penalized both by the value of \bar{D} , which favors a good fit, but also (similar to AIC and BIC) by the effective number of parameters pD . The advantage of DIC over other criteria, for Bayesian model selection, is that the DIC is easily calculated from the MCMC samples. In contrast, AIC and BIC require calculating the likelihood at its maximum values, which are not easily available from the MCMC simulation.

4. Application

4.1. Data

We consider an application to tobacco consumption behavior of individuals based on the 2001 household Tobacco Prevalence survey data from Bangladesh. The Survey was conducted in two administrative districts of paramount interest for tobacco production and consumption in the country. Data on daily consumption of smoking and chewing tobacco along with other socioeconomic and demographic characteristics and parental tobacco consumption habits were collected from respondents of 10 years of age and above. The data set has been used previously by Gurmu and Yunus [26] in the context of binary response models. Here we focus on a sample consisting of 6000 individual respondents aged between 10 and 101 years.

The ordinal outcomes $y_r = 0, 1, 2, 3$ used in this paper correspond roughly to zero, low, moderate, and high levels of tobacco consumption in the form of smoking (y_1) or chewing tobacco (y_2), respectively. The first dependent variable y_1 for an individual's daily

Table 1: Bivariate frequency distribution for intensity of tobacco use.

Smoke group	Chew group			Total (N)
	0	1	2	
0	3931	302	324	4557
1	265	12	6	283
2	526	35	37	598
3	498	29	35	562
Total (N)	5220	378	402	6000

cigarette smoking intensities assumes the following 4 choices: $y_1 = 0$ if nonsmoker, $y_1 = 1$ if smoking up to 7 cigarettes per day, $y_1 = 2$ if smoking between 8 and 12 cigarettes daily, and $y_1 = 3$ if smoking more than 12 cigarettes daily; likewise, for the intensity of chewing tobacco, $y_2 = 0$ if reported not chewing tobacco, $y_2 = 1$ if uses up to 7 chewing tobacco, and $y_2 = 2$ if consuming 7 or more chewing tobacco. The frequency distribution of cigarette smoking and tobacco chewing choices in Table 1 shows that nearly 66% of the respondents identify themselves as nonusers of tobacco. Our modeling strategy recognizes that these self-identified current nonusers of tobacco may include either individuals who never smoke or chew tobacco (genuine nonusers) or those who do, but not during the reporting period (potential users of tobacco). For example, potential tobacco users may include those who wrongly claim to be nonusers, previous tobacco users that are currently nonusers, and those most likely to use tobacco in the future due to changes in, say, prices and income. Table 1 also shows that 76% of the respondents are non-smokers and nearly 87% identify themselves as nonusers of tobacco for chewing. Given the extremely high proportion of observed zeros coupled with sparse cells on the right tail, we employ the zero-inflated bivariate ordered probit framework.

Table 2 gives definition of the explanatory variables as well as their means and standard deviations. The respondents are more likely to be Muslim, married, in early thirties, live in rural area, and have about 7 years of formal schooling. Although the country is mostly agrarian, only around 11% of the respondents were related to agricultural occupation in either doing agricultural operations on their own farms or working as agricultural wage laborers. About 12% of the respondents belong to the service occupation. The benchmark occupational group consists of business and other occupations. More than one-half of the fathers and slightly less than two-thirds of the mothers of the respondents currently use or have used tobacco products in the past.

Among the variables given in Table 2, the two indicators of parental use of tobacco products are included in \mathbf{z} as part of the participation equation (2.4). The rest of the variables are included in \mathbf{x}_r and \mathbf{z} of (2.1) and (2.4). To allow for nonlinear effects, age and education enter all three equations using a quadratic form. Due to lack of data on prices, our analysis is limited to the study of other economic and demographic determinants of participation, smoking, and chewing tobacco.

4.2. Results

We estimate the standard bivariate ordered probit (BOP) and zero-inflated bivariate ordered probit regression models for smoking and chewing tobacco and report estimation results for parameters, marginal effects, and choice probabilities, along measures of model selection.

Table 2: Definition and summary statistics for independent variables.

Name	Definition	Mean ^b	St. Dev.
Age ^a	Age in years	30.35	(14.9)
Education ^a	Number of years of formal schooling	6.83	(4.7)
Income	Monthly family income in 1000s of Taka	7.57	(10.3)
Male	= 1 if male	54.6	
Married	= 1 if married	57.2	
Muslim	= 1 if religion is Islam	78.4	
Father use	= 1 if father uses tobacco	54.0	
Mother use	= 1 if mother uses tobacco	65.1	
Region	= 1 if Rangpur resident, = 0 if Chittagong resident	49.7	
Urban	= 1 if urban resident	38.0	
Agriservice	= 1 if agriculture labor or service occupation	23.2	
Self-employed	= 1 if self-employed or household chores	30.7	
Student	= 1 if student	26.8	
Other	= 1 if business or other occupations (control)	19.3	

^a In implementation, we also include age squared and education squared.

^b The means for binary variables are in percentage.

Table 3: Goodness-of-fit statistics via DIC.

Model	Dbar	Dhat	pD	DIC
Bivariate ordered probit (BOP)	11417.1	11386.9	30.1	11447.2
Zero-inflated BOP	11301.1	11270.3	29.8	11329.9

Dbar: Posterior mean of deviance, Dhat: Deviance evaluated at the posterior mean of the parameters, pD: Dbar-Dhat, the effective number of parameters, and DIC: Deviance information criterion.

An earlier version of this paper reports results from the standard ordered probit model as well as the uncorrelated and correlated versions of the univariate zero-inflated ordered probit model for smoking tobacco. Convergence of the generated samples is assessed using standard tools (such as trace plots and ACF plots) within WinBUGS software. After initial 10,000 burn-in iterations, every 10th MCMC sample thereafter was retained from the next 100,000 iterations, obtaining 10,000 samples for subsequent posterior inference of the unknown parameters. The slowest convergence is observed for some parameters in the inflation submodel. By contrast, the autocorrelations functions for most of the marginal effects die out quickly relative to those for the associated parameters.

Table 3 reports the goodness-of-fit statistics for the standard bivariate ordered probit model and its zero-inflated version, ZIBOP. The ZIBOP regression model clearly dominates BOP in terms of DIC and its components; compare the DIC of 11330 for the former and 11447 for the latter model. Table 4 gives posterior means, standard deviations, medians, and the 95 percent credible intervals (in terms of the 2.5 and 97.5 percentiles) of the parameters and choice probabilities from ZIBOP model. For comparison, the corresponding results from BOP

Table 4: Posterior mean, standard deviation, and 95% credible intervals of parameters from zibop for smoking and chewing tobacco.

Variable	Mean	St. dev.	2.50%	Median	97.50%
Main (β_1, α_1): smoking (y_1):					
Age/10	0.672	0.119	0.444	0.685	0.894
Age square/100	-0.070	0.012	-0.093	-0.071	-0.046
Education	-0.071	0.014	-0.097	-0.071	-0.042
Education square	0.001	0.001	-0.002	0.001	0.003
Income	0.000	0.002	-0.005	0.000	0.005
Male	2.092	0.086	1.925	2.091	2.269
Married	0.213	0.070	0.074	0.213	0.353
Muslim	-0.053	0.052	-0.157	-0.053	0.049
Region	-0.007	0.048	-0.102	-0.007	0.086
Urban	-0.096	0.051	-0.198	-0.097	0.004
Agriservice	-0.234	0.056	-0.345	-0.233	-0.125
Self-employed	-0.246	0.087	-0.414	-0.247	-0.069
student	-0.476	0.137	-0.742	-0.478	-0.204
α_{12}	0.284	0.017	0.252	0.283	0.318
α_{13}	0.987	0.030	0.928	0.987	1.048
Main (β_2, α_2): chewing (y_2):					
Age/10	0.649	0.133	0.382	0.658	0.893
Age square/100	-0.046	0.013	-0.071	-0.046	-0.019
Education	-0.020	0.016	-0.052	-0.020	0.012
Education square	-0.002	0.001	-0.005	-0.002	0.000
Income	0.001	0.003	-0.004	0.002	0.007
Male	-0.479	0.081	-0.641	-0.479	-0.320
Married	-0.025	0.075	-0.171	-0.025	0.122
Muslim	-0.072	0.056	-0.181	-0.072	0.039
Region	0.417	0.051	0.317	0.418	0.517
Urban	-0.080	0.058	-0.194	-0.079	0.035
Agriservice	0.052	0.074	-0.096	0.052	0.194
Self-employed	0.127	0.092	-0.058	0.126	0.309
Student	-0.450	0.221	-0.887	-0.448	-0.023
α_{22}	0.484	0.023	0.439	0.484	0.531
Inflation (γ):					
Age/10	-0.012	2.044	-4.755	0.253	2.861
Age square/100	0.509	0.552	-0.197	0.398	1.812
Education	-0.218	0.115	-0.476	-0.204	-0.024
Education square	0.028	0.011	0.010	0.026	0.053
Income	0.006	0.022	-0.027	0.003	0.059
Male	0.239	0.827	-1.582	0.417	1.379
Married	2.306	4.478	-0.416	0.500	16.900
Muslim	-0.528	0.356	-1.331	-0.494	0.068
Mother	-0.170	0.267	-0.716	-0.164	0.345
Father	-0.119	0.330	-0.664	-0.160	0.605
Region	0.630	0.291	0.061	0.625	1.222

Table 4: Continued.

Variable	Mean	St. dev.	2.50%	Median	97.50%
Urban	0.040	0.357	-0.737	0.071	0.675
Agrservice	5.312	5.416	1.017	2.674	20.470
Self-employed	3.783	5.025	0.124	1.275	17.990
Sstudent	-0.344	0.411	-1.154	-0.339	0.466
ρ_{12}	-0.185	0.033	-0.249	-0.186	-0.119
Select probabilities:					
$P(y_1 = 0)$	0.760	0.004	0.752	0.760	0.768
$P(y_2 = 0)$	0.871	0.004	0.864	0.871	0.879
$P(y_1 = 0, y_2 = 0)$	0.662	0.005	0.652	0.662	0.671
$P(\text{zero-inflation})$	0.242	0.048	0.151	0.243	0.323

Results for the constant terms in the main and inflation parts have been suppressed for brevity.

are shown in Table 6 of the appendix. Both models predict significant negative correlation between the likelihood of smoking and chewing tobacco. The posterior estimates of the cut-off points are qualitatively similar across models. In what follows, we focus on discussion of results from the preferred ZIBOP model. The 95% credible interval for the correlation parameter ρ_{12} from the zero-inflated model is in the range -0.25 to -0.12 , indicating that smoking and chewing tobacco are generally substitutes. Results of selected predicted choice probabilities (bottom of Table 4) show that the ZIBOP regression model provides very good fit to the data. The posterior mean for the probability of (zero, zero)-inflation is about 24% while the 95% credible interval is $[0.15, 0.32]$, indicating that a substantial proportion of zeros may be attributed to nonparticipants. These results underscore the importance of modeling excess zeros in bivariate ordered probit models.

To facilitate interpretation of results, we report in Tables 5 and 7 the same set of posterior estimates for the marginal effects from ZIBOP and BOP models, respectively. Since age and education enter the three equations non-linearly, we report the total marginal effects coming from the linear and quadratic parts. We examine closely the marginal effects on the unconditional marginal probabilities at all levels of smoking and chewing tobacco ($y_1 = 0, 1, 2, 3$; $y_2 = 0, 1, 2$). The marginal effects reported in Table 5 show that the results for covariates are generally plausible. Age has a negative impact on probabilities of moderate and heavy use of tobacco. For heavy smokers, education has a significant negative impact on the probability of smoking cigarettes. An additional year of schooling on average decreases probability of smoking by about 6.9% for heavy smokers. Among participants, being male or married has positive impact on probability of smoking, while the effects for being Muslim, urban resident, and student are largely negative. Male respondents are more likely to smoke cigarettes while women respondents are more likely to use chewing tobacco with heavy intensity, a result which is in line with custom of the country [26].

Using (2.13), we decompose the marginal effect on probability of observing zero-consumption into two components: the effect on nonparticipation (zero inflation) and zero-consumption. For each explanatory variable, this decomposition is shown in Table 5 in the first three rows for smoking and in rows 1, 7, and 8 for chewing tobacco. For most variables, the effects on probabilities of nonparticipation and zero-consumption are on average opposite in sign, but this difference seems to diminish at the upper tail of the distribution. For example, looking at the posterior mean for age under smoking, getting older by one more year

Table 5: Posterior mean, standard deviation, and 95% credible intervals of marginal effects of covariates on probability of smoking and chewing tobacco (ZIBOP model).

Variable	Probability	Mean	St. dev.	2.50%	Median	97.50%	
Age	Nonparticipation	-0.0259	0.0129	-0.0556	-0.0236	-0.0078	
	Zero-consumption, y_1	0.0463	0.0102	0.0294	0.0453	0.0687	
	All zeros, $y_1 = 0$	0.0204	0.0059	0.0078	0.0213	0.0304	
	$y_1 = 1$	0.0058	0.0035	0.0009	0.0053	0.0138	
	$y_1 = 2$	-0.0014	0.0029	-0.0057	-0.0019	0.0055	
	$y_1 = 3$	-0.0690	0.0235	-0.1223	-0.0658	-0.0344	
	Zero-consumption, y_2	0.0403	0.0116	0.0195	0.0386	0.0675	
	All zeros, $y_2 = 0$	0.0145	0.0064	0.0018	0.0149	0.0264	
	$y_2 = 1$	-0.0034	0.0021	-0.0071	-0.0035	0.0008	
	$y_2 = 2$	-0.0019	0.0014	-0.0043	-0.0020	0.0011	
	Education	Nonparticipation	-0.2823	0.0768	-0.4260	-0.2837	-0.1252
		Zero-consumption, y_1	0.2447	0.0749	0.0917	0.2459	0.3851
All zeros, $y_1 = 0$		-0.0377	0.0241	-0.0853	-0.0374	0.0094	
$y_1 = 1$		0.0498	0.0141	0.0231	0.0494	0.0789	
$y_1 = 2$		0.0241	0.0102	0.0045	0.0239	0.0444	
$y_1 = 3$		-0.5557	0.1536	-0.8415	-0.5588	-0.2417	
Zero-consumption, y_2		0.3136	0.0772	0.1561	0.3159	0.4546	
All zeros, $y_2 = 0$		0.0313	0.0161	-0.0009	0.0315	0.0618	
$y_2 = 1$		-0.0134	0.0080	-0.0288	-0.0135	0.0027	
$y_2 = 2$		-0.0222	0.0119	-0.0455	-0.0221	0.0009	
Income		Nonparticipation	-0.0004	0.0015	-0.0038	-0.0002	0.0022
		Zero-consumption, y_1	0.0003	0.0014	-0.0022	0.0002	0.0035
	All zeros, $y_1 = 0$	-0.0001	0.0004	-0.0009	-0.0001	0.0008	
	$y_1 = 1$	0.0001	0.0003	-0.0005	0.0000	0.0008	
	$y_1 = 2$	0.0000	0.0002	-0.0003	0.0000	0.0004	
	$y_1 = 3$	-0.0007	0.0030	-0.0075	-0.0004	0.0044	
	Zero-consumption, y_2	0.0001	0.0016	-0.0025	0.0000	0.0036	
	All zeros, $y_2 = 0$	-0.0002	0.0005	-0.0011	-0.0002	0.0007	
	$y_2 = 1$	0.0001	0.0002	-0.0003	0.0001	0.0004	
	$y_2 = 2$	0.0001	0.0001	-0.0002	0.0001	0.0003	
	Male	Nonparticipation	-0.0254	0.0599	-0.1268	-0.0305	0.1012
		Zero-consumption, y_1	-0.3595	0.0611	-0.4900	-0.3540	-0.2565
All zeros, $y_1 = 0$		-0.3849	0.0116	-0.4078	-0.3849	-0.3618	
$y_1 = 1$		0.0630	0.0040	0.0555	0.0630	0.0711	
$y_1 = 2$		0.1560	0.0065	0.1435	0.1559	0.1689	
$y_1 = 3$		0.1659	0.0083	0.1503	0.1657	0.1829	
Zero-consumption, y_2		0.1012	0.0623	-0.0309	0.1064	0.2075	
All zeros, $y_2 = 0$		0.0758	0.0126	0.0511	0.0759	0.1004	
$y_2 = 1$		0.0501	0.0033	0.0438	0.0500	0.0567	
$y_2 = 2$		-0.1258	0.0112	-0.1478	-0.1258	-0.1040	
Married		Nonparticipation	-0.0680	0.0777	-0.2274	-0.0433	0.0346
		Zero-consumption, y_1	0.0200	0.0705	-0.0778	-0.0001	0.1692
	All zeros, $y_1 = 0$	-0.0480	0.0149	-0.0796	-0.0472	-0.0207	

Table 5: Continued.

Variable	Probability	Mean	St. dev.	2.50%	Median	97.50%
	$y_1 = 1$	0.0056	0.0035	0.0006	0.0047	0.0132
	$y_1 = 2$	0.0161	0.0061	0.0060	0.0154	0.0296
	$y_1 = 3$	0.0263	0.0073	0.0116	0.0264	0.0406
	Zero-consumption, y_2	0.0709	0.0791	-0.0371	0.0474	0.2349
	All zeros, $y_2 = 0$	0.0028	0.0119	-0.0200	0.0026	0.0269
	$y_2 = 1$	0.0628	0.0032	0.0566	0.0627	0.0693
	$y_2 = 2$	-0.0656	0.0115	-0.0888	-0.0654	-0.0434
Muslim	Nonparticipation	0.0393	0.0243	-0.0050	0.0384	0.0900
	Zero-consumption, y_1	-0.0239	0.0247	-0.0752	-0.0231	0.0216
	All zeros, $y_1 = 0$	0.0154	0.0090	-0.0016	0.0153	0.0334
	$y_1 = 1$	-0.0023	0.0011	-0.0044	-0.0022	-0.0002
	$y_1 = 2$	-0.0053	0.0027	-0.0106	-0.0053	-0.0001
	$y_1 = 3$	-0.0078	0.0060	-0.0200	-0.0077	0.0036
	Zero-consumption, y_2	-0.0260	0.0258	-0.0797	-0.0253	0.0222
	All zeros, $y_2 = 0$	0.0133	0.0092	-0.0046	0.0133	0.0315
	$y_2 = 1$	0.0613	0.0030	0.0554	0.0613	0.0674
	$y_2 = 2$	-0.0746	0.0091	-0.0926	-0.0746	-0.0569
Father use	Nonparticipation	0.0122	0.0187	-0.0251	0.0124	0.0487
	Zero-consumption, y_1	-0.0102	0.0158	-0.0411	-0.0104	0.0214
	All zeros, $y_1 = 0$	0.0020	0.0030	-0.0040	0.0019	0.0082
	$y_1 = 1$	-0.0005	0.0008	-0.0022	-0.0005	0.0011
	$y_1 = 2$	-0.0009	0.0014	-0.0037	-0.0009	0.0018
	$y_1 = 3$	-0.0005	0.0008	-0.0023	-0.0005	0.0011
	Zero-consumption, y_2	-0.0116	0.0179	-0.0464	-0.0118	0.0240
	All zeros, $y_2 = 0$	0.0006	0.0011	-0.0012	0.0003	0.0033
	$y_2 = 1$	-0.0003	0.0006	-0.0019	-0.0002	0.0007
	$y_2 = 2$	-0.0002	0.0005	-0.0014	-0.0001	0.0005
Mother use	Nonparticipation	0.0129	0.0257	-0.0343	0.0123	0.0634
	Zero-consumption, y_1	-0.0106	0.0215	-0.0527	-0.0103	0.0298
	All zeros, $y_1 = 0$	0.0024	0.0043	-0.0047	0.0020	0.0115
	$y_1 = 1$	-0.0006	0.0012	-0.0031	-0.0006	0.0014
	$y_1 = 2$	-0.0011	0.0019	-0.0051	-0.0009	0.0022
	$y_1 = 3$	-0.0007	0.0012	-0.0033	-0.0005	0.0012
	Zero-consumption, y_2	-0.0119	0.0242	-0.0587	-0.0118	0.0338
	All zeros, $y_2 = 0$	0.0010	0.0016	-0.0007	0.0004	0.0053
	$y_2 = 1$	-0.0006	0.0009	-0.0030	-0.0002	0.0005
	$y_2 = 2$	-0.0004	0.0007	-0.0023	-0.0001	0.0002
Region	Nonparticipation	-0.0480	0.0240	-0.0963	-0.0470	-0.0040
	Zero-consumption, y_1	0.0412	0.0237	-0.0039	0.0406	0.0889
	All zeros, $y_1 = 0$	-0.0068	0.0079	-0.0222	-0.0068	0.0086
	$y_1 = 1$	0.0021	0.0011	0.0001	0.0021	0.0046
	$y_1 = 2$	0.0033	0.0025	-0.0016	0.0033	0.0083
	$y_1 = 3$	0.0013	0.0052	-0.0087	0.0014	0.0114

Table 5: Continued.

Variable	Probability	Mean	St. dev.	2.50%	Median	97.50%
Urban	Zero-consumption, y_2	-0.0206	0.0252	-0.0672	-0.0217	0.0301
	All zeros, $y_2 = 0$	-0.0686	0.0078	-0.0840	-0.0686	-0.0533
	$y_2 = 1$	0.0756	0.0038	0.0682	0.0755	0.0832
	$y_2 = 2$	-0.0070	0.0070	-0.0207	-0.0070	0.0072
	Nonparticipation	-0.0062	0.0261	-0.0595	-0.0054	0.0428
	Zero-consumption, y_1	0.0217	0.0258	-0.0271	0.0211	0.0733
	All zeros, $y_1 = 0$	0.0155	0.0088	-0.0018	0.0155	0.0324
	$y_1 = 1$	-0.0007	0.0012	-0.0029	-0.0008	0.0017
	$y_1 = 2$	-0.0042	0.0028	-0.0096	-0.0042	0.0014
	$y_1 = 3$	-0.0106	0.0056	-0.0215	-0.0106	0.0006
Agriservice	Zero-consumption, y_2	0.0181	0.0275	-0.0337	0.0178	0.0739
	All zeros, $y_2 = 0$	0.0119	0.0090	-0.0062	0.0120	0.0295
	$y_2 = 1$	0.0597	0.0036	0.0528	0.0597	0.0668
	$y_2 = 2$	-0.0716	0.0075	-0.0864	-0.0717	-0.0566
	Nonparticipation	-0.1989	0.0521	-0.3092	-0.1960	-0.1062
	Zero-consumption, y_1	0.2102	0.0506	0.1202	0.2075	0.3161
	All zeros, $y_1 = 0$	0.0113	0.0098	-0.0084	0.0115	0.0297
	$y_1 = 1$	0.0058	0.0018	0.0026	0.0057	0.0097
	$y_1 = 2$	0.0023	0.0033	-0.0039	0.0021	0.0092
	$y_1 = 3$	-0.0194	0.0060	-0.0311	-0.0194	-0.0077
Self-employed	Zero-consumption, y_2	0.1838	0.0530	0.0871	0.1811	0.2940
	All zeros, $y_2 = 0$	-0.0151	0.0126	-0.0400	-0.0150	0.0091
	$y_2 = 1$	0.0680	0.0049	0.0588	0.0678	0.0782
	$y_2 = 2$	-0.0529	0.0096	-0.0716	-0.0530	-0.0338
	Nonparticipation	-0.1287	0.0693	-0.2542	-0.1191	-0.0122
	Zero-consumption, y_1	0.1590	0.0686	0.0431	0.1508	0.2845
	All zeros, $y_1 = 0$	0.0303	0.0166	-0.0034	0.0305	0.0627
	$y_1 = 1$	0.0005	0.0025	-0.0042	0.0005	0.0058
	$y_1 = 2$	-0.0075	0.0060	-0.0192	-0.0075	0.0043
	$y_1 = 3$	-0.0233	0.0089	-0.0398	-0.0237	-0.0046
Student	Zero-consumption, y_2	0.1034	0.0704	-0.0179	0.0941	0.2327
	All zeros, $y_2=0$	-0.0254	0.0147	-0.0546	-0.0251	0.0035
	$y_2 = 1$	0.0684	0.0047	0.0594	0.0681	0.0781
	$y_2 = 2$	-0.0430	0.0118	-0.0660	-0.0431	-0.0195
	Nonparticipation	0.0305	0.0357	-0.0312	0.0270	0.1076
	Zero-consumption, y_1	0.0548	0.0434	-0.0353	0.0564	0.1354
	All zeros, $y_1 = 0$	0.0852	0.0206	0.0437	0.0855	0.1247
	$y_1 = 1$	-0.0090	0.0027	-0.0149	-0.0089	-0.0041
	$y_1 = 2$	-0.0295	0.0079	-0.0455	-0.0294	-0.0143
	$y_1 = 3$	-0.0468	0.0106	-0.0657	-0.0475	-0.0244
	Zero-consumption, y_2	0.0284	0.0448	-0.0686	0.0313	0.1073
	All zeros, $y_2 = 0$	0.0588	0.0239	0.0065	0.0610	0.0995
	$y_2 = 1$	0.0390	0.0102	0.0207	0.0383	0.0604
	$y_2 = 2$	-0.0979	0.0142	-0.1211	-0.0994	-0.0659

Table 6: Posterior mean, standard deviation and 95% credible intervals of parameters from BOP for smoking and chewing tobacco.

Variable	Mean	St. Dev.	2.50%	Median	97.50%
Smoking (y_1) equation, (β_1, α_1)					
Age/10	1.029	0.095	0.828	1.030	1.199
Age square/100	-0.104	0.010	-0.123	-0.105	-0.082
Education	-0.078	0.014	-0.105	-0.078	-0.050
Education square	0.002	0.001	0.000	0.002	0.004
Income	0.000	0.002	-0.004	0.000	0.005
Male	2.066	0.091	1.888	2.067	2.245
Married	0.221	0.064	0.093	0.220	0.349
Muslim	-0.083	0.049	-0.177	-0.083	0.015
Region	0.041	0.043	-0.044	0.041	0.125
Urban	-0.091	0.048	-0.186	-0.091	0.002
Agriservice	-0.121	0.050	-0.219	-0.122	-0.023
Self-employed	-0.149	0.087	-0.318	-0.150	0.021
Sstudent	-0.720	0.093	-0.905	-0.719	-0.538
α_{12}	0.270	0.015	0.241	0.270	0.300
α_{13}	0.956	0.028	0.901	0.956	1.012
Chewing (y_2) equation, (β_2, α_2)					
Age/10	0.797	0.091	0.609	0.801	0.977
Age square/100	-0.059	0.010	-0.079	-0.059	-0.039
Education	-0.023	0.016	-0.055	-0.023	0.008
Education square	-0.002	0.001	-0.005	-0.002	0.001
Income	0.002	0.003	-0.004	0.002	0.007
Male	-0.441	0.074	-0.586	-0.441	-0.295
Married	-0.010	0.073	-0.153	-0.011	0.134
Muslim	-0.077	0.056	-0.187	-0.077	0.033
Region	0.430	0.049	0.334	0.430	0.528
Urban	-0.082	0.056	-0.193	-0.081	0.026
Agriservice	0.078	0.073	-0.067	0.078	0.222
Self employed	0.177	0.087	0.010	0.176	0.351
Student	-0.715	0.177	-1.070	-0.710	-0.378
α_{22}	0.480	0.023	0.436	0.480	0.525
ρ_{12}	-0.178	0.034	-0.244	-0.179	-0.111

Each equation includes father use and mother use variables as well as a constant term.

decreases probability of nonparticipation by about 2.6% but increases probability of zero-consumption by 4.6%, implying a net increase of 2.0% in predicted probability of observing zero. The effect of age in the case of chewing tobacco is qualitatively similar, negative effect on genuine nonusers and positive effect on potential tobacco users, with the latter dominating in the overall effect.

Income has opposite effects on probability of nonparticipation and zero-consumption, predicting on average that tobacco is an inferior good for nonparticipants and a normal good for participants. However, the 95% credible interval contains zero, suggesting that the

Table 7: Posterior mean, standard deviation, and 95% credible intervals of marginal effects of covariates on probability of smoking and chewing tobacco (BOP model).

Variable	Probability	Mean	St. dev.	2.50%	Median	97.50%
Age	All zeros, $y_1 = 0$	0.0368	0.0038	0.0288	0.0369	0.0438
	$y_1 = 1$	-0.0004	0.0003	-0.0009	-0.0004	0.0002
	$y_1 = 2$	-0.0073	0.0010	-0.0093	-0.0073	-0.0054
	$y_1 = 3$	-0.0292	0.0029	-0.0345	-0.0292	-0.0230
	All zeros, $y_2 = 0$	0.0213	0.0043	0.0125	0.0215	0.0298
	$y_2 = 1$	-0.0056	0.0013	-0.0082	-0.0057	-0.0031
	$y_2 = 2$	-0.0030	0.0008	-0.0047	-0.0030	-0.0014
Education	All zeros, $y_1 = 0$	-0.0342	0.0236	-0.0803	-0.0340	0.0126
	$y_1 = 1$	0.0038	0.0025	-0.0011	0.0039	0.0086
	$y_1 = 2$	0.0130	0.0084	-0.0039	0.0129	0.0293
	$y_1 = 3$	0.0174	0.0128	-0.0076	0.0172	0.0428
	All zeros, $y_2 = 0$	0.0322	0.0156	0.0002	0.0326	0.0616
	$y_2 = 1$	-0.0150	0.0077	-0.0296	-0.0151	0.0009
	$y_2 = 2$	-0.0201	0.0113	-0.0418	-0.0203	0.0024
Income	All zeros, $y_1 = 0$	-0.0001	0.0004	-0.0009	-0.0001	0.0007
	$y_1 = 1$	0.0000	0.0000	-0.0001	0.0000	0.0001
	$y_1 = 2$	0.0000	0.0001	-0.0002	0.0000	0.0002
	$y_1 = 3$	0.0000	0.0002	-0.0004	0.0000	0.0005
	All zeros, $y_2 = 0$	-0.0003	0.0005	-0.0012	-0.0003	0.0007
	$y_2 = 1$	0.0001	0.0002	-0.0002	0.0001	0.0004
	$y_2 = 2$	0.0001	0.0002	-0.0002	0.0001	0.0004
Male	All zeros, $y_1 = 0$	-0.3824	0.0121	-0.4064	-0.3826	-0.3586
	$y_1 = 1$	0.0641	0.0040	0.0567	0.0640	0.0722
	$y_1 = 2$	0.1540	0.0065	0.1416	0.1540	0.1667
	$y_1 = 3$	0.1643	0.0083	0.1487	0.1641	0.1807
	All zeros, $y_2 = 0$	0.0721	0.0123	0.0481	0.0721	0.0962
	$y_2 = 1$	0.0500	0.0032	0.0438	0.0500	0.0565
	$y_2 = 2$	-0.1222	0.0108	-0.1430	-0.1220	-0.1014
Married	All zeros, $y_1 = 0$	-0.0416	0.0124	-0.0666	-0.0415	-0.0174
	$y_1 = 1$	0.0039	0.0013	0.0015	0.0038	0.0067
	$y_1 = 2$	0.0131	0.0042	0.0053	0.0130	0.0218
	$y_1 = 3$	0.0246	0.0070	0.0106	0.0247	0.0385
	All zeros, $y_2 = 0$	0.0018	0.0118	-0.0207	0.0018	0.0254
	$y_2 = 1$	0.0622	0.0031	0.0563	0.0622	0.0685
	$y_2 = 2$	-0.0640	0.0114	-0.0873	-0.0640	-0.0420
Muslim	All zeros, $y_1 = 0$	0.0154	0.0092	-0.0029	0.0154	0.0331
	$y_1 = 1$	-0.0013	0.0008	-0.0028	-0.0013	0.0002
	$y_1 = 2$	-0.0044	0.0026	-0.0093	-0.0044	0.0008
	$y_1 = 3$	-0.0097	0.0058	-0.0211	-0.0097	0.0018
	All zeros, $y_2 = 0$	0.0126	0.0093	-0.0051	0.0125	0.0313
	$y_2 = 1$	0.0613	0.0031	0.0555	0.0613	0.0675
	$y_2 = 2$	-0.0739	0.0091	-0.0922	-0.0739	-0.0562
Father use	All zeros, $y_1 = 0$	0.7604	0.0042	0.7521	0.7604	0.7684

Table 7: Continued.

Variable	Probability	Mean	St. dev.	2.50%	Median	97.50%
	$y_1 = 1$	0.0477	0.0027	0.0426	0.0477	0.0531
	$y_1 = 2$	0.0982	0.0035	0.0915	0.0982	0.1051
	$y_1 = 3$	0.0937	0.0032	0.0874	0.0936	0.1000
	All zeros, $y_2 = 0$	0.8713	0.0039	0.8635	0.8713	0.8789
	$y_2 = 1$	0.0623	0.0030	0.0566	0.0623	0.0684
	$y_2 = 2$	0.0664	0.0030	0.0607	0.0664	0.0724
Mother use	All zeros, $y_1 = 0$	0.7604	0.0042	0.7521	0.7604	0.7684
	$y_1 = 1$	0.0477	0.0027	0.0426	0.0477	0.0531
	$y_1 = 2$	0.0982	0.0035	0.0915	0.0982	0.1051
	$y_1 = 3$	0.0937	0.0032	0.0874	0.0936	0.1000
	All zeros, $y_2 = 0$	0.8713	0.0039	0.8635	0.8713	0.8789
	$y_2 = 1$	0.0623	0.0030	0.0566	0.0623	0.0684
	$y_2 = 2$	0.0664	0.0030	0.0607	0.0664	0.0724
Region	All zeros, $y_1 = 0$	-0.0075	0.0079	-0.0229	-0.0075	0.0080
	$y_1 = 1$	0.0006	0.0007	-0.0007	0.0006	0.0020
	$y_1 = 2$	0.0022	0.0023	-0.0023	0.0022	0.0067
	$y_1 = 3$	0.0047	0.0049	-0.0050	0.0047	0.0144
	All zeros, $y_2 = 0$	-0.0691	0.0078	-0.0846	-0.0691	-0.0539
	$y_2 = 1$	0.0756	0.0038	0.0684	0.0755	0.0832
	$y_2 = 2$	-0.0065	0.0070	-0.0200	-0.0065	0.0072
Urban	All zeros, $y_1 = 0$	0.0167	0.0087	-0.0003	0.0167	0.0339
	$y_1 = 1$	-0.0014	0.0008	-0.0030	-0.0014	0.0000
	$y_1 = 2$	-0.0049	0.0026	-0.0100	-0.0049	0.0001
	$y_1 = 3$	-0.0104	0.0054	-0.0210	-0.0104	0.0002
	All zeros, $y_2 = 0$	0.0130	0.0088	-0.0041	0.0129	0.0303
	$y_2 = 1$	0.0592	0.0036	0.0524	0.0591	0.0664
	$y_2 = 2$	-0.0721	0.0074	-0.0866	-0.0721	-0.0576
Agriservice	All zeros, $y_1 = 0$	0.0218	0.0088	0.0043	0.0219	0.0390
	$y_1 = 1$	-0.0018	0.0007	-0.0032	-0.0018	-0.0004
	$y_1 = 2$	-0.0062	0.0025	-0.0110	-0.0062	-0.0013
	$y_1 = 3$	-0.0138	0.0057	-0.0250	-0.0139	-0.0027
	All zeros, $y_2 = 0$	-0.0127	0.0119	-0.0366	-0.0127	0.0106
	$y_2 = 1$	0.0656	0.0043	0.0572	0.0655	0.0742
	$y_2 = 2$	-0.0528	0.0094	-0.0711	-0.0529	-0.0338
Self employed	All zeros, $y_1 = 0$	0.0277	0.0162	-0.0039	0.0277	0.0592
	$y_1 = 1$	-0.0028	0.0018	-0.0065	-0.0027	0.0003
	$y_1 = 2$	-0.0087	0.0053	-0.0194	-0.0086	0.0012
	$y_1 = 3$	-0.0163	0.0093	-0.0335	-0.0165	0.0025
	All zeros, $y_2 = 0$	-0.0290	0.0144	-0.0578	-0.0286	-0.0017
	$y_2 = 1$	0.0686	0.0046	0.0600	0.0685	0.0779
	$y_2 = 2$	-0.0396	0.0116	-0.0617	-0.0398	-0.0162
Student	All zeros, $y_1 = 0$	0.1287	0.0155	0.0980	0.1286	0.1588
	$y_1 = 1$	-0.0173	0.0030	-0.0235	-0.0171	-0.0118

Table 7: Continued.

Variable	Probability	Mean	St. dev.	2.50%	Median	97.50%
	$y_1 = 2$	-0.0475	0.0069	-0.0614	-0.0475	-0.0343
	$y_1 = 3$	-0.0639	0.0063	-0.0758	-0.0640	-0.0510
	All zeros, $y_2 = 0$	0.0855	0.0151	0.0531	0.0866	0.1117
	$y_2 = 1$	0.0278	0.0071	0.0154	0.0274	0.0428
	$y_2 = 2$	-0.1133	0.0088	-0.1288	-0.1140	-0.0944

effect of income is weak. Generally, the opposing effects on probabilities of nonparticipation and zeroconsumption would have repercussions on both the magnitude and the statistical significance of the full effect of observing zero-consumption. Similar considerations apply to positive levels of consumption since the marginal effect on probability of observing consumption level j ($j = 1, 2, \dots$) can be decomposed into the marginal effects on (i) participation $P(s_i = 1)$ and (ii) levels of consumption conditional on participation, $P(y_{ri} = j \mid s_i = 1)$. These results show that policy recommendations that ignore excess zeros may lead to misleading conclusions.

5. Conclusion

In this paper we analyze the zero-inflated bivariate ordered probit model in a Bayesian framework. The underlying model arises as a mixture of a point mass distribution at $(0, 0)$ for nonparticipants and the bivariate ordered probit distribution for participants. The Bayesian analysis is carried out using MCMC techniques to approximate the posterior distribution of the parameters. Using household tobacco survey data with substantial proportion of zeros, we analyze the socioeconomic determinants of individual problem of smoking and chewing tobacco. In our illustration, we find evidence that accounting for excess zeros provides very good fit to the data. The use of a model that ignores zero-inflation masks differential effects of covariates on nonusers and users at various levels of consumption, including zeros. The Bayesian approach to modeling excess zeros provides computational flexibility of generalizing to multivariate ordered response models as well as ordinal panel data models.

The proposed zero-inflated bivariate model is particularly useful when most of the bivariate ordered outcomes are zero ($y_1 = 0, y_2 = 0$). In addition to allowing for inflation in the double-zero state, our approach can be extended to allow for zero inflation in each component. If needed, other states in an ordered regression model may be inflated as well. These extensions need to be justified empirically on a case-by-case basis and are beyond the scope of this paper.

Appendices

A.

For more details see Tables 6 and 7.

B.

WinBUGS Code for Fitting the Proposed Models (see Algorithm 1).

```

#Variable names in the tobacco data are given in y[,1:21]
model {
  for(h in 1:N) {
    ## participation model ###
    cov2[h]<- gama[1]+gama[2]*y[h,6]+gama[3]*y[h,7]+gama[4]*y[h,8]+gama[5]*y[h,9]
    cov3[h]<- gama[6]*y[h,10]+gama[7]*y[h,11]+gama[8]*y[h,12]+gama[9]*y[h,13]
    cov4[h]<- gama[10]*y[h,14] +gama[11]*y[h,15]+gama[12]*y[h,16]
    cov5[h]<- gama[13]*y[h,17]+gama[14]*y[h,18]+gama[15]*y[h,19] +gama[16]*y[h,20]
    cov[h] <- cov2[h]+cov3[h]+cov4[h] +cov5[h]
    pi[h] <- phi(-cov[h])
    ph.5a[h] <- phi(cov[h])
    ### consumption model #####
    #Smoking #
    covar2[h]<-beta[1]+beta[2]*y[h,6]+beta[3]*y[h,7]+beta[4]*y[h,8]+beta[5]*y[h,9]
    covar3[h]<-beta[6]*y[h,10]+beta[7]*y[h,11]+beta[8]*y[h,12]+beta[9]*y[h,13]
    +beta2[1]*y[h,16]
    covar4[h]<-beta2[2]*y[h,17]+beta2[3]*y[h,18]+beta2[4]*y[h,19]+beta2[5]*y[h,20]
    covar[h] <- covar2[h]+covar3[h]+covar4[h]
    #Chewing #
    covar2.chew[h]<-beta.chew[1]+beta.chew[2]*y[h,6]+beta.chew[3]*y[h,7]
    +beta.chew[4]*y[h,8]
    covar3.chew[h] <- beta.chew[5]*y[h,9]+beta.chew[6]*y[h,10]+beta.chew[7]*y[h,11]
    covar4.chew[h] <- beta.chew[8]*y[h,12]+beta.chew[9]*y[h,13]
    covar5.chew[h] <- beta2.chew[1]*y[h,16]+beta2.chew[2]*y[h,17]
    +beta2.chew[3]*y[h,18]
    covar6.chew[h] <- beta2.chew[4]*y[h,19]+ beta2.chew[5]*y[h,20]
    covar.chew2[h] <-covar2.chew[h]+covar3.chew[h]+covar4.chew[h]
    covar.chew3[h] <-covar5.chew[h]+covar6.chew[h]+covar7.chew[h]
    covar.chew[h] <- covar.chew2[h]+covar.chew3[h]
    # Cumulative probability of < j
    ph.2[h] <- (1/sqrt(2*3.14159))*exp(-0.5*covar[h]*covar[h])
    ph.3[h] <- (1/sqrt(2*3.14159))*exp(-0.5*(alpha[1]-covar[h])*(alpha[1]-covar[h]))
    ph.4[h] <- (1/sqrt(2*3.14159))*exp(-0.5*(alpha[2]-covar[h])*(alpha[2]-covar[h]))
    ph.5b[h]<- phi(-covar[h])
    #joint CDF probability for ((y1,y2)=(0,0))
    nu.0[h] <- -rho12*ph.2[h]/phi(-covar[h])
    s2.0[h] <-1+rho12*(-covar[h])*nu.0[h]-nu.0[h]*nu.0[h]
    Q.00[h] <-ph.5b[h]*phi((-covar.chew[h]-nu.0[h])/sqrt(s2.0[h]))

    #joint CDF probability for ((y1,y2)=(0,1))
    Q.01[h] <-ph.5b[h]*phi((alpha.chew-covar.chew[h]-nu.0[h])/sqrt(s2.0[h]))
    .....
    #joint CDF probability for ((y1,y2)=(3,2))
    Q.32[h] <-1

    mu[h,1] <- pi[h] + ph.5a[h]*Q.00[h] #p[0,0]
    mu[h,2] <- ph.5a[h]*(Q.01[h]-Q.00[h]) #p[0,1]
    mu[h,3] <- ph.5a[h]*(Q.02[h]-Q.01[h]) #p[0,2]
    mu[h,4] <- ph.5a[h]*(Q.10[h]-Q.00[h]) #p[1,0]
    mu[h,5] <- ph.5a[h]*(Q.11[h]-Q.10[h]-Q.01[h]+Q.00[h]) #p[1,1]
    mu[h,6] <- ph.5a[h]*(Q.12[h]-Q.11[h]-Q.02[h]+Q.01[h]) #p[1,2]
    mu[h,7] <- ph.5a[h]*(Q.20[h]-Q.10[h]) #p[2,0]
    mu[h,8] <- ph.5a[h]*(Q.21[h]-Q.20[h]-Q.11[h]+Q.10[h]) #p[2,1]
    mu[h,9] <- ph.5a[h]*(Q.22[h]-Q.21[h]-Q.12[h]+Q.11[h]) #p[2,2]
    mu[h,10] <- ph.5a[h]*(Q.30[h]-Q.20[h]) #p[3,0]
    mu[h,11] <- ph.5a[h]*(Q.31[h]-Q.30[h]-Q.21[h]+Q.20[h]) #p[3,1]
    mu[h,12] <- ph.5a[h]*(Q.32[h]-Q.31[h]-Q.22[h]+Q.21[h]) #p[3,2]
    y[h,21] ~dcat(mu[h,1:12])}

```

Algorithm 1

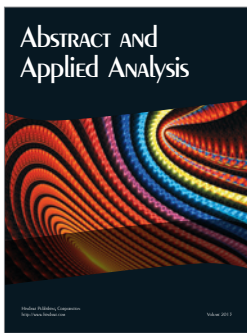
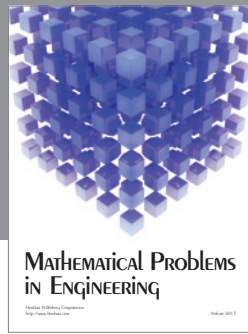
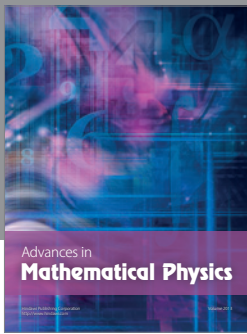
Acknowledgments

The authors thank Alfonso Flores-Lagunes, the editor, two anonymous referees and seminar participants at the Conference on Bayesian Inference in Econometrics and Statistics, the Joint Statistical Meetings, the Southern Economics Association Conference, and Syracuse University for useful comments. Mohammad Yunus graciously provided the data used in this paper.

References

- [1] C. Calhoun, "Estimating the distribution of desired family size and excess fertility," *The Journal of Human Resources*, vol. 24, pp. 709–724, 1989.
- [2] C. Calhoun, "Desired and excess fertility in Europe and the United States: indirect estimates from World Fertility Survey data," *European Journal of Population*, vol. 7, no. 1, pp. 29–57, 1991.
- [3] A. A. Weiss, "A bivariate ordered probit model with truncation: helmet use and motor cycle injuries," *Applied Statistics*, vol. 42, pp. 487–499, 1993.
- [4] J. S. Butler and P. Chatterjee, "Tests of the specification of univariate and bivariate ordered probit," *The Review of Economics and Statistics*, vol. 79, pp. 343–347, 1997.
- [5] A. Biswas and K. Das, "A bayesian analysis of bivariate ordinal data. Wisconsin epidemiologic study of diabetic retinopathy revisited," *Statistics in Medicine*, vol. 21, no. 4, pp. 549–559, 2002.
- [6] Z. Sajaia, "Maximum likelihood estimation of a bivariate ordered probit model: implementation and Monte Carlo simulations," Tech. Rep., The World Bank, Working Paper, 2008.
- [7] M. K. Munkin and P. K. Trivedi, "Bayesian analysis of the ordered probit model with endogenous selection," *Journal of Econometrics*, vol. 143, no. 2, pp. 334–348, 2008.
- [8] P. E. Stephan, S. Gurmu, A. J. Sumell, and G. C. Black, "Who's patenting in the university? Evidence from the Survey of Doctorate Recipients," *Economics of Innovation and New Technology*, vol. 16, pp. 71–99, 2007.
- [9] D. Lambert, "Zero-inflated poisson regression, with an application to defects in manufacturing," *Technometrics*, vol. 31, no. 1, pp. 1–14, 1992.
- [10] S. Gurmu and P. K. Trivedi, "Excess zeros in count models for recreational trips," *Journal of Business and Economic Statistics*, vol. 14, no. 4, pp. 469–477, 1996.
- [11] J. Mullahy, "Heterogeneity, excess zeros, and the structure of count data models," *Journal of Applied Econometrics*, vol. 12, no. 3, pp. 337–350, 1997.
- [12] S. Gurmu and J. Elder, "A bivariate zero-inflated count data regression model with unrestricted correlation," *Economics Letters*, vol. 100, pp. 245–248, 2008.
- [13] D. B. Hall, "Zero-inflated Poisson and binomial regression with random effects: a case study," *Biometrics*, vol. 56, no. 4, pp. 1030–1039, 2000.
- [14] G. A. Dagne, "Hierarchical Bayesian analysis of correlated zero-inflated count data," *Biometrical Journal*, vol. 46, no. 6, pp. 653–663, 2004.
- [15] M. N. Harris and X. Zhao, "A zero-inflated ordered probit model, with an application to modelling tobacco consumption," *Journal of Econometrics*, vol. 141, no. 2, pp. 1073–1099, 2007.
- [16] D. B. Hall and J. Shen, "Robust estimation for zero-inflated Poisson regression," *Scandinavian Journal of Statistics*, vol. 37, no. 2, pp. 237–252, 2010.
- [17] L. Liu, R. L. Strawderman, M. E. Cowen, and T. S. Ya-C. T. Shih, "A flexible two-part random effects model for correlated medical costs," *Journal of Health Economics*, vol. 29, no. 1, pp. 110–123, 2010.
- [18] S. Chib and B. H. Hamilton, "Bayesian analysis of cross-section and clustered data treatment models," *Journal of Econometrics*, vol. 97, no. 1, pp. 25–50, 2000.
- [19] D. Gamerman, *Markov Chain Monte Carlo*, Chapman & Hall, London, UK, 1997.
- [20] A. Gelman, J. Carlin, H. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall, London, UK, 1995.
- [21] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London, UK, 1996.
- [22] L. Tierney, "Markov chains for exploring posterior distributions (with discussion)," *The Annals of Statistics*, vol. 22, no. 4, pp. 1701–1762, 1994.
- [23] C. J. Geyer, "Practical markov chain Monte Carlo (with discussion)," *Statistical Science*, vol. 7, pp. 473–511, 1992.

- [24] A. E. Raftery and S. Lewis, "Comment: one long run with diagnostics: implementation strategies for Markov Chain Monte Carlo," *Statistical Science*, vol. 7, pp. 493–549, 1992.
- [25] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde, "Bayesian measures of model complexity and fit (with discussion)," *Journal of the Royal Statistical Society: Series B*, vol. 64, no. 4, pp. 583–639, 2002.
- [26] S. Gurmú and M. Yunus, "Tobacco chewing, smoking and health knowledge: evidence from Bangladesh," *Economics Bulletin*, vol. 9, no. 12, pp. 1–9, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

